



Expected utility theory with probability grids and preference formation

Mamoru Kaneko¹ 

Received: 30 May 2018 / Accepted: 21 August 2019 / Published online: 28 August 2019
© The Author(s) 2019

Abstract

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

Keywords Expected utility · Measurement of utility · Bounded rationality · Probability grids · Cognitive bound · Incomparabilities

JEL Classification C72 · C79 · C91

The author thanks J. J. Kline, P. Wakker, M. Lewandowski, R. Ishikawa, S. Shiba, M. Cohen, O. Shulte, and Y. Rebillé for helpful comments on earlier versions of this paper. In particular, comments given by the two referees improved the paper significantly and are greatly appreciated. The author is supported by Grant-in-Aids for Scientific Research Nos. 19K21704 and 17H02258, Ministry of Education, Science and Culture.

✉ Mamoru Kaneko
mkanekoepi@waseda.jp

¹ Faculty of Political Science and Economics, Waseda University, Shinjuku, Tokyo 169-0850, Japan

1 Introduction

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

The concepts of *probability grids* and *cognitive bounds* are introduced based on the idea of “bounded rationality.” The idea can be interpreted in many ways such as bounded logical inference, bounded perception ability, though Simon’s (1956) original concept meant a relaxation of utility maximization. The mathematical components involved in EU theory are classified into two types; object components used by the decision maker and meta-components used by the outside analyst and possibly by the decision maker himself. The former are primary targets in EU theory, and the latter such as highly complex rational as well as irrational probabilities are added for analytic convenience. A free use of the latter leads to a critique that the theory presumes “super rationality” (Simon 1983).

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

Turing (1937) in his attempt to define computable numbers faced a similar situation: “...The differences from our point of view between the single and compound symbols is that the compound symbols, if they are too lengthy, cannot be observed at one glance. This is in accordance with experience. We cannot tell at a glance whether 0.9999999999999999 and 0.9999999999999999 (the underlined parts by the author) are the same (1937, p. 250).”¹ In contrast, it is easy for us to distinguish between 0.999 and 0.99. Turing’s theory tried to abstract calculation in the human mind, but a built machine has no such problem since it reads each primitive symbol one by one. This is, so far, due to the difference in human and machine cognition. A cognitive bound ρ is a bound on such distinguishability and also on how deep the decision maker cares about those probabilities.

The set of probability grids up to depth k is given as $\Pi_k = \{\frac{0}{10^k}, \frac{1}{10^k}, \dots, \frac{10^k}{10^k}\}$. The decision maker thinks about his preferences with Π_k from a small k to a larger k up to bound ρ ; for example, when $\rho = 2$, Π_0 , Π_1 , and Π_2 are only allowed. This is a

¹ The author thanks Oliver Schulte for mentioning this quotation to me.

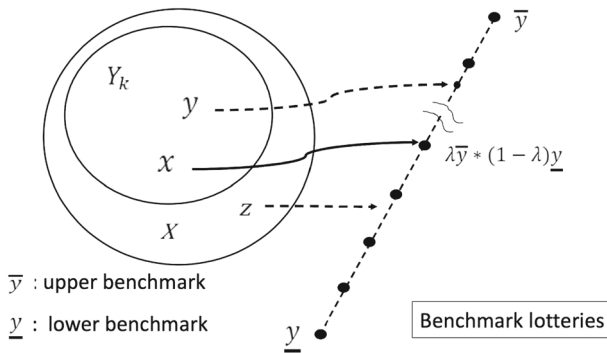


Fig. 1 Step B with the benchmark scale

constructive approach from the viewpoint of the decision maker in the sense that he finds/forms his own preferences.^{2,3}

We turn our attention to the development of our constructive EU theory. Construction needs a basis; we take a hint from Von Neumann and Morgenstern (1944), Section 3.3.2, p.17. They mentioned a separation of their argument into the following two steps, though this was not reflected in their mathematical development:

- Step B: measurements of utilities from pure alternatives in terms of probabilities;
- Step E: extensions of these measurements to lotteries involving more risks.

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

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

² This sounds similar to “constructive decision theory” in Shafer (1986), Shafer (2016) and Blume et al. (2013). These authors study Savage’s (1954) subjective utility/probability theory so as to introduce certain constructive features for decision making. Our theory is constructive more explicitly with the introduction of probability grids and a cognitive bound. The chief difference is that we formulate how a decision maker finds/forms his own preferences, while they add new constructs like “goals” or “frames” that shape the choices of the decision maker.

³ Our concept of probability grids may be interpreted as “imprecise probabilities/similarity” (cf. Augustin et al. 2014; Rubinstein 1988). Imprecision/similarity is defined as an attribute of a probability/a set of probabilities, allowing all real number probabilities. In our approach, however, probability grids in Π_k are exact; the restriction of probabilities to Π_k expresses imprecision in cognitive acts taken by the decision maker.

⁴ Our theory may be regarded as dual to that in terms of certainty equivalent of a lottery (cf. Kontek and Lewandowski 2018 and its references). In our method, the set of benchmark lotteries forms a base scale, while the set of monetary amounts is the base scale in the latter (see Section 4.2 in Kontek and Lewandowski 2018).

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

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

- (i) probability grids used for measurement of a pure alternative in Step B;
- (ii) probability coefficients to pure alternatives.

By these, relevant cognitive depths of lotteries become more complex, especially with a finite cognitive bound; this leads to incomparabilities in preferences and a violation of the EU hypothesis. This is central in our development and is closely related to the issue of “bounded rationality”.

Let us illustrate (i) and (ii) via an example, which is a lottery for choice in the Allais paradox in Sect. 8. Consider one example with the upper and lower benchmarks \bar{y} , \underline{y} , and the third pure alternative y with strict preferences $\bar{y} \succ y \succ \underline{y}$. In Step B, the decision maker looks for a probability λ so that y is indifferent to a lottery $[\bar{y}, \lambda; \underline{y}] = \lambda\bar{y} * (1 - \lambda)\underline{y}$ with probability λ for \bar{y} and $1 - \lambda$ for \underline{y} ; this indifference is denoted by

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

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

We have the other source of cognitive depths as mentioned in (ii). Consider lottery $d = \frac{25}{10^2}y * \frac{75}{10^2}\underline{y}$, which includes the third pure alternative y . The independence condition of the classical EU theory dictates that because of (1), $[\bar{y}, \frac{83}{10^2}; \underline{y}]$ is substituted for y in d , and d is reduced to:

$$d = \frac{25}{10^2}y * \frac{75}{10^2}\underline{y} \sim \frac{25}{10^2} \left[\bar{y}, \frac{83}{10^2}; \underline{y} \right] * \frac{75}{10^2}\underline{y} = \frac{2075}{10^4}\bar{y} * \frac{7925}{10^4}\underline{y}, \tag{2}$$

where $\frac{2075}{10^4} = \frac{25}{10^2} \times \frac{75}{10^2}$ and $\frac{7925}{10^4} = 1 - \frac{2075}{10^4}$. Thus, y is evaluated as being indifferent to $[\bar{y}, \frac{83}{10^2}; \underline{y}]$ in Step B, but y also has a probability coefficient $\frac{25}{10^2}$ in d , which is taken into account in Step E. These steps lead to probability $\frac{2075}{10^4}$, which is much more precise than either of $\frac{83}{10^2}$ and $\frac{25}{10^2}$.

As indicated in (i) and (ii), lottery $d = \frac{25}{10^2}y * \frac{75}{10^2}\underline{y}$ has two types of cognitive depths; one is simply a probability coefficient $\frac{25}{10^2}$ and the other is $\lambda_y = \frac{83}{10^2}$ from (1). Although d itself is expressed as a lottery of depth 2, the total depths including these two are 4, which is beyond the cognitive bound $\rho = 2$. One point is that the resulting probability may become more precise from lotteries of a relatively small bound, and

Table 1 Measurement and extension

Layers	0	1	...	$k - 1$	k	...	ρ		
<i>B</i> : base relations	\supseteq_0	\subseteq	\supseteq_1	$\subseteq \cdots \subseteq$	\supseteq_{k-1}	\subseteq	\supseteq_k	$\subseteq \cdots \subseteq$	\supseteq_ρ
	\downarrow		\downarrow		\downarrow		\downarrow		\downarrow
<i>E</i> : constructed relations	\succsim_0	\rightarrow	\succsim_1	$\rightarrow \cdots \rightarrow$	\succsim_{k-1}	\rightarrow	\succsim_k	$\rightarrow \cdots \rightarrow$	\succsim_ρ

Table 2 Developments of the preference and utility theories

	Preference theory		Utility theory
Step <i>B</i>	B0 to B3	\iff Sec.3	<i>b</i> 0 to <i>b</i> 3
	\Downarrow extension: Sec.4		\Downarrow extension: Sec.5
Step <i>E</i>	E0 to E3	\iff Sec.5	<i>e</i> 0 to <i>e</i> 3

the other is that this is intimately related to the EU hypothesis. When ρ is small, the EU hypothesis does not hold typically, while it would hold more as ρ is getting larger.

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

As stated earlier, we provide another approach in terms of a 2-dimensional vector-valued utility functions $\langle \mathbf{v}_k \rangle_{k < \rho+1} = \langle [\bar{v}_k, \underline{v}_k] \rangle_{k < \rho+1}$ and $\langle \mathbf{u}_k \rangle_{k < \rho+1} = \langle [\bar{u}_k, \underline{u}_k] \rangle_{k < \rho+1}$ with Fishburn’s (1970) interval order \geq_I . In each of Steps *B* and *E*, this approach is entirely equivalent to the preference approach, as depicted in Table 2. This may be interpreted as what Von Neumann and Morgenstern (1944), p.29, indicated. The approaches in terms of preferences and utilities enable us to view Steps *B* and *E* in different ways as well as they serve different analytic tools for studies of incomparabilities/comparabilities involved.

Our theory enjoys a weak form of the expected utility hypothesis. This will be discussed in Sect. 6. In the case of $\rho = \infty$, restricting our attention to the set of measurable pure alternatives, Sect. 7 shows that our theory exhibits a form of the classical EU theory. We provide a further extension of \succsim_∞ to have the full form of the classical EU theory; this extension involves some unavoidable non-constructive step, which may be interpreted as the criticism of “super rationality” by Simon (1983).

A remark is on the relationship between *k* and ρ exhibiting a layer and a cognitive bound. The former is a variable in our theory, and the latter is a parameter of the theory. We talk about the sequences $\langle \succsim_k \rangle_{k < \rho+1}$ and $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ describing the process of preference/utility formation, layer to layer, up to ρ . Nevertheless, the final target preferences and utilities are \succsim_ρ and \mathbf{u}_ρ . In the context of the quotation from Turing (1937), within the layers up to ρ , the decision maker can distinguish each probability as a single symbol, but beyond ρ , he would have a difficulty; here, it is assumed that he does not think about his decision problem beyond ρ . When $\rho = \infty$, he can treat

any grid probability as a single entity. This remark leads to the view that our theory is a generalization of the classical EU theory, which is discussed in Sect. 7.

In Sect. 8, we apply our theory to the Allais paradox, specifically to an experimental result from Kahneman and Tversky (1979). We show that the paradoxical results remains when the cognitive bound $\rho \geq 3$. However, when $\rho = 2$, the resultant preference relation \succsim_ρ is compatible with their experimental result, where incomparabilities play crucial roles in explaining them.

The paper is organized as follows: Section 2 explains the concept of probability grids and other basic concepts. Section 3 formulates Step *B* in terms of preferences and utilities and states their equivalence. Section 4 discusses Step *E* in terms of preferences, and Sect. 5 does it in terms of utilities. Section 6 discusses the measurable/non-measurable lotteries and shows that the expected utility hypothesis holds for the measurable lotteries. Section 7 discusses the connection from our theory to the classical EU theory. In Sect. 8, we exemplify our theory with an experimental result in Kahneman and Tversky (1979). Section 9 concludes this paper with comments on further possible studies. Proofs of all the results in each section are given in a separate subsection; only proof of Lemma 2.1 is given in “Appendix.”

2 Preliminaries

Our theory is about preference formation in the context of EU theory. The classical EU theory is the reference point, but our theory deviates from it in various manners. To have clear relations between the classical EU theory and our development, we first mention the classical theory (cf. Herstein and Minor 1953; Fishburn 1982), and then, we start our development. In Sect. 2.2, we give various basic concepts for our theory and one basic lemma. In Sect. 2.3, we give definitions of preferences, indifferences, incomparabilities, and their counterparts in terms of vector-valued utility functions.

2.1 Classical EU theory

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

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

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

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

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

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

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

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

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

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

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

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

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

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

This theory is silent about how a decision maker finds/forms his preferences and does not separate between object components and meta-components for decision making. The above two theorems are meta-components, and some structural components such as lotteries and a preference relation \succsim_E in $(L_{[0,1]}, \succsim_E)$ are object components, but some include both such as highly complex probabilities. Another difficulty is the presumption that a well-formed preference relation \succsim_E already exists somewhere in the mind of decision maker. Here, our theory studies a formation of such a preference relation by reflecting upon his mind (with past experiences) from the simplest case to complex cases. We target to describe this process; Sects. 2.2 and 2.3 prepare basic concepts for the description of the process. Using these basic concepts, Sect. 3 describes Measurement Step *B* and Sects. 4 and 5 describe Extension Step *E*.

Since we target a process of preference formation, a preference relation contains almost necessarily some or many incomparabilities at least in the beginning. In the literature, we find some studies on expected utility theory without the completeness axiom. Aumann (1962) and Fishburn (1971) considered one-way representation theorem [i.e., the *only-if* of (3)], dropping completeness. See Fishburn (1972) for further studies. Dubra and Ok (2002) and Dubra et al. (2004) developed representation theorems in terms of utility comparisons based on all possible expected utility functions for the relation without completeness. In this literature, incomparabilities are given in the preference relation. In contrast, in our approach, incomparabilities are changing with a cognitive bound and may disappear when there are no cognitive bounds.

2.2 Probability grids, lotteries, and decompositions

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

$$\Pi_k = \left\{ \frac{v}{\ell^k} : v = 0, 1, \dots, \ell^k \right\} \text{ for any finite } k \geq 0. \tag{5}$$

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

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

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

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

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

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

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

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

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

We identify each pure alternative x with the lottery having x as its support; so X is regarded as a subset of $L_k(X)$. Specifically, $L_0(X) = X$. Since Π_k is a finite set, every

⁵ See Mendelson (1973) for related basic mathematics.

$f \in L_k(X)$ has a finite support. Since $\Pi_k \subseteq \Pi_{k+1}$, it holds that $L_k(X) \subseteq L_{k+1}(X)$. We denote $L_\infty(X) = \cup_{k < \infty} L_k(X)$. As long as X is finite, $L_k(X)$ is also a finite set, but $L_\infty(X)$ is a countable set and is dense in $L_{[0,1]}(X)$.

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

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

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

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

We denote this by $\sum_{t=1}^{\ell} \frac{1}{\ell} * f_t$, and letting $\widehat{e} = (\frac{1}{\ell}, \dots, \frac{1}{\ell})$, it is written as $\widehat{e} * \widehat{f}$. We can regard $\widehat{e} * \widehat{f}$ as a compound lottery connecting $L_{k-1}(X)$ to $L_k(X)$ by reducing $\widehat{e} * \widehat{f}$ to f in (9). Our theory allows only this form of compound lotteries and reduction with the depth constraint. The next lemma states that $L_k(X)$ is generated from $L_{k-1}(X)$ by taking all compound lotteries of this kind. It facilitates our induction method described in Table 1 reducing an assertion in layer k to layer $k - 1$. A proof of Lemma 2.1 is given in ‘‘Appendix.’’⁶

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

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

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

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

The right-hand side of (10) is the set of composed lotteries from $L_{k-1}(X)$ with the equal weights. The inclusion \supseteq states that the composed lotteries from $L_{k-1}(X)$ belong to $L_k(X)$. The converse inclusion \subseteq is essential and means that each lottery in $L_k(X)$ is decomposed to an equally weighted sum of some (f_1, \dots, f_ℓ) in $L_{k-1}(X)^\ell$ with the depth constraint in (9). In the trivial case that $f = x \in L_0(X)$ is decomposed to $\widehat{f} = (x, \dots, x)$. This will be used in Lemma 4.2. The latter with (11) asserts the choice of a strictly shallower decomposition for f with $\delta(f) > 0$.

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

Here, we give three remarks. One is that when f is a benchmark lottery in $B_k(\bar{y}; \underline{y})$, for its decomposition $\hat{f} = (f_1, \dots, f_\ell)$, each f_i is a benchmark lottery in $B_{k-1}(\bar{y}; \underline{y})$. This fact will be used without referring. The second is that when $f \in L_{k-1}(\bar{X})$, $\hat{f} = (f, \dots, f)$ is a decomposition of f . To allow this triviality, we require only the weak inequality in the depth constraint in (9). The third is that for any subset X' of X , we define $L_k(X') = \{f \in L_k(X) : f(x) > 0 \text{ implies } x \in X'\}$. Hence, $L_k(X')$ is a subset of $L_k(X)$. Lemma 2.1 holds for $L_k(X')$ and $L_{k-1}(X')$.

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

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

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

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

2.3 Incomplete preference relations and vector-valued utility functions

We consider two methods to express the decision maker’s desires: a preference relation and a utility function. We starts with an incomplete preference relation and then works on a vector-valued utility function with the interval order. These are the first departures from the classical EU theory.

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

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

All the axioms are given on the relations \succ, \succsim, \sim , and the relation \bowtie is defined as the residual part of \succsim . Although \sim and \bowtie are sometimes regarded as closely related (cf. Shafer 1986, p. 469), they are well separated in Theorem 6.2 in our theory.

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

Another method of measurement of desires is by a vector-valued function \mathbf{u} with the interval order introduced by Fishburn (1970). Let $\mathbf{u}(f) = [\bar{u}(f), \underline{u}(f)]$ be a 2-dimensional vector-valued function from its domain A to the set $\mathbb{Q}^2 = \mathbb{Q} \times \mathbb{Q}$ with $\bar{u}(f) \geq \underline{u}(f)$ for each $f \in A$. The components $\bar{u}(f)$ and $\underline{u}(f)$ are interpreted as the least upper and greatest lower bounds of possible utilities from f . We say that $\mathbf{u}(f)$ is *effectively single-valued* iff $\bar{u}(f) = \underline{u}(f)$; in this case, we write $\bar{u}(f) = \underline{u}(f) = u(f)$, dropping the upper and lower bars. We use the interval order \geq_I over the values of \mathbf{u} ; for $f, g \in A$,

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

That is, f and g are ordered if and only if the greatest lower bound $\underline{u}(f)$ from f is larger than or equal to the least upper bound $\bar{u}(g)$ from g . This \geq_I allows incomparabilities, for example, if $\mathbf{u}(f) = [\frac{9}{10}, \frac{7}{10}]$ and $\mathbf{u}(g) = [\frac{83}{102}, \frac{83}{102}]$, then f and g are incomparable by \geq_I . The relation \geq_I is transitive, but is not reflexive; $\mathbf{u}(f) \geq_I \mathbf{u}(g)$ and $\mathbf{u}(g) \geq_I \mathbf{u}(f)$ are equivalent to $u(f) = \underline{u}(f) = \bar{u}(g) = \underline{u}(g)$, i.e., this is the case only when the values $\mathbf{u}(f)$ and $\mathbf{u}(g)$ are effectively single-valued and identical.

3 Measurement Step B

We formulate Step B of measurement of pure alternatives up to cognitive bound ρ . This has two sides: in terms of preference relations $\langle \triangleright_k \rangle_{k < \rho+1}$ and in terms of vector-valued utility functions $\langle \mathbf{v}_k \rangle_{k < \rho+1}$. We show the representation theorem on $\langle \triangleright_k \rangle_{k < \rho+1}$ by $\langle \mathbf{v}_k \rangle_{k < \rho+1}$, and the uniqueness theorem on $\langle \mathbf{v}_k \rangle_{k < \rho+1}$ up to positive linear transformations. Finally, we mention that these are well interpreted in terms of Simon’s (1956) satisficing/aspiration argument.

3.1 Base preference streams

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

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

Let \succeq_k be a subset of

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

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

A stream of basic preference relations is expressed as $\langle \succeq_k \rangle_{k < \rho + 1}$; when $\rho < \infty$, it is expressed as $\langle \succeq_0, \succeq_1, \dots, \succeq_\rho \rangle$, and when $\rho = \infty$, it is as $\langle \succeq_0, \succeq_1, \dots \rangle$. We make four axioms on $\langle \succeq_k \rangle_{k < \rho + 1}$; the first requires a property on measurement on layer 0, the second and third require properties on measurement on each layer k , and the fourth connects measurements in layers k and $k + 1$.

Specifically, Axiom B0 requires pure alternatives to be between the upper and lower benchmarks \bar{y}, \underline{y} .

Axiom B0 (Benchmarks): $\bar{y} \succeq_0 x$ and $x \succeq_0 \underline{y}$ for all $x \in X$.

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

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

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

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

Also, $\lambda = \lambda'$ if and only if $[\bar{y}, \lambda; \underline{y}]$ and $[\bar{y}, \lambda'; \underline{y}]$ are indifferent. Thus, \succeq_k is a complete relation over $B_k(\bar{y}; \underline{y})$ by (6). This is the scale part of \succeq_k and is precise up to Π_k . Since $\bar{y} = [\bar{y}, 1; \underline{y}]$ and $\underline{y} = [\bar{y}, 0; \underline{y}]$, it follows from (16) that $\bar{y} \succ_0 \underline{y}$.

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

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

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

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

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

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

3.2 Base utility streams

Let $\langle \mathbf{v}_k \rangle_{k < \rho+1} = \langle [\bar{v}_k, \underline{v}_k] \rangle_{k < \rho+1}$ be a sequence of vector-valued functions, where $\langle \mathbf{v}_k \rangle_{k < \rho+1} = \langle \mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_\rho \rangle$ if $\rho < \infty$ and $\langle \mathbf{v}_k \rangle_{k < \rho+1} = \langle \mathbf{v}_0, \mathbf{v}_1, \dots \rangle$ if $\rho = \infty$. For each $k < \rho + 1$, $\mathbf{v}_k = [\bar{v}_k, \underline{v}_k]$ is a function from $B_k(\bar{y}; \underline{y}) \cup X$ to \mathbb{Q}^2 such that $\bar{v}_k(f) \geq \underline{v}_k(f)$ for all $f \in B_k(\bar{y}; \underline{y}) \cup X$, which are intended to be the least upper utility and greatest lower utility from lottery f . Recall that when $\mathbf{v}_k(f)$ is effectively single-valued, we write $\bar{v}_k(f) = \underline{v}_k(f) = v_k(f)$. The following conditions on $\langle \mathbf{v}_k \rangle_{k < \rho+1}$ are not exactly parallel to Axioms B0 to B3, but these two systems are equivalent, which is stated in Theorem 3.1 :

- b0:** $v_0(\bar{y}) > v_0(\underline{y})$;
- b1:** for $k < \rho + 1$, $v_k([\bar{y}, \lambda; \underline{y}]) = \lambda v_k(\bar{y}) + (1 - \lambda)v_k(\underline{y})$ for all $[\bar{y}, \lambda; \underline{y}] \in B_k(\bar{y}; \underline{y})$;
- b2:** for $k < \rho + 1$ and $x \in X$, $\bar{v}_k(x) = v_k([\bar{y}, \bar{\lambda}_x; \underline{y}])$ and $\underline{v}_k(x) = v_k([\bar{y}, \underline{\lambda}_x; \underline{y}])$ for some $\bar{\lambda}_x$ and $\underline{\lambda}_x$ in Π_k ;
- b3:** for $k < \rho$ and $x \in X$, $\bar{v}_k(x) \geq \bar{v}_{k+1}(x) \geq \underline{v}_{k+1}(x) \geq \underline{v}_k(x)$.

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

We observe that by *b3*, $\bar{v}_k(x) = \underline{v}_k(x)$ implies $\bar{v}_k(x) = \bar{v}_{k+1}(x) = \underline{v}_{k+1}(x) = \underline{v}_k(x)$, i.e., if $\mathbf{v}_k(x) = [\bar{v}_k(x), \underline{v}_k(x)]$ is effectively single-valued, then $\mathbf{v}_{k'}(x)$ is constant and effectively single-valued for any $k' > k$. In particular,

$$v_0(\bar{y}) = v_k(\bar{y}) \text{ and } v_0(\underline{y}) = v_k(\underline{y}) \text{ for all } k < \rho + 1. \tag{17}$$

This and *b1* imply that a benchmark lottery takes the same utility value as long as it belongs $B_k(\bar{y}; \underline{y})$. Also, the values $\lambda_{\bar{y}}$ and $\lambda_{\underline{y}}$ given by *b2* are 1 and 0, since $\bar{y} = [\bar{y}, 1; \underline{y}]$ and $\underline{y} = [\bar{y}, 0; \underline{y}]$. Hence,

$$\delta(\lambda_{\bar{y}}) = \delta(\lambda_{\underline{y}}) = 0. \tag{18}$$

These observations will be used later.

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

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

$$f \succeq_k g \text{ if and only if } \mathbf{v}_k(f) \geq_I \mathbf{v}_k(g). \tag{19}$$

Without difficulty, we can construct a sequence $\langle \mathbf{v}_k \rangle_{k < \rho+1}$ satisfying b_0 to b_3 . Thus, by Theorem 3.1, there is a sequence $\langle \underline{\triangleright}_k \rangle_{k < \rho+1}$ satisfying B_0 to B_3 , which implies the consistency of Axioms B_0 to B_3 .

We have the uniqueness theorem.

Theorem 3.2 (Uniqueness up to affine transformations). *Let $\langle \underline{\triangleright}_k \rangle_{k < \rho+1}$ satisfy Axioms B_0 to B_3 . If $\langle \mathbf{v}_k \rangle_{k < \rho+1}$ and $\langle \mathbf{v}'_k \rangle_{k < \rho+1}$ satisfying b_0 to b_3 represent $\langle \underline{\triangleright}_k \rangle_{k < \rho+1}$ in the sense of (19), there are rational numbers $\alpha > 0$ and β such that $\mathbf{v}'_k(x) = \alpha \mathbf{v}_k(x) + \beta = [\alpha \bar{v}_k(x) + \beta, \alpha \underline{v}_k(x) + \beta]$ for all $x \in X$ and $k < \rho + 1$.*

Conditions b_0 to b_3 require $\mathbf{v}_k(x) = [\bar{v}_k(x), \underline{v}_k(x)]$ to be represented essentially by two values $\bar{\lambda}_x$ and $\underline{\lambda}_x$ in Π_k with $v_0(\bar{y})$ and $v_0(\underline{y})$. However, $v_0(\bar{y})$ and $v_0(\underline{y})$ are allowed to take any values in \mathbb{Q} only with $v_0(\bar{y}) > v_0(\underline{y})$. Hence, for two streams $\langle \mathbf{v}_k \rangle_{k < \rho+1}$ and $\langle \mathbf{v}'_k \rangle_{k < \rho+1}$ representing the same $\langle \underline{\triangleright}_k \rangle_{k < \rho+1}$, one is expressed as a positive linear transformation of the other, which is the above uniqueness result.

The uniqueness up to a positive linear transformation plays a crucial role in the literature on bargaining such as Nash (1950) and also the Nash welfare function theory by Kaneko and Nakamura (1979). The rational number scalars are enough for the 2-person case (cf. Kaneko 1992) and the real-algebraic numbers are enough for the general n -person case. It is easy to generalize Theorem 3.2 for the real numbers scalars, but the problem is how much we can restrict the scalars. Theorem 3.2 is suggestive of how bounded rationality is incorporated to these theories.

The processes described in terms of $\langle \underline{\triangleright}_k \rangle_{k < \rho+1}$ and/or $\langle \mathbf{v}_k \rangle_{k < \rho+1}$ are regarded as thought experiments by the decision maker to search preferences/ utilities in his mind. These preference and utility comparisons are object components for the decision maker, but the above theorems belong to the level of meta-components. From the viewpoint of “bounded rationality”, he may stop his search when he is satisfied and/or tired. This is the same as Simon’s (1956) argument of satisficing/aspiration. We consider one example and exemplify the satisficing/aspiration argument.

Example 3.1 Let $X = \{\bar{y}, y, \underline{y}\}$, $\mathbf{v}_0(\bar{y}) = [1, 1]$, $\mathbf{v}_0(\underline{y}) = [0, 0]$, and $\mathbf{v}_0(y) = [1, 0]$. Also, let $\mathbf{v}_1(y) = [\frac{9}{10}, \frac{7}{10}]$. For $f = [\bar{y}, \frac{8}{10}; \underline{y}]$, $\mathbf{v}_1(f) = [\frac{8}{10}, \frac{8}{10}]$ by (17) and b_1 . Then $\mathbf{v}_1(y) \not\preceq_I \mathbf{v}_1(f)$ and $\mathbf{v}_1(f) \not\preceq_I \mathbf{v}_1(y)$; so y and f are incomparable with respect to $\underline{\triangleright}_1$ by (19). In Fig. 2, $\langle \mathbf{v}_k(y) \rangle_{k < \rho+1} = \langle [\bar{v}_k(y), \underline{v}_k(y)] \rangle_{k < \rho+1}$ is described as solid lines in cases A, B , and C . Since $\mathbf{v}_0(y) = [1, 0]$, we have $\bar{y} \triangleright_0 y \triangleright_0 \underline{y}$ by (19). For $k = 2$, in A , $\mathbf{v}_2(y) = [\frac{77}{10^2}, \frac{77}{10^2}]$ and the decision maker prefers $f = [\bar{y}, \frac{8}{10}; \underline{y}]$ to y ; and in B , $\mathbf{v}_2(y) = [\frac{83}{10^2}, \frac{83}{10^2}]$; he prefers y to f . In C , $\mathbf{v}_k(y) = [\frac{9}{10}, \frac{7}{10}]$ is constant for $k \geq 2$; he gives up comparisons between y and f after $k = 1$.

This example is interpreted in terms of Simon’s satisficing/aspiration. The decision maker starts evaluating the pure alternative y the benchmark scale $B_0(\bar{y}; \underline{y})$. Suppose that he finds $\mathbf{v}_0(y) = [1, 0]$, i.e., he attaches the upper value 1 and lower value 0 to y . When $\rho = 0$, he immediately concludes that y is between \bar{y} and \underline{y} . When $\rho \geq 1$, he goes to layer 1 and uses the more precise scale $B_1(\bar{y}; \underline{y})$ to measure y . Since $\mathbf{v}_1(y) = [\frac{9}{10}, \frac{7}{10}]$, y is better than $[\bar{y}, \frac{7}{10}; \underline{y}]$ but worse than $[\bar{y}, \frac{9}{10}; \underline{y}]$. Still, he has not reached an exact measurement. If $\rho = 1$, he stops introspection and is satisfied with these evaluations of y . If $\rho \geq 2$, he goes to layer $k = 2$; in A of Fig. 2, he

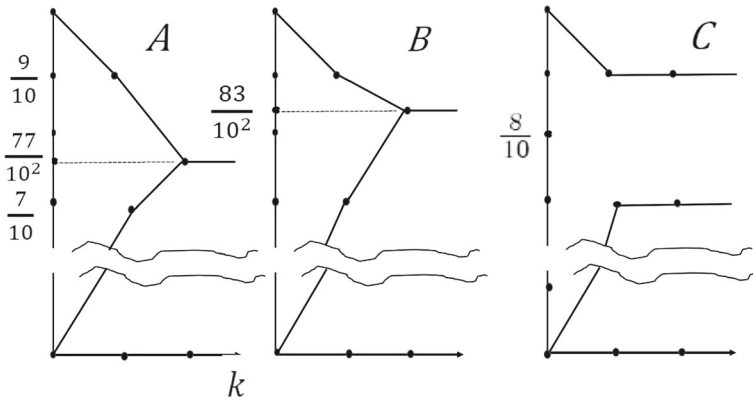


Fig. 2 Upper and lower bounds of utilities

reaches the exact utility value $v_2(y) = [\frac{77}{10^2}, \frac{77}{10^2}]$,⁷ but in C, he has still imprecise values $v_2(y) = [\frac{9}{10}, \frac{7}{10}]$ and does not improve them any more even for $k > 2$.

3.3 Proofs

Proof of Theorem 3.1 (If): Suppose that $\langle v_k \rangle_{k < \rho+1}$ satisfies b0 to b3 and that (19) holds for $\langle \underline{v}_k \rangle_{k < \rho+1}$ and $\langle v_k \rangle_{k < \rho+1}$

B0: We have, by b0, b1, and b2, $v_0(\bar{y}) \geq \bar{v}_0(x)$ and $\underline{v}_0(x) \geq v_0(\underline{y})$, i.e., $\bar{y} \succeq_0 x \succeq_0 \underline{y}$. Thus, B0.

B1: By (19), b1, and b2, we have $[\bar{y}, \lambda; \underline{y}] \succeq_k [\bar{y}, \lambda'; \underline{y}]$ if and only if $v_k([\bar{y}, \lambda; \underline{y}]) \geq v_k([\bar{y}, \lambda'; \underline{y}])$ if and only if $\lambda v_0(\bar{y}) + (1 - \lambda)v_0(\underline{y}) \geq \lambda' v_0(\bar{y}) + (1 - \lambda')v_0(\underline{y})$. By b0, this is equivalent to $\lambda \geq \lambda'$. That is, B1.

B2: Let $[\bar{y}, \lambda; \underline{y}] \succeq_k x$ and $\lambda' > \lambda$. By b2, (19), (17), and b0, we have $\lambda' v_0(\bar{y}) + (1 - \lambda')v_0(\underline{y}) > \lambda v_0(\bar{y}) + (1 - \lambda)v_0(\underline{y}) \geq \bar{v}_k(x)$. Thus, $v_k([\bar{y}, \lambda'; \underline{y}]) = \lambda' v_k(\bar{y}) + (1 - \lambda')v_k(\underline{y}) > \bar{v}_k(x)$. By (19), we have $[\bar{y}, \lambda'; \underline{y}] \succ_k x$. The other case is symmetric.

B3: Let $f \succeq_k g$. By (19), we have $\underline{v}_k(f) \geq \bar{v}_k(g)$. Let $f = x \in X$ and $g = [\bar{y}, \lambda; \underline{y}] \in B_k(\bar{y}; \underline{y})$. Then, by b3, we have $\underline{v}_{k+1}(f) \geq \underline{v}_k(f) \geq v_k(g) = v_{k+1}(g)$. By (19), $f \succeq_{k+1} g$. The case $f \in B_k(\bar{y}; \underline{y})$, $g = x \in X$ is parallel. The case $f = [\bar{y}, \lambda; \underline{y}]$, $g = [\bar{y}, \lambda'; \underline{y}] \in B_k(\bar{y}; \underline{y})$ is similar.

(Only-if): Suppose that $\langle \underline{v}_k \rangle_{k < \rho+1}$ satisfying Axioms B0 to B3 is given. We construct a base utility stream $\langle v_k \rangle_{k < \rho+1} = [\bar{v}_k, \underline{v}_k]$ satisfying (19), as follows: for any $f \in$

⁷ This may be interpreted as including some imprecision: for example, the lower and upper values may be $\frac{77}{10^2}$ and $\frac{78}{10^2}$, but according his aspiration level, the difference $\frac{78}{10^2} - \frac{77}{10^2} = \frac{1}{10^2}$ is tiny and he does not care about the choice between $\frac{77}{10^2}$ and $\frac{78}{10^2}$. By chance, he chooses $v_2(y) = [\frac{77}{10^2}, \frac{77}{10^2}]$.

$$B_k(\bar{y}; \underline{y}) \cup X,$$

$$\begin{aligned} \bar{v}_k(f) &= \min\{\lambda \in \Pi_k : [\bar{y}, \lambda; \underline{y}] \succeq_k f\}; \\ \underline{v}_k(f) &= \max\{\lambda \in \Pi_k : f \succeq_k [\bar{y}, \lambda; \underline{y}]\}. \end{aligned} \tag{20}$$

Note that when $f = [\bar{y}, \lambda_f; \underline{y}] \in B_k(\bar{y}; \underline{y})$, we have $\bar{v}_k(f) = \underline{v}_k(f) = \lambda_f$ for $f \in B_k(\bar{y}; \underline{y})$ by (20). Consider $f = [\bar{y}, \lambda_f; \underline{y}]$, $g = [\bar{y}, \lambda_g; \underline{y}] \in B_k(\bar{y}; \underline{y})$. Then, $f \succeq_k g$ if and only if $[\bar{y}, \lambda_f; \underline{y}] \succeq_k [\bar{y}, \lambda_g; \underline{y}]$ if and only if $\lambda_f \geq \lambda_g$, i.e., $\mathbf{v}_k(f) \geq_I \mathbf{v}_k(g)$ by B1. Let $f \in B_k(\bar{y}; \underline{y})$ and $g = x \in X$. Denote $v_k(f) = \lambda_f$ and $\bar{v}_k(x) = \bar{\lambda}_x$. Suppose $f \succeq_k x$. By (20), $[\bar{y}, \lambda_f; \underline{y}] = f \succeq_k [\bar{y}, \bar{\lambda}_x; \underline{y}]$. By B1, $\lambda_f \geq \bar{\lambda}_x$, i.e., $\mathbf{v}_k(f) \geq_I \mathbf{v}_k(x)$. The converse is obtained by tracing this back. Thus, $f \succeq_k x$ if and only if $\mathbf{v}_k(f) \geq_I \mathbf{v}_k(x)$. The case $f = x \in X$, $g \in B_k(\bar{y}; \underline{y})$ is parallel.

By (16) and B0, we have $b0$. Consider $b1$. Since $v_k(\bar{y}) = \bar{1}$ and $v_k(\underline{y}) = 0$, it follows from the note immediately after (20) that $v_k(f) = \lambda = \lambda v_k(\bar{y}) + (1 - \lambda)v_k(\underline{y})$, which is $b1$. By (20), we have $b2$ and $b3$. □

Proof of Theorem 3.2 Let $\alpha = (v'_0(\bar{y}) - v'_0(\underline{y})) / (v_0(\bar{y}) - v_0(\underline{y}))$ and $\beta = (v_0(\bar{y})v'_0(\underline{y}) - v'_0(\bar{y})v_0(\underline{y})) / (v_0(\bar{y}) - v_0(\underline{y}))$. By $b1$ and (17), we have $\mathbf{v}'_k(\bar{y}) = \alpha \mathbf{v}_k(\bar{y}) + \beta$ and $\mathbf{v}'_k(\underline{y}) = \alpha \mathbf{v}_k(\underline{y}) + \beta$. For $[\bar{y}, \lambda; \underline{y}] \in B_k(\bar{y}; \underline{y})$, we have $\mathbf{v}'_k([\bar{y}, \lambda; \underline{y}]) = \lambda \mathbf{v}'_k(\bar{y}) + (1 - \lambda)\mathbf{v}'_k(\underline{y}) = \alpha \mathbf{v}_k([\bar{y}, \lambda; \underline{y}]) + \beta$ by $b1$.

For $x \in X$, we have $\bar{\lambda}_x$ and $\underline{\lambda}_x$ in Π_k by $b2$ for \mathbf{v}_k such that $\mathbf{v}_k(x) = [v_k([\bar{y}, \bar{\lambda}_x; \underline{y}]), v_k([\bar{y}, \underline{\lambda}_x; \underline{y}])]$. Let $\bar{\lambda}'_x$ and $\underline{\lambda}'_x$ be given by $b2$ for \mathbf{v}'_k . Suppose $\bar{\lambda}_x \neq \bar{\lambda}'_x$, say, $\bar{\lambda}_x > \bar{\lambda}'_x$. Then, $\mathbf{v}_k([\bar{y}, \bar{\lambda}_x; \underline{y}]) \geq_I \mathbf{v}_k(x)$, but $\bar{v}_k(x) = v_k([\bar{y}, \bar{\lambda}_x; \underline{y}]) > v_k([\bar{y}, \bar{\lambda}'_x; \underline{y}])$. Hence, $\mathbf{v}_k([\bar{y}, \bar{\lambda}'_x; \underline{y}]) \not\geq_I \mathbf{v}_k(x)$. However, by definition of $\bar{\lambda}'_x$, we have $\mathbf{v}'_k([\bar{y}, \bar{\lambda}'_x; \underline{y}]) \geq_I \mathbf{v}'_k(x)$. This is impossible since \mathbf{v}_k and \mathbf{v}'_k represent the same \succeq_k . The case $\bar{\lambda}_x < \bar{\lambda}'_x$ is parallel. Thus, $\bar{\lambda}'_x = \bar{\lambda}_x$, and similarly, $\underline{\lambda}'_x = \underline{\lambda}_x$. It was shown in the above paragraph that $\mathbf{v}_k(f) = \alpha \mathbf{v}'_k(f) + \beta$ for any $f \in B_k(\bar{y}; \underline{y})$. This together with $\bar{\lambda}'_x = \bar{\lambda}_x$ and $\underline{\lambda}'_x = \underline{\lambda}_x$ implies $\mathbf{v}'_k(x) = [v'_k([\bar{y}, \bar{\lambda}_x; \underline{y}]), v'_k([\bar{y}, \underline{\lambda}_x; \underline{y}])] = \alpha[v_k([\bar{y}, \bar{\lambda}_x; \underline{y}]), v_k([\bar{y}, \underline{\lambda}_x; \underline{y}])] + \beta = \alpha[\bar{v}_k(x), \underline{v}_k(x)] + \beta = \mathbf{v}_k(x) + \beta$. □

4 Extension Step E: preferences

Step B is an introspection process to find preferences hidden in the mind of the decision maker. On the other hand, Step E is a logical process to extend base preferences found in Step B ; it involves a possible difficulty generated by a new type of probability depths, $\delta(f(x))$, interacting with a finite cognitive bound ρ . This requires our axiomatic system, Axiom E1 in particular, to take a certain specific form. Keeping this remark in mind, we present our axiomatic system for Step E . Throughout this section, let $\langle \succeq_k \rangle_{k < \rho + 1}$ be a given base preference stream satisfying Axioms B0 to B3.

4.1 Extended preference streams

We consider how \sqsupseteq_k is extended to $L_k(X)$ for $k < \rho + 1$. We formulate this derivation by a kind of mathematical induction from the base preferences \sqsupseteq_k and the previously derived relation \succsim_{k-1} ; this process starts at layer 0 to the last layer ρ (or goes to any layer if $\rho = \infty$). These are formulated by three axioms; the first axiom E0 corresponds to the start, and the other two, E1 and E2, describe the extension process. It is shown by Theorem 4.1 that our formalism involves no logical difficulty. Then, we give one additional axiom to capture the central part determined by E0 to E2.

First, Axiom E0 is to convert base preferences \sqsupseteq_k to \succsim_k for each $k < \rho + 1$, depicted as the vertical arrows in Table 1.

- Axiom E0 (Extension)**(i): For any $(f, g) \in D_0$, $f \sqsupseteq_0 g$ if and only if $f \succsim_0 g$.
 (ii): For any $k(1 \leq k < \rho + 1)$ and $(f, g) \in D_k$, if $f \sqsupseteq_k g$, then $f \succsim_k g$.

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

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

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

In layer $k - 1$, each f_t of $\widehat{f} = (f_1, \dots, f_\ell)$ is compared with the corresponding benchmark lottery g_t of $\widehat{g} = (g_1, \dots, g_\ell)$. These preferences are extended to layer k . In Table 1, the horizontal arrows indicate this derivation. A lottery $f \in L_k(X)$ may involve the depth $\delta(f(x))$ of the probability value $f(x)$ and the depths of $\delta(\overline{\lambda}_x)$, $\delta(\underline{\lambda}_x)$ of $\overline{\lambda}_x, \underline{\lambda}_x$ given in b2, for each pure alternative $x \in X$ with $f(x) > 0$. In the lottery $d = \frac{25}{10^2}y * \frac{75}{10^2}\underline{y}$ in Example 3.1, the former is $\delta(\frac{25}{10^2}) = 2$ and the latter is $\delta(\overline{\lambda}_x) = \delta(\underline{\lambda}_x) = \delta(\frac{77}{10^2}) = 2$ in case A. On the other hand, benchmark lotteries involve only the former depths since $\delta(\overline{\lambda}_{\overline{y}}) = \delta(\overline{\lambda}_{\underline{y}}) = 0$ by (18). In Axiom E1, extension is always made based on the benchmark scale. In fact, Lemma 4.1 does not take this constraint into account, but Theorem 4.1 does.

Preferences extended through the benchmark scale $B_k(\overline{y}; \underline{y})$ in E1 are further extended by transitivity, which is the next axiom. Let $0 \leq k < \rho + 1$.

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

We interpret Axioms E1 and E2 as inference rules with Axiom E0 as the bases for \succsim_k . This means that the decision maker constructs $\succsim_0, \succsim_1, \dots$, step by step, using these axioms. This may involve some subtlety; Axioms E0 to E2 may lead to new

unintended preferences. We will show Theorem 4.1, implying that this is not the case for the constructed preference relations.

The following are strengthened versions of E0 and E1:

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

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

Condition E0* states that $\langle \succsim_k \rangle_{k < \rho + 1}$ is a faithful extension of $\langle \succeq_k \rangle_{k < \rho + 1}$ as long as the pairs of lotteries in D_k are concerned. The other, E1*, is a strengthening of E1, too. Without these, some preferences would be added in the derivation process of $\succsim_0, \succsim_1, \dots$. Note that E2 (transitivity) preserves strict preferences in the same way as E1*.

To prove that our extended stream $\langle \succsim_k \rangle_{k < \rho + 1}$ enjoys E0*, E1*, and E2, we first show the following lemma using the EU hypothesis, which is an auxiliary step to Theorem 4.1. A by-product is the consistency of E0*, E1*, and E2. For the lemma, a base utility stream $\langle \mathbf{v}_k \rangle_{k < \rho + 1}$ satisfying (19) in Theorem 3.1 is given.

Lemma 4.1 (Direct application of the EU hypothesis). *Let $\langle \succsim_k^* \rangle_{k < \rho + 1}$ be defined as follows: for all $k < \rho + 1$,*

$$f \succsim_k^* g \text{ if and only if } E_f(\mathbf{v}_k) \geq_I E_g(\mathbf{v}_k). \tag{21}$$

Then, $\langle \succsim_k^ \rangle_{k < \rho + 1}$ satisfying Axioms E0*, E1*, and E2.*

The right-hand side is given by comparisons of the expected values of the vector-valued utility function \mathbf{v}_k . The point of the lemma is not the representation of expected utility values; instead, it is the consistency of E0*, E1*, and E2, which will be used in Theorem 4.1. The consistency of Axioms E0, E1, and E2 is straightforward since E0 takes only preferences given by B0 to B3, and E1 and E2 introduce new preferences from them. On the other hand, E0* and E1* may generate strict preferences, including negations. Hence, the consistency implied by Lemma 4.1 is a basis of our development.

It will be argued in Sect. 7 that when $\rho = \infty$, the limit preference relation \succsim_∞^* is determined E0, E1, and E2 under some additional condition on $L_\infty(X) = \cup_{k < \infty} L_k(X)$.

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

$$f \succsim_k g \text{ implies } f \succsim'_k g. \tag{22}$$

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

Using these concepts, we construct the smallest stream $\langle \succsim_k \rangle_{k < \rho + 1}$, each of which is a binary relation on $L_k(X)$, and show, using Lemma 4.1, that it satisfies E0* and E1* as well as E2.

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

$$\begin{aligned} \succsim_0 &= (\supseteq_0)^{tr}; \text{ and} \\ \succsim_k &= [(\succsim_{k-1})^{E1} \cup (\supseteq_k)]^{tr} \text{ for each } k \ (1 \leq k < \rho + 1) \end{aligned} \tag{23}$$

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

The construction of $\langle \succsim_k \rangle_{k < \rho+1}$ starts with $\succsim_0 = (\supseteq_0)^{tr}$, which is well defined since \supseteq_0 is a binary relation in D_0 . Then, provided that \succsim_{k-1} and \supseteq_k are already given, \succsim_k is defined to be $[(\succsim_{k-1})^{E1} \cup (\supseteq_k)]^{tr}$. This is a subset of $L_k(X)^2$; thus, it is a binary relation. Thus, the stream $\langle \succsim_k \rangle_{k < \rho+1}$ is unique and is the smallest among the streams satisfying E0 to E2. Furthermore, it satisfies E0* and E1*, where Lemma 4.1 is used. This assertion guarantees that the constructed stream $\langle \succsim_k \rangle_{k < \rho+1}$ is a faithful extension of $\langle \supseteq_k \rangle_{k < \rho+1}$.⁸

The stream $\langle \succsim_k \rangle_{k < \rho+1}$ constructed in Theorem 4.1 differs from the stream $\langle \succsim_k^* \rangle_{k < \rho+1}$ given in Lemma 4.1 in that the two types of depths, $\delta(\bar{\lambda}_x)$, $\delta(\underline{\lambda}_x)$, and $\delta(f(x))$ are taken into account in the former, but for the latter, (21) defines the expected utility, ignoring them. These depths are interactive with cognitive bound ρ . When $\rho = \infty$, $\succsim_\infty = \cup_{k < \rho+1} \succsim_k$ and $\succsim_\infty^* = \cup_{k < \rho+1} \succsim_k^*$ coincide under some additional restriction on $L_\infty(X)$, which will be discussed in Sect. 7.1. In general, $\langle \succsim_k \rangle_{k < \rho+1}$ and $\langle \succsim_k^* \rangle_{k < \rho+1}$ differ and coincide only in partial domains, which will be discussed in Sect. 6.

In the constructing process of $\langle \succsim_k \rangle_{k < \rho+1}$, E0 to E2 are all extension axioms, and the resulting stream of (23) is uniquely constructed, while there are multiple preference streams satisfying E0 to E2. It would be easier for various purposes to extract the essence of (23) by formulating one axiom. It is formulated as Axiom E3, which requires that a preference $f \succsim_k g$ be based on comparisons with the benchmark scale $B_k(\bar{y}; \underline{y})$ with either \supseteq_k or \succsim_{k-1} , that is, it eliminates preferences from other possible sources. Note that h in E3 may be the same as f or g .

Axiom E3. Let $k < \rho + 1$ and $f, g \in L_k(X)$ with $f \succsim_k g$. Then $f \succsim_k h \succsim_k g$ for some $h \in B_k(\bar{y}; \underline{y})$. When $k \geq 1$, for the pair (f, h) , $f \supseteq_k h$ holds or f, h have decompositions \hat{f}, \hat{h} with $\hat{f} \succsim_{k-1} \hat{h}$; the same holds for the pair (h, g) .

The stream $\langle \succsim_k \rangle_{k < \rho+1}$ given by (23) is characterized by adding E3 to E0 to E2.

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

Throughout the following, the stream given by (23) is denoted by $\langle \succsim_k \rangle_{k < \rho+1}$. Other streams may have some additional superscripts such as ', *.

Lemma 4.2 will be used in the subsequent analyses; (1) is the horizontal arrows in Table 1; and (2) means that \succsim_k is bounded in $L_k(X)$ by the upper and lower benchmarks \bar{y} and \underline{y} .

⁸ We extend an already built preference relation \succsim_{k-1} and a given base relation \supseteq_k to \succsim_k by Axioms E1 and E2, which is the weakest relation. This extension is somewhat similar to Dubra and Ok's (2002) argument: they extend a preference relation on a finite set of lotteries to the smallest relation satisfying Axiom NM2-ID1, and they show the extended relation is represented by a set of expected utilities.

Lemma 4.2 *Let $\langle \succsim_k \rangle_{k < \rho+1}$ satisfies E0 to E3, and $1 \leq k < \rho + 1$.*

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

The EU hypothesis is included in Axiom B1 and condition b1 along the benchmark scale $B_k(\bar{y}; \underline{y})$, and E1 is a weak form of Axiom NM2 (independence). It follows from Lemma 4.1 and Theorem 4.1 that there are possibly multiple preference streams satisfying E0 to E2, among which some satisfy the EU representation in (21) but not in general. This is caused by two types of depths included in a lottery. For example, lottery $d = \frac{25}{10^2} y * \frac{75}{10^2} \underline{y}$ involves the depths of coefficient $\frac{25}{10^2}$ and of evaluation λ_y . This is the reason for the EU hypothesis to hold only for some partial domain, which will be explicitly studied in Sect. 6.

4.2 Proofs

Proof of Lemma 4.1 We show that $\langle \succsim_k^* \rangle_{k < \rho+1}$ given by (21) satisfies E0*, E1*, and E2. By Theorem 3.2, we can assume that $v_k(\bar{y}) = 1$ and $v_k(\underline{y}) = 0$.

Since $E_f(v_k) = \lambda$ if $f = [\bar{y}, \lambda; \underline{y}] \in B_k(\bar{y}; \underline{y})$ and $E_x(\bar{v}_k) = \bar{v}_k(x)$ if $f = x \in X$. Hence, by (19) and b2, $f \triangleright_k x$ if and only if $\lambda \geq \bar{v}_k(x)$ if and only if $E_f(v_k) \geq E_x(\bar{v}_k)$. The other cases are symmetric. Thus, E0* holds for any $(f, g) \in D_k$.

It remains to show that $\langle \succsim_k^* \rangle_{k < \rho+1}$ satisfies E1* and E2. Since (21) gives the interval order over the set $\{[E_f(\bar{v}_k), E_f(v_k)] : f \in L_k(X)\}$, E2 holds. We show E1*. Let $f \in L_k(X)$, $g \in B_k[\bar{y}; \underline{y}]$ and their decompositions \hat{f} and \hat{g} with $\hat{f} \succsim_{k-1}^* \hat{g}$. By (21), $E_{f_t}(v_k) \geq E_{g_t}(\bar{v}_k)$ for all $t = 1, \dots, \ell$. Then, $E_f(v_k) = E_{\hat{e} * \hat{f}}(v_k) = \sum_{t=1}^{\ell} \frac{1}{\ell} E_{f_t}(v_k) \geq \sum_{t=1}^{\ell} \frac{1}{\ell} E_{g_t}(\bar{v}_k) = E_{\hat{e} * \hat{g}}(\bar{v}_k) = E_g(\bar{v}_k)$. If strict preferences are included in the decompositions, the conclusion is strict; thus, we have E1*. □

Proof of Theorem 4.1 This has the three assertions: (a) $\langle \succsim_k \rangle_{k < \rho+1}$ is a sequence of a binary relations satisfying Axioms E0 to E2; (b) it is the smallest in the sense of (22) among the streams $\langle \succsim'_k \rangle_{k < \rho+1}$ satisfying E0 to E2; and (c) E0*, E1* hold for $\langle \succsim_k \rangle_{k < \rho+1}$.

(a): E2 follows directly from (23). Consider E0. (ii) follows from (23). We show that $\succsim_0 = (\triangleright_0)^{tr}$ satisfies that for any $(f, g) \in D_0$, $f \succsim_0 g$ implies $f \triangleright_0 g$. Since $\succsim_0 = (\triangleright_0)^{tr}$, there is a sequence $f = h_0 \triangleright_0 \dots \triangleright_0 h_m = g$. If $h_t \in X - B_0(\bar{y}; \underline{y})$, then $h_{t-1} \in B_0(\bar{y}; \underline{y})$ and $h_{t+1} \in B_0(\bar{y}; \underline{y})$. By B2, $\lambda_{t-1} \geq \lambda_{t+1}$, where $h_{t-1} = [\bar{y}; \lambda_{t-1}, \underline{y}]$ and $h_{t+1} = [\bar{y}; \lambda_{t+1}, \underline{y}]$. If $h_t, h_{t+1} \in B_0(\bar{y}; \underline{y})$, then $\lambda_t \geq \lambda_{t+1}$. Hence, we can shorten the sequence to $f = h_0 \triangleright_0 h_m = g$. Thus, $f \triangleright_0 g$.

Consider E1. Suppose that $f \in L_k(X)$ and $g \in B_k[\bar{y}; \underline{y}]$ have decompositions \hat{f}, \hat{g} with $\hat{f} \succsim_{k-1} \hat{g}$. By (23), we have $f = e * \hat{f} \succsim_k e * \hat{g} = g$. This causes no difficulty, even if $f \in B_k(X)$ and $g \in B_k[\bar{y}; \underline{y}]$. The symmetric case $\hat{g} \succsim_{k-1} \hat{f}$ is similar.

(b): We prove by induction on k that $\langle \succsim_k \rangle_{k < \rho+1}$ satisfies (22) for any $\langle \succsim'_k \rangle_{k < \rho+1}$ satisfying E0 to E2. When $k = 0$, we have $\succsim_0 = (\triangleright_0)^{tr}$ by (23). Let $f \succsim_0 g$, i.e., $f(\triangleright_0)^{tr} g$, which implies that there is a sequence $f = h_0 \triangleright_0 h_1 \triangleright_0 \dots \triangleright_0 h_m = g$. By E0.(i), we have $f = h_0 \succsim'_0 h_1 \succsim'_0 \dots \succsim'_0 h_m = g$. By E2 for \succsim'_0 , we have $f \succsim'_0 g$.

Now, we assume that (22) holds for $k - 1$. Let $f \succsim_k g$. By (23), there is a sequence $f = h_0 \succsim_k \dots \succsim_k h_m = g$ such that each $h_t \succsim_k h_{t+1}$ is a consequence of E1 or $h_t \succsim_k h_{t+1}$ is $\widehat{h}_t \supseteq_k h_{t+1}$. In the first case, there are decompositions $\widehat{h}_t, \widehat{h}_{t+1}$ of h_t, h_{t+1} such that $\widehat{h}_t \succsim_{k-1} \widehat{h}_{t+1}$. By the induction hypothesis, we have $\widehat{h}_t \succsim'_{k-1} \widehat{h}_{t+1}$. Thus, $h_t \succsim'_k h_{t+1}$ by E1 for \succsim'_k . In the second case, $h_t \supseteq_k h_{t+1}$ implies $h_t \succsim'_k h_{t+1}$ by E0.(ii) for \succsim'_k . Hence, $f \succsim'_k g$ by E2 for \succsim'_k .

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

$$f \succsim_k g \text{ implies } f \succsim_k^* g. \tag{24}$$

E0*: Since E0* holds for $\widehat{\succsim}_k^*$ by Lemma 4.1, for any $(f, g) \in D_k$, $f \succsim_k^* g$ implies $f \supseteq_k g$. Thus, if $f \succsim_k g$, then $f \succsim_k^* g$ by (24), which implies $f \supseteq_k g$. For the converse, $f \supseteq_k g$ implies $f \succsim_k g$ by (23).

E1*: First, we prove by induction on k that for all $k < \rho + 1$ and $f, g \in L_k(X)$,

$$f \succ_k g \text{ implies } f \succ_k^* g. \tag{25}$$

We make the induction hypothesis that (25) holds for $k - 1$. Now, let $f, g \in L_k(X)$ with $f \succ_k g$. By (23), there are $h_0 = f, h_1, \dots, h_m = g$ in $L_k(X)$ such that $(h_l, h_{l+1}) \in (\succ_{k-1})^{E1}$ or $(h_l, h_{l+1}) \in (\supseteq_k)$ for each $l = 0, \dots, m - 1$.

Consider case (i) : $(h_l, h_{l+1}) \in (\supseteq_k)$. Then by E0* for $\langle \widehat{\succsim}_k^* \rangle_{k < \rho + 1}$, it holds that $h_l \succsim_k^* h_{l+1}$, and also, if the premise is strict, it follows from E0* for $\langle \widehat{\succsim}_k^* \rangle_{k < \rho + 1}$ that $h_l \succ_k^* h_{l+1}$. Now, consider case (ii) : $(h_l, h_{l+1}) \in (\succ_{k-1})^{E1}$. Let $\widehat{h}_l, \widehat{h}_{k+1}$ be decompositions of h_l, h_{l+1} so that $\widehat{h}_l \succsim_{k-1} \widehat{h}_{l+1}$ with/without strict preferences for some components. Hence, by (24) and (25) (the induction hypothesis), the same holds for $\widehat{\succsim}_{k-1}^*$. By E1* for $\langle \widehat{\succsim}_k^* \rangle_{k < \rho + 1}$, we have $h_l \succsim_k^* h_{l+1}$ and $h_l \succ_k^* h_{l+1}$ if strict preferences hold for some components. At least one of $l = 0, \dots, m - 1$, we have strict preferences for $(h_l, h_{l+1}) \in (\supseteq_k)$ or $\widehat{h}_l \succsim_{k-1} \widehat{h}_{l+1}$, because of $f \succ_k g$. This and the above assertions in (i) and (ii) imply $f \succ_k^* g$.

Finally, we verify E1* for $\langle \widehat{\succsim}_k^* \rangle_{k < \rho + 1}$. Let $f \in L_k(X)$ and $g \in B_k(\bar{y}; \underline{y})$, and let their decompositions be \widehat{f}, \widehat{g} with $\widehat{f} \succ_{k-1} \widehat{g}$. Suppose that one of these preferences is strict. By E1, we have $f \succsim_k g$. It suffices to show not $g \succsim_k f$. However, $\widehat{f} \succ_{k-1} \widehat{g}$ implies $\widehat{f} \succ_{k-1}^* \widehat{g}$ by (24), and some components of the latter hold strictly by (25). By E1* for $\langle \widehat{\succsim}_k^* \rangle_{k < \rho + 1}$, we have $f \succ_k^* g$, implying not $g \succsim_k^* f$. By the contrapositive of (24), we have not $g \succsim_k f$. □

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

$$f \widehat{\succsim}_k^* g \text{ if and only if } f \widehat{\succsim}_k g. \tag{26}$$

Since $\langle \widehat{\succsim}_k^* \rangle_{k < \rho + 1}$ is the smallest stream satisfying E0 to E2 by Theorem 4.1, the *if* part holds for any $k < \rho + 1$.

Consider the *only-if* part. Let $k = 0$. Let $f, g \in L_0(X)$ with $f \widehat{\succsim}_0^* g$. Then, by E3 for $\widehat{\succsim}_0^*$, we have an $h \in B_0(\bar{y}; \underline{y})$ with $f \widehat{\succsim}_0^* h \widehat{\succsim}_0^* g$. Thus, $f \supseteq_0 h \supseteq_0 g$

by E0 for \succsim_0^* . Hence, $f(\geq_0)^{tr} g$, i.e., $f \succsim_0 g$ by (23). Now, we make the induction hypothesis that the *only-if* part holds for $k - 1$. Let $f \succsim_k^* g$. Then, by E3, we have $h_0 := f \succsim_k^* h_1 \succsim_k^* h_2 := g$ for some $h_1 \in B_k(\bar{y}; \underline{y})$. Let $h_0 = x \in X$. If $h_0 \geq_k h_1$, then $h_0 \succsim_k h_1$ by E0* for \succsim_k . Suppose $\widehat{h}_0 \succsim_{k-1}^* \widehat{h}_1$ for some decompositions $\widehat{h}_0, \widehat{h}_1$ of h_0, h_1 . By the induction hypothesis, we have $\widehat{h}_0 \succsim_{k-1} \widehat{h}_1$. Thus, by E1 for \succsim_k , we have $h_0 \succsim_k h_1$. In the above two cases, we have $h_0 \succsim_k h_1$. By the same argument, we have $h_1 \succsim_k h_2$. Thus, by E2 for \succsim_k , we have $f = h_0 \succsim_k h_2 = g$, i.e., $f \succsim_k g$. \square

Proof of Lemma 4.2 (1): Let $f \in L_{k-1}(X)$ and $g \in B_{k-1}(\bar{y}; \underline{y})$. Suppose $f \succsim_{k-1} g$. Then, $f, g \in L_{k-1}(X) \subseteq L_k(X)$. Let $f_1 = \dots = f_\ell = f$ and $g_1 = \dots = g_\ell = g$. Then, $f = \sum_{t=1}^\ell \frac{1}{\ell} * f_t$ and $g = \sum_{t=1}^\ell \frac{1}{\ell} * g_t$. By E1, we have $f \succsim_k g$. The case $g \succsim_{k-1} f$ is similar.

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

(2): Let $f \in L_0(X) = X$. By B0 and $\succsim_0 = \geq_0$, we have the assertion for $k = 0$. Suppose the induction hypothesis that $\bar{y} \succsim_{k-1} f \succsim_{k-1} \underline{y}$ for any $f \in L_{k-1}(X)$. Consider $\widehat{f} \in L_k(X)$. Then, by Lemma 2.1, there is a vector $\widehat{f} \in L_{k-1}(X)^\ell$ such that $f = \widehat{e} * \widehat{f}$. By the induction hypothesis, $\bar{y} \succsim_{k-1} f_t \succsim_{k-1} \underline{y}$ for any $t \leq \ell$. By E1, $\bar{y} = \widehat{e} * \bar{y} \succsim_k f = \widehat{e} * \widehat{f} \succsim_k \widehat{e} * \underline{y} = \underline{y}$. \square

5 Extension Step E: utilities

We extend a base utility stream $\langle \mathbf{v}_k \rangle_{k < \rho+1}$ to $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ so that each \mathbf{u}_k is a function over $L_k(X)$ to \mathbb{Q}^2 . We show that this approach is equivalent to that given in Sect. 4. It provides clear-cut interpretations and mathematical tractability of the entire theory.

5.1 Extended utility streams

Let a base utility stream $\langle \mathbf{v}_k \rangle_{k < \rho+1}$ satisfying b0 to b3 be given. Consider a stream of functions $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ so that each $\mathbf{u}_k = [\bar{u}_k, \underline{u}_k]$ is a function from $L_k(X)$ to \mathbb{Q}^2 with $\bar{u}_k(f) \geq \underline{u}_k(f)$ for all $f \in L_k(X)$. As for a base utility stream, the values $\bar{u}_k(f)$ and $\underline{u}_k(f)$ are interpreted as the least upper and greatest lower bounds of possible utilities from f . Recall that when $\mathbf{u}_k = [\bar{u}_k, \underline{u}_k]$ is effectively single-valued for f , we drop the upper and lower bars from $\bar{u}_k(f) = \underline{u}_k(f)$ as $u_k(f)$. For $\widehat{f} = (f_1, \dots, f_\ell)$, we write $\bar{u}_k(\widehat{f}) = (\bar{u}_k(f_1), \dots, \bar{u}_k(f_\ell))$ and $\underline{u}_k(\widehat{f}) = (\underline{u}_k(f_1), \dots, \underline{u}_k(f_\ell))$. For \widehat{f}, \widehat{g} , we write $\underline{u}_k(\widehat{f}) \geq \bar{u}_k(\widehat{g})$ to mean “ $\underline{u}_k(f_t) \geq \bar{u}_k(g_t)$ for $t \leq \ell$ ”.

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

- e0:** The restriction of \mathbf{u}_k to $B_k(\bar{y}; \underline{y}) \cup X$ coincides with \mathbf{v}_k .
- e1:** Let $f \in L_k(X) - B_k(\bar{y}; \underline{y}) \cup X$, $g \in B_k(\bar{y}; \underline{y})$, and \widehat{f}, \widehat{g} their decompositions. If $u_{k-1}(\widehat{g}) \geq \bar{u}_{k-1}(\widehat{f})$ or $\underline{u}_{k-1}(\widehat{f}) \geq u_{k-1}(\widehat{g})$, then $u_k(g) \geq \bar{u}_k(f)$ or $\underline{u}_k(f) \geq u_k(g)$, respectively.

- e2: For any $f \in L_k(X) - B_k(\bar{y}; \underline{y}) \cup X$, there are decompositions $\widehat{f}, \widehat{f}'$ of f such that $\bar{u}_k(f) = \widehat{e} * \bar{u}_{k-1}(\widehat{f})$ and $\underline{u}_k(f) = \widehat{e} * \underline{u}_{k-1}(\widehat{f}')$.
- e3: For any $f \in L_k(X) - B_k(\bar{y}; \underline{y}) \cup X$, there are g, h in $B_k(\bar{y}; \underline{y})$ such that $\bar{u}_k(f) = u_k(g)$ and $\underline{u}_k(f) = u_k(h)$.

Conditions e0 and e1 correspond to E0 and E1. By b1 and e0, u_k is effectively single-valued for $g \in B_k(\bar{y}; \underline{y})$, i.e., $u_k(g) = v_k(g)$. Since the interval order \geq_I satisfies transitivity, no condition corresponding to E2 is assumed. Instead, e2 requires that the least upper and greatest lower utilities $\bar{u}_k(f)$ and $\underline{u}_k(f)$ come from those of some decompositions. Condition e3 requires $\bar{u}_k(f)$ and $\underline{u}_k(f)$ be measured by the benchmark scale $B_k(\bar{y}; \underline{y})$ of the same depth k , which is a depth constraint corresponding to Axiom E3. Under e0, b1, and b2, this is equivalent to

- e3*: For any $f \in L_k(X)$, there are g, h in $B_k(\bar{y}; \underline{y})$ such that $\bar{u}_k(f) = u_k(g)$ and $\underline{u}_k(f) = u_k(h)$.

We may use either e3 or e3* whichever is more convenient.

Also, it holds that

$$\begin{aligned} &\text{If } g \in B_k(\bar{y}; \underline{y}) \text{ and } g = \widehat{e} * \widehat{g} \text{ for some } \widehat{g} \in B_{k-1}(\bar{y}; \underline{y})^\ell, \\ &\text{then } u_k(g) = \widehat{e} * u_{k-1}(\widehat{g}). \end{aligned} \tag{27}$$

That is, if a benchmark lottery g is decomposed into \widehat{g} , then the utility value $u_k(g)$ is obtained from the weighted sum of $u_{k-1}(\widehat{g})$. This will be proved in Sect. 5.2.

First, we present the unique determination of a possible utility stream $\langle u_k \rangle_{k < \rho+1}$ extended from a given base utility stream $\langle v_k \rangle_{k < \rho+1}$.

Theorem 5.1 (Unique extension). *Let $\langle u_k \rangle_{k < \rho+1}$ be an extended utility stream satisfying e0 to e3. Then, it holds that for any $f \in L_k(X)$ with $k \geq 0$,*

$$\bar{u}_k(f) = \begin{cases} \bar{v}_k(f) & \text{if } f \in B_k(\bar{y}; \underline{y}) \cup X \\ \min\{\widehat{e} * \bar{u}_{k-1}(\widehat{f}) : \widehat{f} \text{ is a decomposition of } f\} & \text{otherwise} \end{cases} \tag{28}$$

$$\underline{u}_k(f) = \begin{cases} \underline{v}_k(f) & \text{if } f \in B_k(\bar{y}; \underline{y}) \cup X \\ \max\{\widehat{e} * \underline{u}_{k-1}(\widehat{f}) : \widehat{f} \text{ is a decomposition of } f\} & \text{otherwise.} \end{cases} \tag{29}$$

Conversely, when a base utility stream $\langle v_k \rangle_{k < \rho+1}$ is given, an extended utility stream $\langle u_k \rangle_{k < \rho+1}$ is uniquely determined from $\langle v_k \rangle_{k < \rho+1}$ by (28) and (29), and it satisfies e0 to e3.

Since $\langle v_k \rangle_{k < \rho+1}$ representing $\langle \triangleright_k \rangle_{k < \rho+1}$ is uniquely determined up to positive linear transformations, stated in Theorem 3.2, and since the extended utility stream $\langle u_k \rangle_{k < \rho+1}$ is uniquely determined by $\langle v_k \rangle_{k < \rho+1}$, the stream of pair $\langle v_k, u_k \rangle_{k < \rho+1}$ is uniquely determined up to positive linear transformations.

The existence of $\langle u_k \rangle_{k < \rho+1}$ representing $\langle \widetilde{\triangleright}_k \rangle_{k < \rho+1}$ is guaranteed by the next theorem. Recall that $\langle \triangleright_k \rangle_{k < \rho+1}$ with B0 to B3 is assumed behind $\langle \widetilde{\triangleright}_k \rangle_{k < \rho+1}$ by E0 and that $\langle v_k \rangle_{k < \rho+1}$ with b0 to b3 is assumed behind $\langle u_k \rangle_{k < \rho+1}$ by e0. These $\langle \triangleright_k \rangle_{k < \rho+1}$ and $\langle v_k \rangle_{k < \rho+1}$ are connected by (19) in Theorem 3.1. Axiom E3 and

condition $e3$ are separately treated in Theorem 5.2, because this separation will be needed for Theorem 7.1.

Theorem 5.2 (Representation of $\langle \succsim_k \rangle_{k < \rho+1}$ by $\langle \mathbf{u}_k \rangle_{k < \rho+1}$). *A preference stream $\langle \succsim_k \rangle_{k < \rho+1}$ satisfies E0 to E2 (and E3, respectively) if and only if there is a utility stream $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ satisfying $e0$ to $e2$ (and $e3$) such that for any $k < \rho + 1$ and $f, g \in L_k(X)$,*

$$f \succsim_k g \text{ if and only if } \mathbf{u}_k(f) \geq \mathbf{u}_k(g). \tag{30}$$

Table 2 summarizes the results in Sects. 3 to 5; here E3 and $e3$ are included 5.2. Section 3 is started with the theory of a base preference stream $\langle \triangleright_k \rangle_{k < \rho+1}$ and of a base utility stream $\langle \mathbf{v}_k \rangle_{k < \rho+1}$. In Sect. 4, $\langle \triangleright_k \rangle_{k < \rho+1}$ is extended to $\langle \succsim_k \rangle_{k < \rho+1}$; Theorems 4.1 and 4.2 show the unique existence of $\langle \succsim_k \rangle_{k < \rho+1}$ satisfying E0 to E3, relative to $\langle \triangleright_k \rangle_{k < \rho+1}$. Correspondingly, Theorem 5.1 implies that $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ is uniquely determined relative to $\langle \mathbf{v}_k \rangle_{k < \rho+1}$. Theorem 5.2 states the existence of an extended utility stream $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ when $\langle \succsim_k \rangle_{k < \rho+1}$ satisfies E0 to E3.

The above theorems can be regarded as a substantiation of the indication, by Von Neumann and Morgenstern (1944), p. 29, of a possibility of a representation of a preference relation involving incomparabilities in terms of a higher-dimensional vector-valued function.

The utility stream $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ given by Theorem 5.2 with E3 and $e3$ differs from the EU representation in Lemma 4.1. This causes some difficulty in practical calculation of $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ for the case $\rho < \infty$, but Theorem 5.1 is useful for practical purpose. We show how to calculate $\bar{\mathbf{u}}_k$ and $\underline{\mathbf{u}}_k$ in Example 3.1, which will be used in Sect. 8.

Example 5.1 Consider A and B of Example 3.1 with $X = \{\bar{y}, \underline{y}, y\}$. Recall $\mathbf{v}_0(\bar{y}) = \mathbf{v}_1(\bar{y}) = [1, 1]$, $\mathbf{v}_0(\underline{y}) = \mathbf{v}_1(\underline{y}) = [0, 0]$, $\mathbf{v}_0(y) = [1, 0]$, $\mathbf{v}_1(y) = [\frac{9}{10}, \frac{7}{10}]$. These values are the same as $\mathbf{u}_0, \mathbf{u}_1$ by $e0$. Also, we have:

$$A : \mathbf{v}_2(y) = \mathbf{u}_2(y) = \left[\frac{77}{10^2}, \frac{77}{10^2} \right] \text{ and } B : \mathbf{v}'_2(y) = \mathbf{u}'_2(y) = \left[\frac{83}{10^2}, \frac{83}{10^2} \right]. \tag{31}$$

Consider how to calculate $\mathbf{u}_2(d)$ and $\mathbf{u}_3(d)$ for $d = \frac{25}{10^2} y * \frac{75}{10^2} \underline{y}$. We mainly consider case A and will adjust the calculation for case B .

The lottery d has the three types of decompositions $\frac{t}{10} * y + \frac{5-2t}{10} * [y, \frac{5}{10}; \underline{y}] + \frac{5+t}{10} * \underline{y}$ for $t = 0, 1, 2$ in (12). Among these, the one with $t = 2$ gives \mathbf{u}_2 and \mathbf{u}_3 ; since y is evaluated in a shallower layer through $\frac{5-2t}{10} * [y, \frac{5}{10}; \underline{y}]$ than $\frac{t}{10} * y$, it is more accurate to use $\frac{t}{10} * y$ than $\frac{5-2t}{10} * [y, \frac{5}{10}; \underline{y}]$. Thus, we take the largest weight $t = 2$ to $\frac{t}{10} * y$ for the min and max operators in (28) and (29).

The second term $[y, \frac{5}{10}; \underline{y}] = \frac{5}{10} y * \frac{5}{10} \underline{y}$ itself is regarded as a unique decomposition; we have, by (28) and (29),

$$\mathbf{u}_1 \left(\frac{5}{10} y * \frac{5}{10} \underline{y} \right) = \frac{5}{10} \mathbf{v}_0(y) + \frac{5}{10} \mathbf{v}_0(\underline{y}) = \left[\frac{5}{10}, 0 \right]. \tag{32}$$

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

$$\begin{aligned} u_2(d) &= u_2\left(\frac{2}{10}y * \frac{1}{10}\left(\frac{5}{10}y * \frac{5}{10}\underline{y}\right) * \frac{7}{10}\underline{y}\right) \\ &= \frac{2}{10}u_1(y) + \frac{1}{10}u_1\left(\frac{5}{10}y * \frac{5}{10}\underline{y}\right) + \frac{7}{10}u_1(\underline{y}) \\ &= \frac{2}{10}\left[\frac{9}{10}, \frac{7}{10}\right] + \frac{1}{10}\left[\frac{5}{10}, 0\right] + \frac{1}{10}[0, 0] = \left[\frac{23}{10^2}, \frac{14}{10^2}\right]. \end{aligned} \tag{33}$$

This is compared with $u_2(c) = u_2(\frac{2}{10}\bar{y} * \frac{8}{10}\underline{y}) = [\frac{2}{10}, \frac{2}{10}]$, and thus, by Theorem 5.2, c and d are incomparable with respect to \succsim_2 .

In the same as (32), we have $u_2(\frac{5}{10}y * \frac{5}{10}\underline{y}) = \frac{5}{10}v_1(y) + \frac{5}{10}v_1(\underline{y}) = [\frac{45}{10^2}, \frac{35}{10^2}]$. Then, we calculate $u_3(d)$ for case $\rho = 3$:

$$\begin{aligned} u_3(d) &= u_3\left(\frac{2}{10}y * \frac{1}{10}\left(\frac{5}{10}y * \frac{5}{10}\underline{y}\right) * \frac{7}{10}\underline{y}\right) \\ &= \frac{2}{10}u_2(y) + \frac{1}{10}u_2\left(\frac{5}{10}y * \frac{5}{10}\underline{y}\right) + \frac{7}{10}u_2(\underline{y}) \\ &= \frac{2}{10}\left[\frac{77}{10^2}, \frac{77}{10^2}\right] + \frac{1}{10}\left[\frac{45}{10^2}, \frac{35}{10^2}\right] + \frac{7}{10}[0, 0] = \left[\frac{199}{10^3}, \frac{189}{10^3}\right]. \end{aligned} \tag{34}$$

Here, c is strictly preferred to d , since $u_3(c) = u_3(\frac{2}{10}\bar{y} * \frac{8}{10}\underline{y}) = [\frac{2}{10}, \frac{2}{10}]$. For $k \geq 4$, since $u_k(d) = [\frac{1925}{10^4}, \frac{1925}{10^4}]$, c is strictly preferred to d .

Consider case $B : u'_2(y) = [\frac{83}{10^2}, \frac{83}{10^2}]$ for $\rho = 2$. Then, the above calculation (33) for $u_2(d)$ remains the same for $u'_2(d)$ with $u'_2(d) = [\frac{23}{10^2}, \frac{14}{10^2}]$, but for $\rho = 3$, $u'_3(d)$ is calculated as follows:

$$\begin{aligned} u'_3(d) &= u'_3\left(\frac{2}{10}y * \frac{1}{10}\left(\frac{5}{10}y * \frac{5}{10}\underline{y}\right) * \frac{7}{10}\underline{y}\right) \\ &= \frac{2}{10}u_2(y) + \frac{1}{10}u_2\left(\frac{5}{10}y * \frac{5}{10}\underline{y}\right) + \frac{7}{10}u_2(\underline{y}) \\ &= \frac{2}{10}\left[\frac{83}{10^2}, \frac{83}{10^2}\right] + \frac{1}{10}\left[\frac{45}{10^2}, \frac{35}{10^2}\right] + \frac{7}{10}[0, 0] = \left[\frac{211}{10^3}, \frac{201}{10^3}\right] \end{aligned} \tag{35}$$

Here, d is strictly preferred to c . This holds for $k \geq 4$ since $u'_k(d)$ is calculated as $[\frac{2075}{10^4}, \frac{2075}{10^4}]$.

5.2 Proofs

Proof of (27) Suppose $g = [\bar{y}, \lambda; \underline{y}] \in B_k(\bar{y}; \underline{y})$. Let $g_t = [\bar{y}, \lambda_t; \underline{y}]$ for $t \leq \ell$. Since $g = \widehat{e} * \widehat{g}$, we have $\lambda = \sum_{t \leq \ell} \frac{1}{\ell} \lambda_t$. Since $\lambda_t \in \Pi_{k-1} \subseteq \Pi_k$ for $t \leq \ell$, we have $v_{k-1}(g_t) = \lambda_t v_0(\bar{y}) + (1 - \lambda_t)v_0(\underline{y}) = v_k(g_t)$ by (17). Since $v_k(g) = \lambda v_0(\bar{y}) + (1 - \lambda)v_0(\underline{y}) = (\sum_{t \leq \ell} \frac{1}{\ell} \lambda_t)v_0(\bar{y}) + (1 - (\sum_{t \leq \ell} \frac{1}{\ell} \lambda_t))v_0(\underline{y}) = \sum_{t \leq \ell} \frac{1}{\ell} (\lambda_t v_0(\bar{y}) + (1 - \lambda_t)v_0(\underline{y})) = \sum_{t \leq \ell} \frac{1}{\ell} v_{k-1}(g_t) = \widehat{e} * u_{k-1}(\widehat{g})$. \square

Proof of Theorem 5.1 Since (28) and (29) are dual, we consider only (28). Let $f \in B_k(\bar{y}; \underline{y}) \cup X$. Then, by $e0$, $\bar{u}_k(f) = \bar{v}_k(f)$. Consider the case where $f \in L_k(X) -$

$B_k(\bar{y}; \underline{y}) \cup X$. We prove $\bar{u}_k(f) = \min\{\widehat{e} * \bar{u}_{k-1}(\widehat{f}) : \widehat{f} \text{ is a decomposition of } f\}$. By *e2*, there is a decomposition \widehat{f} of f such that $\bar{u}_k(f) = \widehat{e} * \bar{u}_{k-1}(\widehat{f})$. Let \widehat{f}' be any decomposition of f . We show $\widehat{e} * \bar{u}_{k-1}(\widehat{f}') \geq \widehat{e} * \bar{u}_{k-1}(\widehat{f})$. By *e3*, for each $t \leq \ell$, there is a benchmark lottery $g'_t \in B_{k-1}(\bar{y}; \underline{y})$ such that $\bar{u}_{k-1}(\widehat{f}'_t) = u_{k-1}(g'_t)$. Let $\widehat{g}' = (g'_1, \dots, g'_\ell)$, which is a decomposition of $g' := \widehat{e} * \widehat{g}'$. By applying *e1* to $u_{k-1}(\widehat{g}') = \bar{u}_{k-1}(\widehat{f}')$, we obtain $u_k(g') \geq \bar{u}_k(f)$. Since $\widehat{e} * u_{k-1}(\widehat{g}') = u_k(g')$ by (27), we have $\widehat{e} * \bar{u}_{k-1}(\widehat{f}') = \widehat{e} * u_{k-1}(\widehat{g}') = u_k(g') \geq \bar{u}_k(f) = \widehat{e} * \bar{u}_{k-1}(\widehat{f})$.

For the latter assertion, we construct inductively $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ from $k = 0$ by (28) and (29). It is easy to see that $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ satisfies *e0* to *e2*. We need an inductive proof for *e3*. □

Proof of Theorem 5.2 (If): Let $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ be the extended utility stream satisfying *e0* to *e2* (and/or *e3*). Let \succsim_k be the binary relation over $L_k(X)$ defined by (30) and let its restriction to D_k be \sqsupseteq_k . Since \mathbf{v}_k the restriction of \mathbf{u}_k to $B_k(\bar{y}; \underline{y}) \cup X$ for each $k < \rho + 1$, $\langle \mathbf{v}_k \rangle_{k < \rho+1}$ satisfies *b0* to *b3* by Theorem 3.1. Thus, we have *E0*. It remains to show that $\langle \succsim_k \rangle_{k < \rho+1}$ satisfies *E1* to *E2* (and/or *E3*). Since the relation \geq_I over $\{\mathbf{u}_k(f) : f \in L_k(X)\}$ is transitive, *E2* is satisfied.

Consider *E1*. Let $\widehat{f} \in L_{k-1}(X)^\ell$ and $\widehat{g} \in B_{k-1}(\bar{y}; \underline{y})^\ell$ be decompositions of $f \in L_k(X)$ and $g \in B_k(\bar{y}; \underline{y})$. Suppose that $\widehat{f} \succsim_{k-1} \widehat{g}$, which implies $\mathbf{u}_{k-1}(\widehat{f}) \geq_I \mathbf{u}_{k-1}(\widehat{g})$ by (30).

Suppose $f \in B_k(\bar{y}; \underline{y}) \cup X$. Let $f = x \in X$. Then $\widehat{f} = (x, \dots, x)$. Then, $\underline{v}_{k-1}(x) \geq u_{k-1}(g_t)$ for all $t \leq \ell$, which, *b3*, implies $\underline{v}_k(x) \geq \underline{v}_{k-1}(x) \geq u_{k-1}(g_t)$ for all $t \leq \ell$. Thus, $\underline{v}_k(x) \geq \sum_{t \leq \ell} \frac{1}{\ell} u_{k-1}(g_t) = \widehat{e} * \bar{u}_{k-1}(\widehat{g})$. By (27), $\widehat{e} * \bar{u}_{k-1}(\widehat{g}) = u_k(g)$. Hence, by *b0*, $\underline{u}_k(x) = \underline{v}_k(x) \geq u_k(g)$. Thus, $f \succsim_k g$ by (30).

Suppose $f \in B_k(\bar{y}; \underline{y})$. By (27), $u_k(f) = \widehat{e} * \bar{u}_{k-1}(\widehat{f})$, and again, by (27), $\widehat{e} * \bar{u}_{k-1}(\widehat{f}) = u_k(g)$. Since $u_{k-1}(g_t) \geq u_{k-1}(g_t)$ for $t \leq \ell$, we have $u_k(f) \geq u_k(g)$. Thus, $f \succsim_k g$ by (30).

Finally, consider the case $f \in L_k(X) - B_k(\bar{y}; \underline{y}) \cup X$. Thus, $\widehat{e} * \underline{u}_{k-1}(\widehat{f}) \geq \widehat{e} * \mathbf{u}_{k-1}(\widehat{g}) = u_{k-1}(\widehat{e} * \widehat{g}) = u_k(g)$ by (27). By *e1*, we have $\underline{u}_k(f) \geq \widehat{e} * \mathbf{u}_{k-1}(\widehat{f}) \geq u_k(g)$. Hence, $\mathbf{u}_k(f) \geq_I \mathbf{u}_k(g)$; thus, $f \succsim_k g$ by (30). The other case of $\widehat{g} \succsim_{k-1} \widehat{f}$ is symmetric.

To prove Axiom *E3*, we prove the following assertion. Then, we obtain Axiom *E3* from *E3*($\langle \mathbf{u}_k \rangle_{k < \rho+1}$) by translating it in terms of the preference stream $\langle \succsim_k \rangle_{k < \rho+1}$.

E3 ($\langle \mathbf{u}_k \rangle_{k < \rho+1}$): Let $k < \rho + 1$ and $f, g \in L_k(X)$ with $\mathbf{u}_k(f) \geq_I \mathbf{u}_k(g)$. There is an $h \in B_k(\bar{y}; \underline{y})$ such that $\mathbf{u}_k(f) \geq_I \mathbf{u}_k(h) \geq_I \mathbf{u}_k(g)$. When $k \geq 1$, $\mathbf{v}_k(f) \geq_I \mathbf{v}_k(h)$ or f, h have decompositions \widehat{f}, \widehat{h} with $\mathbf{u}_{k-1}(\widehat{f}) \geq_I \mathbf{u}_{k-1}(\widehat{h})$, and the same holds for h, g .

Since $\underline{u}_k(f) \geq \bar{u}_k(g)$, by *e3**, there is an $h \in B_k(\bar{y}; \underline{y})$ such that $\underline{u}_k(f) = u_k(h)$. Hence, $\underline{u}_k(f) = u_k(h) \geq \bar{u}_k(g)$, i.e., $\mathbf{u}_k(f) \geq_I \mathbf{u}_k(h) \geq_I \mathbf{u}_k(g)$.

Let $k \geq 1$. Let $f \in B_k(\bar{y}; \underline{y}) \cup X$. By *e0*, $\mathbf{u}_k(f) = \mathbf{v}_k(f)$ and $\mathbf{u}_k(h) = \mathbf{v}_k(h)$ since $h \in B_k(\bar{y}; \underline{y})$. Hence, $\mathbf{u}_k(f) \geq_I \mathbf{u}_k(h)$ implies $\mathbf{v}_k(f) \geq_I \mathbf{v}_k(h)$. Now, suppose $f, g \notin B_k(\bar{y}; \underline{y}) \cup X$. We show that f, h have decompositions \widehat{f}, \widehat{h} such that $\mathbf{u}_{k-1}(\widehat{f}) \geq_I \mathbf{u}_{k-1}(\widehat{h})$. By *e2*, there is a decomposition of \widehat{f} of f such that $\widehat{e} * \underline{u}_{k-1}(\widehat{f}) = \underline{u}_k(f)$. By *e3*, there are $\underline{\lambda}_{f_1}, \dots, \underline{\lambda}_{f_\ell}$ in Π_k such that $\underline{u}_{k-1}(f_i) = \underline{\lambda}_{f_i}$ for

$t \leq \ell$. Let $h_t = [\bar{y}, \underline{\lambda}_{f_t}; \underline{y}]$ for $t \leq \ell$ and let $\hat{h} = (h_1, \dots, h_\ell)$. Since $\hat{f} \in L_{k-1}(X)^{\ell-1}$, $\hat{h} = (h_1, \dots, h_\ell)$ belongs to $L_{k-1}(X)^{\ell-1}$. Since $\underline{\lambda}_f = \underline{u}_k(f) = \hat{e} * \underline{u}_{k-1}(\hat{f}) = \sum \frac{1}{\ell} \cdot \underline{\lambda}_{f_t}$, we have $h = [\bar{y}, \underline{\lambda}_f; \underline{y}] = \sum_t \frac{1}{\ell} * h_t$. Hence, \hat{h} is a decomposition of h with $\underline{u}_{k-1}(\hat{f}) = \underline{u}_{k-1}(\hat{h})$. This implies $\underline{u}_{k-1}(\hat{f}) \geq_I \underline{u}_{k-1}(\hat{h})$.

(Only-if): Suppose that $\langle \succsim_k \rangle_{k < \rho+1}$ satisfies Axioms E0 to E2 (and/or E3) with its base preference stream $\langle \supseteq_k \rangle_{k < \rho+1}$. Then, $\langle \supseteq_k \rangle_{k < \rho+1}$ satisfies B0 to B3 by E0. Although this part can be proved using Theorem 5.1, the following direct proof is clearer.

We define $\langle \underline{u}_k \rangle_{k < \rho+1} = \langle [\bar{u}_k, \underline{u}_k] \rangle_{k < \rho+1}$ as follows: for each $f \in L_k(X)$,

$$\begin{aligned} \bar{u}_k(f) &= \min\{\lambda \in \Pi_k : [\bar{y}, \lambda; \underline{y}] \succsim_k f\}; \\ \underline{u}_k(f) &= \max\{\lambda \in \Pi_k : f \succsim_k [\bar{y}, \lambda; \underline{y}]\}. \end{aligned} \tag{36}$$

By Lemma 4.1.(2), these are well defined. Let $f, g \in L_k(X)$, $\underline{u}_k(f) := \underline{\lambda}_f$, and $\bar{u}_k(g) := \bar{\lambda}_g$. Then, $\underline{u}_k(f) \geq_I \underline{u}_k(g)$ if and only if $\underline{\lambda}_f \geq \bar{\lambda}_g$. It holds that $\underline{u}_k(f) \geq_I \underline{u}_k(g)$ if and only if $f \succsim_k g$; indeed, if $\underline{u}_k(f) \geq_I \underline{u}_k(g)$, then $f \succsim_k [\bar{y}, \underline{\lambda}_f; \underline{y}] \succsim_k [\bar{y}, \bar{\lambda}_g; \underline{y}] \succsim_k g$, i.e., $f \succsim_k g$ by E2, and conversely, if $f \succsim_k g$, then $\underline{u}_k(f) \geq \bar{u}_k(g)$ by (36).

When $f \in B_k(\bar{y}; \underline{y})$, we have $\bar{u}_k(f) = \lambda_f = v_k(f)$. Thus, e0 holds. Consider e1 to e2 (and e3).

e1 : Let $\hat{f} \in L_{k-1}(X)^\ell$. By (36) for $k - 1$, $\bar{u}_{k-1}(f_t)$ is written as $\lambda_{f_t} \in \Pi_{k-1}$ for all $t = 1, \dots, \ell$. Let $\hat{\lambda} = (\lambda_1, \dots, \lambda_\ell)$. Then, $\hat{e} * \bar{u}_{k-1}(\hat{f}) = \hat{e} * \hat{\lambda} \in \Pi_k$. By (36) for k , it holds that $\hat{e} * \bar{u}_{k-1}(\hat{f}) = \hat{e} * \hat{\lambda} \geq \bar{u}_k(\hat{e} * \hat{f})$. The other assertion that $\underline{u}_k(\hat{e} * \hat{f}) \geq \hat{e} * \underline{u}_{k-1}(\hat{f})$ is similarly proved.

e2 : Let $f \in L_k(X) - B_k(\bar{y}; \underline{y}) \cup X$. We prove that $\hat{e} * \bar{u}_{k-1}(\hat{f}) = \bar{u}_k(f)$ for some decomposition \hat{f} of f . The other half can be proved in the dual manner. By (36), we have $\bar{u}_k(f) = \bar{\lambda}_f \in \Pi_k$. Let $g := [\bar{y}, \bar{\lambda}_f; \underline{y}]$. By (36) and B1, g is the least preferred among $g' \in B_k(\bar{y}; \underline{y})$ with $g' \succsim_k f$, that is,

$$g' \succsim_k g \text{ for any } g' \in B_k(\bar{y}; \underline{y}) \text{ with } g' \succsim_k f. \tag{37}$$

Consider the preference $g \succsim_k f$. By (23), there are lotteries $g = h_0, h_1, \dots, h_m = f$ such that each (h_l, h_{l+1}) either belongs to \supseteq_k or is derived by E1. Since $h_m = f \in L_k(X) - B_k(\bar{y}; \underline{y}) \cup X$, the preference $h_{m-1} \succsim_k h_m = f$ is derived by E1, which implies $h_{m-1} \in B_k(\bar{y}; \underline{y})$. Since $g = h_0 \succsim_k h_{m-1}$ and $g \in B_k(\bar{y}; \underline{y})$, we have $\bar{\lambda}_f = \lambda_0 \geq \lambda_{m-1}$. However, since $\bar{\lambda}_f$ is given by the minimization (36), it is impossible that $\bar{\lambda}_f = \lambda_0 > \lambda_{m-1}$. Thus, $\lambda_0 = \bar{\lambda}_f = \lambda_{m-1}$, i.e., $h_0 = g = h_{m-1}$.

Since $h_0 = g = h_{m-1} \succsim_k f$ is derived by E1, there are decompositions \hat{g}, \hat{f} of g, \hat{f} such that $\hat{g} \succsim_{k-1} \hat{f}$ and $g = \hat{e} * \hat{g} \succsim_k \hat{e} * \hat{f} = f$. Then, by (36), each component f_t of \hat{f} has $g'_t \in B_{k-1}(\bar{y}; \underline{y})$ such that $\bar{u}_{t-1}(f_t) = u_{t-1}(g'_t) = \lambda_{g'_t}$. Since $g_t \succsim_{k-1} f_t$, we have $g_t \succsim_{k-1} g'_t$ by (36), which implies $u_{k-1}(g_t) = \lambda_{g_t} \geq \lambda_{g'_t} = u_{t-1}(g'_t) = \bar{u}_{t-1}(f_t)$. Since this holds for all $t \leq \ell$, we have

$$u_k(g) = \hat{e} * u_{k-1}(\hat{g}) \geq \hat{e} * u_{k-1}(\hat{g}') = \hat{e} * \bar{u}_{k-1}(\hat{f}). \tag{38}$$

On the other hand, each g'_t satisfies $g'_t \succsim_{k-1} f_t$ for all $t \leq \ell$ by definition. By E1, $\widehat{e} * \widehat{g}' \succsim_k \widehat{e} * \widehat{f} = f$. By (37), $\widehat{e} * \widehat{g}' \succsim_k g$. This implies $\widehat{e} * u_{k-1}(\widehat{g}') = u_k(\widehat{e} * \widehat{g}') \geq u_k(g)$, which implies that all the terms in (38) are identical. Thus, $\bar{u}_k(f) = u_k(g) = \widehat{e} * \bar{u}_{k-1}(\widehat{f})$.

e3 : By (36), $\bar{u}_k(x) = \lambda \in \Pi_k$ and $\underline{u}_k(x) = \lambda' \in \Pi_k$ for some λ and λ' in Π_k . Hence, $\bar{u}_k(x) = u_k([\bar{y}, \lambda; \underline{y}])$ and $\underline{u}_k(x) = u_k([\bar{y}, \lambda'; \underline{y}])$. This is the conclusion of e3. □

6 Measurability, comparability, and the EU hypothesis

Here, we study the concepts of measurable and non-measurable lotteries. Comparability and the EU hypothesis hold for measurable lotteries, while incomparability is intimately related to non-measurable lotteries. We assume that $\langle \succsim_k \rangle_{k < \rho+1}$ satisfies E0 to E3, relative to a base preference stream $\langle \succeq_k \rangle_{k < \rho+1}$ satisfying B0 to B3.

6.1 Measurable and non-measurable lotteries

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

$$M_k = \{f \in L_k(X) : f \sim_k g \text{ for some } g \in B_k(\bar{y}; \underline{y})\}. \tag{39}$$

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

Under our axioms, it holds that

$$\text{for each } f \in M_k, \text{ the probability weight } \lambda \text{ with } f \sim_k [\bar{y}, \lambda; \underline{y}] \text{ is unique,} \tag{40}$$

which we denote by λ_f . In M_k , no incomparabilities are observed; that is, if $f, g \in M_k$ with $\lambda_f \geq \lambda_g$, then $f \sim_k [\bar{y}, \lambda_f; \underline{y}] \succsim_k [\bar{y}, \lambda_g; \underline{y}] \sim_k g$. It also holds by Lemma 4.2.(1) that

$$M_k \subseteq M_{k+1} \text{ for all } k < \rho + 1. \tag{41}$$

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

Let $Y = \cup_{k < \infty} Y_k$. Each lottery $f \in L_\infty(Y) = \cup_{k < \infty} L_k(Y) = \cup_{k < \infty} L_k(Y_k)$ involves two types of depths, i.e., the measurement depth $\delta(\lambda_y)$ of $y \in Y_k$ with

$f(y) > 0$ and the depth $\delta(f(y))$ of the probability value $f(y)$. The sum of these two types gives a criterion for measurability, which is given in Theorem 6.1. Before it, we need Lemma 6.1. The first implies that any lottery $f \in M_k$ has a support in Y_k , which is the second.

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

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

For any $f \in L_\infty(X)$, we define k_f by

$$k_f = \begin{cases} \max\{\delta(\lambda_y) + \delta(f(y)) : f(y) > 0\} & \text{if } f \in L_\infty(Y) \\ \infty & \text{if } f \notin L_\infty(Y). \end{cases} \tag{42}$$

That is, if $f \in L_\infty(Y)$, then k_f is the maximum of the sum of $\delta(\lambda_y)$ and $\delta(f(y))$ over the support of f , and if $f \in L_\infty(X) - L_\infty(Y)$, then λ_y is undefined and k_f is assumed to take the value ∞ .

Theorem 6.1 (Measurability criterion). *Let $k < \rho + 1$ and $f \in L_\infty(X)$. Then,*

$$f \in M_k \text{ if and only if } k_f \leq k. \tag{43}$$

A lottery $f \in L_\infty(X) - L_\infty(Y)$ is neither measurable because of Lemma 6.1.(2), nor the right-hand side of (43) holds. When $f \in L_\infty(Y)$, (43) can be read in two ways. One is to fix a lottery $f \in L_\infty(Y)$ but to change (increase) k . Lottery f becomes measurable when k is large enough. For example, when $f = \frac{25}{10^2}y * \frac{75}{10^2}\underline{y}$ and $y \sim_2 [\bar{y}, \frac{83}{10^2}; \underline{y}]$, we have $k_f = \delta(\frac{83}{10^2}) + \delta(\frac{25}{10^2}) = 4$; by (43), $f \in M_k$ if and only if $4 \leq k$. The other reading is to fix a k and to change f . If $\delta(\lambda_y) > 0$ for some $y \in Y_k$, there is an $f \in L_k(Y_k)$ such that $\delta(\lambda_y) + \delta(f(y)) > k$; so $f \notin M_k$ by (43). Thus, non-measurable lotteries exist as long as $\{\bar{y}, \underline{y}\} \subsetneq Y_k$.

Incomparability \bowtie_k and indifference \sim_k may appear similar: indeed, Shafer (1986), p. 469, discussed whether \bowtie_k and \sim_k could be defined together and pointed out a difficulty from the constructive point of view. Theorem 6.2 gives a clear distinction between \sim_k and \bowtie_k . By E2, \sim_k is transitive, but \bowtie_k is not; indeed, we have distinct $f, h \in M_k$ with $f \sim_k h$, but by Theorem 6.2, for any $g \notin M_k$, $f \bowtie_k g$ and $g \bowtie_k h$. Also, reflexivity holds for the measurable domain M_k but not at all for $L_k(X) - M_k$.

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

(1) (No indifferences outside M_k): *If $f \notin M_k$, then $f \not\sim_k g$.*

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

6.2 EU hypothesis for measurable lotteries

Our theory is closely related to the expected utility hypothesis. It is explicitly assumed for the benchmark scale, i.e., B1 and b1. For the other part, it is only partially observed

by looking at conditions $e1, e2$ for $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ as well as Axiom E1 for preference relations $\langle \succsim_k \rangle_{k < \rho+1}$. In fact, the EU hypothesis holds for the measurable domain M_k , which is now shown.

Let $\langle \mathbf{u}_k \rangle_{k < \rho+1}$ be the extended utility stream satisfying $e0$ to $e3$, given Theorem 5.2, relative to a base utility stream $\langle \mathbf{v}_k \rangle_{k < \rho+1}$. It follows from Theorem 6.2.(2) and (30) that for any $f \in L_k(X)$,

$$\bar{u}_k(f) = \underline{u}_k(f) \text{ if and only if } f \in M_k. \tag{44}$$

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

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

Thus, the EU hypothesis holds for measurable lotteries, which gives a simple method of calculation of $u_k(f)$. On the other hand, by Theorem 6.2.(2) and (30), it holds that

$$\bar{u}_k(f) > \underline{u}_k(f) \text{ if and only if } f \in L_k(X) - M_k. \tag{45}$$

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

6.3 Proofs

Proof of Lemma 6.1 As mentioned, the latter assertion of (1) implies that any $f \in M_k$ has a support in Y_k , which is (2). We show (1) 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) for $k = 0$. Suppose the induction hypothesis that (1) hold for k . Now, we take any $f \in M_{k+1}$.

Suppose, on the contrary, that $0 < f(y_o) < 1$ for some $y_o \in Y_{k+1} - Y_k$ or $0 < f(y_o) \leq 1$ for some $y_o \in X - Y_{k+1}$. If $f(y_o) = 1$ and $y_o \in X - Y_{k+1}$, there is no $g \in B_{k+1}(\bar{y}; \underline{y})$ with $f \sim_{k+1} g$, a contradiction to $f \in M_{k+1}$. Hence, $0 < f(y_o) < 1$. By $f \in M_{k+1}$, we have a $g \in B_{k+1}(\bar{y}; \underline{y})$ with $f \sim_{k+1} g$. Since $0 < f(y_o) < 1$, it holds that $f \in L_{k+1}(X) - X$. Hence, E3 is applied to $f \sim_{k+1} g$ with the middle $h = g \in B_{k+1}(\bar{y}; \underline{y})$; we have decompositions \hat{f}, \hat{g} of f, g with $\hat{f} \succsim_k \hat{g}$. If one preference was strict, then $f \succ_{k+1} g$ by E1*, which is impossible; hence, $\hat{f} \sim_k \hat{g}$, which implies $f_t \in M_k$ for each t . Applying the induction hypothesis to f_t , we have $\hat{f} \in L_k(Y_k)^\ell$, and by E1, $f = \hat{e} * \hat{f} \in L_{k+1}(Y_k)$. This is impossible since $0 < f(y_o) < 1$ and $(y_o \in Y_{k+1} - Y_k \text{ or } y_o \in X - Y_{k+1})$. Hence, $f(y) = 0$ or 1 if $y \in Y_{k+1} - Y_k$ and $f(y) = 0$ if $y \in X - Y_{k+1}$. Now, we have (1) for $k + 1$. \square

Proof of Theorem 6.1 By (42) and Lemma 6.1.(2), it suffices to prove (43) for any $f \in L_\infty(Y)$. We prove (43) 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, (43) holds for $k = 0$. Now, suppose the induction hypothesis that (43) holds for k . We prove (43) for $k + 1$. In the following, let $f \in L_{k+1}(Y)$ with $\delta(f) > 0$.

Suppose $k_f \leq k + 1$. By the latter assertion of Lemma 2.1, there is a decomposition \widehat{f} of f so that $\delta(f_t(x)) < \delta(f(x))$ for all x with $\delta(f(x)) > 0$. Hence, $k_{f_t} = \max\{\delta(\lambda_x) + \delta(f_t(x)) : \delta(f_t(x)) > 0\} < \max\{\delta(\lambda_x) + \delta(f(x)) : \delta(f(x)) > 0\} \leq k_f$ for all t ; so, $k_{f_t} < k_f \leq k + 1$, which implies $k_{f_t} \leq k$ for all t . By the induction hypothesis, it holds that $f_t \in M_k$ for all t . Thus, for each t , we have $g_t \in B_k(\bar{y}, \underline{y})$ such that $f_t \sim_k g_t$. By E1, $f = \widehat{e} * \widehat{f} \sim_{k+1} \widehat{e} * \widehat{g} \in B_{k+1}(\bar{y}; \underline{y})$, which implies $f \in M_{k+1}$.

Consider the converse. Let $f \in M_{k+1}$. If $f \in M_k$, we have $k_f \leq k < k + 1$ by the induction hypothesis. Consider the case $f \in M_{k+1} - M_k$. Then, $f \sim_{k+1} g$ for some $g \in B_{k+1}(\bar{y}; \underline{y})$. By E3, there are decompositions \widehat{f}, \widehat{g} of f, g such that $\widehat{f} \sim_k \widehat{g}$. Thus, $\widehat{f} \in (M_k)^\ell \subseteq L_k(Y_k)^\ell$ by Lemma 6.1.(2). By the induction hypothesis, we have $\delta(\lambda_y) + \delta(f_t(y)) \leq k_{f_t} \leq k$ for all $y \in Y_k$ with $f_t(y) > 0$ and $t \leq \ell$. Since $f = \widehat{e} * \widehat{f}$, it holds that $\delta(f(y)) \leq \max_{t \leq \ell} \delta(f_t(y)) + 1$ for all $y \in Y_k$ with $f(y) > 0$. Thus, for $y \in Y_k$ with $f(y) > 0$, $\delta(\lambda_y) + \delta(f(y)) \leq \delta(\lambda_y) + \max_{t \leq \ell} \delta(f_t(y)) + 1 \leq \max_{t \leq \ell} k_{f_t} + 1 \leq k + 1$. Since $\delta(f(y)) = 0$ for $y \in Y_{k+1} - Y_k$ by Lemma 6.1.(1), we have $k_f = \max\{\delta(\lambda_y) + \delta(f(y)) : f(y) > 0\} \leq k + 1$. This means $k_f \leq k + 1$. \square

Proof of Theorem 6.2 (1): Let $f \notin M_k$. Suppose $f \sim_k g$. By E3, $f \succsim_k h \succsim_k g$ for some $h \in B_k(\bar{y}; \underline{y})$. By E2 (transitivity), $f \sim_k h \sim_k g$. This is impossible since $f \notin M_k$. Hence, $f \not\sim_k g$.

(2): The *if* part is by (39) and E2. The *only-if* part (contrapositive) follows from (1). \square

Proof of Theorem 6.3 Let $f \in M_k$. When $f \in B_k(\bar{y}; \underline{y}) \cup X$, we have, by $e0, \bar{u}_k(f) = \underline{u}_k(f) = v_k(f)$; if $f = [\bar{y}, \lambda; \underline{y}] \in B_k(\bar{y}; \underline{y})$, then, by $b1, v_k(f) = \lambda v_k(\bar{y}) + (1 - \lambda)v_k(\underline{y}) = E_f(v_k)$; and if $f = x \in X$, then $v_k(f) = 1 \times v_k(x) = E_f(v_k)$. Now, let $f \notin B_k(\bar{y}; \underline{y}) \cup X$. We show the assertion by induction on $k < \rho + 1$. The case $k = 0$ is included in the case $f \in B_k(\bar{y}; \underline{y}) \cup X, k < \rho + 1$. Suppose that the assertion holds for $k - 1$. Let $f \in M_k$. Then, $\delta(\bar{f}) > 0$, and by Lemma 6.1, $x \in Y_{k-1}$; so, $v_k(x) = v_{k-1}(x)$. Since $f \in M_k$, we have $f \sim_k h$ for some $h \in B_k(\bar{y}; \underline{y})$. By (23), this $f \sim_k h$ is derived by E1 from the decompositions \widehat{f}, \widehat{h} of f, h . By E1*, it holds that $\widehat{f} \sim_{k-1} \widehat{h}$. Hence, $\widehat{f} \in (M_{k-1})^\ell$. By the induction hypothesis, we have $\bar{u}_{k-1}(f_t) = \underline{u}_{k-1}(f_t) = E_{f_t}(v_{k-1})$ for $t \leq \ell$. By $e1, \widehat{e} * \bar{u}_{k-1}(\widehat{f}) \geq \bar{u}_k(f) \geq \underline{u}_k(f) \geq \widehat{e} * \underline{u}_{k-1}(\widehat{f})$; thus, these are all equal. Now, we have $u_k(f) = \widehat{e} * \underline{u}_{k-1}(\widehat{f}) = \sum_t \frac{1}{\ell} E_{f_t}(v_{k-1}) = E_f(v_k)$. \square

7 Toward the classical EU theory

We have focused on our theory from the constructive point of view. In particular, Theorem 4.1 reflects this constructiveness, which is extracted by Axiom E3 as well as condition $b3$. These are constraints on depths and interact with a finite cognitive bound. When we delete this bound, we go toward the classical EU theory. We have still two steps. The first step is to go to the case $\rho = \infty$, where the EU hypothesis holds in the exact form when the set of pure alternatives is restricted to $Y = \cup_{k < \infty} Y_k$, where $Y_k = M_k \cap X$ for $k < \infty$. The second is to allow all real number probabilities for lotteries, i.e., we take lotteries in $L_{[0,1]}(Y)$. There, the classical EU theory is a unique extension of our theory, while it contains highly non-constructive components.

7.1 Two steps to the classical EU theory

Throughout this section, we assume $\rho = \infty$. Let $\langle \succsim_k \rangle_{k < \infty}$ be a preference stream satisfying E0 to E2, relative to a base preference stream $\langle \succeq_k \rangle_{k < \infty}$ satisfying B0 to B3. Also, let $\langle \mathbf{v}_k \rangle_{k < \infty}$ be a base utility stream satisfying (19) of Theorem 3.1, where $\langle \mathbf{v}_k \rangle_{k < \infty}$ satisfies conditions $b0$ to $b3$. In this section, Axiom E3 and condition $e3$ are unnecessary.

We assume that for any $x \in X$, there is a k such that for any $k' \geq k$,

$$\bar{v}_k(x) = \bar{v}_{k'}(x) \text{ and } \underline{v}_k(x) = \underline{v}_{k'}(x). \tag{46}$$

That is, the upper and lower evaluations of x become constant after some k .⁹

Now, the *limit preference relation* of $\langle \succsim_k \rangle_{k < \infty}$ is defined to be $\succsim_\infty = \cup_{k < \infty} \succsim_k$; that is, the decision maker can go to any layer k for his preference comparisons. In parallel, the *limit utility function* of $\langle \mathbf{v}_k \rangle_{k < \infty}$ is defined by, for each $x \in X$,

$$\bar{v}_\infty(x) = \lim_{k \rightarrow \infty} \bar{v}_k(x) \text{ and } \underline{v}_\infty(x) = \lim_{k \rightarrow \infty} \underline{v}_k(x). \tag{47}$$

By (46), for each $x \in X$, there is a k_x such that $\bar{v}_k(x)$ and $\underline{v}_k(x)$ are constant for $k \geq k_x$. Also, we define $\mathbf{u}_\infty(f) = \lim_{k \rightarrow \infty} \mathbf{u}_k(f) = \lim_{k \rightarrow \infty} [\bar{u}_k(f), \underline{u}_k(f)]$ for any $f \in L_\infty(X)$.

We have the following theorem.

Theorem 7.1 (EU hypothesis without cognitive bounds).

- (1): For all $f \in L_\infty(X)$, $\mathbf{u}_\infty(f) = E_f(\mathbf{v}_\infty)$.
- (2): For all $f, g \in L_\infty(X)$, $f \succsim_\infty g$ if and only if $E_f(\mathbf{v}_\infty) \geq_I E_g(\mathbf{v}_\infty)$.

This theorem corresponds to Lemma 4.1 for $\rho < \infty$. Under (46), the limit preference relation \succsim_∞ is represented by the expectation of the vector-valued utility \mathbf{v}_∞ . Thus, without the cognitive restriction, the EU hypothesis holds for the limit relation \succsim_∞ even without completeness. Here, focusing on the limit utility function \mathbf{v}_∞ with (46), condition $e3$ becomes unnecessary and so does $E3$.

Let us restrict the set of pure alternatives X to $Y = \cup_{k < \rho+1} Y_k$. For each $x \in Y$, we can write $\bar{v}_\infty(x) = \underline{v}_\infty(x) = v_\infty(x)$ for $x \in Y$. Theorem 7.1.(2) is written as: for all $f, g \in L_\infty(Y)$,

$$f \succsim_\infty g \text{ if and only if } E_f(v_\infty) \geq E_g(v_\infty). \tag{48}$$

This is exactly the EU hypothesis in the classical sense.

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

⁹ Note that this can be expressed in terms of $\langle \succeq_k \rangle_{k < \infty}$.

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

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

$$f \succsim_E g \text{ if and only if } E_f(v_\infty) \geq E_g(v_\infty). \tag{49}$$

We have the following theorem.

Theorem 7.2 (Unique extension). *The relation \succsim_E defined by (49) is a unique extension of \succsim_∞ to $L_{[0,1]}(Y)$ with NM0 to NM2; that is,*

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

(2): \succsim_E satisfies NM0 to NM2.

This theorem is proved by the denseness of $L_\infty(Y)$ in $L_{[0,1]}(Y)$ and the continuity of $E_f(v_\infty)$ with respect to f relative to point-wise convergence, where $E_f(v_\infty)$ is continuous iff for any sequence $\{f^\nu\}$ in $L_{[0,1]}(Y)$ and $f \in L_{[0,1]}(Y)$, if $f^\nu(y) \rightarrow f(y)$ for each $y \in Y$, then $\lim_{\nu \rightarrow \infty} E_{f^\nu}(v_\infty) = E_f(v_\infty)$. The proof of the theorem may appear to be constructive, but the last extension step to \succsim_E is non-constructive, since probabilities newly involved in $f \in L_{[0,1]}(Y) - L_\infty(Y)$ may be given only in a non-constructive manner.¹⁰

7.2 Proofs

Proof of Theorem 7.1 (1): We show only $\bar{u}_\infty(f) = E_f(\bar{v}_\infty)$ and the other is parallel. Take an $f \in L_\infty(X)$, i.e., $f \in L_{k'}(X)$ for some k' . Let S be the support of f , which is a finite set. We choose a $k \geq k'$ by (46) such that for all $x \in S$, $\bar{v}_{k''}(x) = \bar{v}_k(x)$ for all $k'' \geq k$. It holds that for $l = 0, \dots$,

$$\bar{u}_{k+l}(h) = E_h(\bar{v}_{k+l}) \text{ for any } h \in L_l(S). \tag{50}$$

Once this is proved, this implies $\bar{u}_\infty(f) = E_f(\bar{v}_\infty)$; indeed, both sides become constant for large l ; so, we can choose one large k'' such that $\bar{u}_{k''}(h) = E_h(\bar{v}_{k''})$ for any $h \in L_{k''}(S)$. The above chosen f belongs to $L_{k''}(S)$. Hence, $\bar{u}_\infty(f) = E_f(\bar{v}_\infty)$.

Now, we show (50) by induction on $l = 0, \dots$. Let $l = 0$. Then, any $h \in L_0(S)$ is expressed as $h = x$ for some $x \in S$. By e0, $\bar{u}_k(h) = \bar{v}_k(x) = E_h(\bar{v}_k)$; so, (50) holds for $l = 0$. Suppose that (50) holds for l . Consider it for $l + 1$. Take $h \in L_{l+1}(S)$. By e2, there is a decomposition \hat{h} of h such that $\bar{u}_{k+l+1}(h) = \hat{e} * \bar{u}_{k+l}(\hat{h})$. Since each h_t belongs to $L_l(X)$, we have, by the induction hypothesis, $\bar{u}_{k+l}(h_t) = E_{h_t}(\bar{v}_l)$. Thus, $\bar{u}_{k+l+1}(h) = \hat{e} * \bar{u}_{k+l}(\hat{h}) = \sum_{t=1}^{\hat{\ell}} \frac{1}{\hat{\ell}} * E_{h_t}(\bar{v}_l) = E_h(\bar{v}_l)$. Since $\bar{v}_{l+1}(x) = \bar{v}_l(x)$ for all $x \in S$. Thus, we have $\bar{u}_{k+l+1}(\hat{h}) = E_h(\bar{v}_{l+1})$.

(2): By Theorem 5.2, for each k , we have some u_k representing \succsim_k . Let $f, g \in L_\infty(X)$. By definitions, $f \succsim_\infty g$ if and only if $f \succsim_k g$ for large k . Thus, $u_k(f) \geq_I u_k(g)$, which implies $u_\infty(f) \geq_I u_\infty(g)$. Tracing this argument back, we have the

¹⁰ We avoid the use of a topology for Axiom NM1. This does not change the content of classical EU theory as long as the set of lotteries is given as $L_{[0,1]}(Y)$. However, NM1 allows to restrict it to $L_{[0,1] \cap \mathbb{Q}}(Y)$. In this case, the extension result given in Theorem 7.2 is regarded as approximately constructive in the theoretical sense.

converse that $u_\infty(f) \geq_I u_\infty(g)$ implies $f \succsim_\infty g$. By (1) of this theorem, $E_f(v_\infty) = u_\infty(f) \geq_I u_\infty(g) = E_g(v_\infty)$. Hence, $f \succsim_\infty g$ if and only if $E_f(v_\infty) \geq_I E_g(v_\infty)$. \square

Proof of Lemma 7.1 Take any $f \in L_{[0,1]}(Y)$. This f has a support $S = \{y_0, y_1, \dots, y_m\}$ in Y with $f(y_t) > 0$ for $t \leq m$. We construct a sequence $\{g^v\}_{v=v_0}^\infty$ so that $g^v \in L_\infty(Y)$ for $v \geq v_0$, and for each $y \in Y$, $g^v(y) \rightarrow f(y)$ as $v \rightarrow \infty$. When $m = 0$, let $g^v = f$ for all $v \geq 0$. Now, we assume $m \geq 1$.

For any natural number v , let $z_{v,t} = \max\{\pi_t \in \Pi_v : \pi_t \leq f(y_t)\}$ for all $t \leq m$. There is a v_0 such that for all $v \geq v_0$, $\frac{1}{\ell^v} \leq z_{v,t} \leq 1 - \frac{1}{\ell^v}$ for all $t = 0, \dots, m - 1$ and $\frac{1}{\ell^v} \leq 1 - \sum_{t' < m} z_{v,t'} \leq 1 - \frac{m}{\ell^v}$. Also, we define $u_{v,0}, \dots, u_{v,m}$ by

$$u_{v,t} = \begin{cases} z_{v,t} & \text{if } t < m \\ 1 - \sum_{t' < m} z_{v,t'} & \text{if } t = m. \end{cases}$$

Then, $\sum_{t \leq m} u_{v,t} = 1$ and $u_{v,t} \in \Pi_v$ for all $t \leq m - 1$. Since $\frac{1}{\ell^v} \leq 1 - \sum_{t' < m} z_{v,t'} = u_{v,m} \leq 1 - \frac{m}{\ell^v}$, we have $u_{v,m} \in \Pi_v$.

We define $\{g^v\}_{v=v_0}^\infty$ by

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

Each g^v belongs to $L_v(Y)$. For $t \leq m - 1$, since $g^v(y_t) = u_{v,t} \leq f(y_t) < u_{v,t} + \frac{1}{\ell^v} = g^v(y_t) + \frac{1}{\ell^v}$ for all $v \geq v_0$, we have $\lim_{v \rightarrow \infty} g^v(y_t) = f(y_t)$. Since $g^v(y_m) - \frac{m}{\ell^v} = 1 - \sum_{t' < m} g^v(y_{t'}) - \frac{m}{\ell^v} \leq 1 - \sum_{t' < m} f(y_{t'}) = f(y_m) \leq 1 - \sum_{t' < m} g^v(y_{t'}) = g^v(y_m)$ for all $v \geq v_0$. Thus, $\lim_{v \rightarrow \infty} g^v(y_m) = f(y_m)$. \square

Proof of Theorem 7.2 (1): Let $f, g \in L_\infty(Y)$. Then, if k is large enough, then $f, g \in L_k(Y_k)$ and $v_k(x) = v_\infty(x)$ for all $x \in Y_k$. Now, suppose $f \succsim_\infty g$. Then $f \succsim_k g$, equivalently, $E_f(v_\infty) \geq E_g(v_\infty)$, which implies $f \succsim_E g$. Conversely, if $f \succsim_E g$, then $E_f(v_\infty) \geq E_g(v_\infty)$, equivalently, $f \succsim_k g$ for a large enough k . Thus, $f \succsim_\infty g$.

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

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

Now, let $f, g \in L_{[0,1]}(Y)$ with $f \succsim'_E g$. By Lemma 7.1, there are sequences $\{f^v\}$ and $\{g^v\}$ in $L_\infty(Y)$ such that they point-wise converge to f and g . As stated above,

$E_h(v_\infty)$ is continuous with respect to h . Then, $E_f(v_\infty) = \lim_{v \rightarrow \infty} E_{f^v}(v_\infty) \geq \lim_{v \rightarrow \infty} E_{g^v}(v_\infty) = E_g(v_\infty)$. We have shown that that for any $f, g \in L_{[0,1]}(Y)$, $E_f(v_\infty) \geq E_g(v_\infty)$ if and only if $f \succsim'_E g$. Thus, $f \succsim_E g$ if and only if $f \succsim'_E g$. \square

8 An application to a Kahneman–Tversky Example

We apply our theory to an experimental result reported in Kahneman and Tversky (1979). The experimental instance is formulated as Examples 3.1 and 5.1, and the relevant lotteries are $c = [\bar{y}, \frac{2}{10}; \underline{y}]$ and $d = \frac{25}{100}y * \frac{75}{100}\underline{y}$, which are incomparable for people with $\rho = 2$. It is the key how the observed behaviors are connected to the incomparabilities predicted in our theory. First, we look at the Kahneman–Tversky example, and then we make a certain postulate to have such a connection.

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

$$a = \left[4000, \frac{80}{10^2}; 0\right] \text{ (20\%)} \text{ vs. } b = 3000 \text{ with probability 1 (80\%)}$$

$$c = \left[4000, \frac{20}{10^2}; 0\right] \text{ (65\%)} \text{ vs. } d = [3000, \frac{25}{10^2}; 0] \text{ (35\%).}$$

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

$$0.80u(4000) + 0.20u(0) < u(3000)$$

$$0.20u(4000) + 0.80u(0) > 0.25u(3000) + 0.75u(0). \tag{51}$$

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

In Kahneman and Tversky (1979), no more information is mentioned other than the above percentages. Consider three possible distributions of the answers in terms of percentages over the four cases. In Table 3, the first, second, or third entry in each cell is the percentage derived by assuming 65%, 52%, or 45% for $b \wedge c$. The first 65% is the maximum possibility for $b \wedge c$, which leads to 0% for $a \wedge c$, and these determine the 20% for $a \wedge d$ and 15% for $b \wedge d$. The second entries are based on the assumption that the choices of b and c are stochastically independent, for example, $52 = (0.80 \times 0.65) \times 100$ for $b \wedge c$. In the third entries, 45% is the minimum possibility for $b \wedge c$. We interpret this table as meaning that each cell was observed at a significant level.

Let $\bar{y} = 4000$, $\underline{y} = 0$, $y = b = 3000$, and $\rho \geq 2$. Consider two cases A: $v_2(y) = u_2(y) = [\frac{77}{10^2}, \frac{77}{10^2}]$ and B: $v_2(y) = u_2(y) = [\frac{83}{10^2}, \frac{83}{10^2}]$ in Example 3.1, and recall that $u_2(a) = u_2([\bar{y}, \frac{80}{10^2}; \underline{y}]) = [\frac{80}{10^2}, \frac{80}{10^2}]$. Our theory predicts the choice a (or

Table 3 Three possible distributions

	<i>c</i> : 65%	<i>d</i> : 35%
<i>a</i> : 20%	<i>a</i> ∧ <i>c</i> : EU: 0//13//20	<i>a</i> ∧ <i>d</i> : paradox: 20//7//0
<i>b</i> : 80%	<i>b</i> ∧ <i>c</i> : paradox: 65//52//45	<i>b</i> ∧ <i>d</i> : EU: 15//28//35

b) in case *A* (*B*), independent of ρ . We assume that the distribution of subjects over *A* and *B* is the same as that given in Table 3, i.e.,

$$A : B = 20\% : 80\%. \tag{52}$$

We calculate the distribution of choices *c* and *d* based upon (52) and the distribution of ρ .

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

Let $\rho = 3$. In case *A*, (34) states $\mathbf{u}_3(c) = [\frac{2}{10}, \frac{2}{10}] \geq_I \mathbf{u}_3(d) = [\frac{199}{10^3}, \frac{189}{10^3}]$, and in case *B*, (35) states $\mathbf{u}_3(d) = [\frac{211}{10^3}, \frac{201}{10^3}] \geq_I \mathbf{u}_3(c) = [\frac{2}{10}, \frac{2}{10}]$. Hence, people with $\rho = 3$ behave in the same manner as those with $\rho \geq 4$, though *d* is non-measurable.

In case $\rho = 2$. (33) states that people in cases *A* and *B* show the same base utility evaluation of *d*, i.e., $\mathbf{u}_2(d) = [\frac{23}{10^2}, \frac{14}{10^2}]$. Since $\mathbf{u}_2(c) = [\frac{2}{10}, \frac{2}{10}]$, *c* and *d* are incomparable for these people.

Here, we find a conflict between our theory and the reported experimental result in that every subject chose one lottery in each of the above choice problems, while our theory states that *c* and *d* are incomparable for people with $\rho = 2$. The issue is how a subject behaves for the choice problem when the lotteries are incomparable for him. In such a situation, a person would typically be forced (e.g., following social customs) to make a choice.¹¹ Here, we assume the following postulate for choice behavior for a subject having incomparabilities:

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

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

¹¹ It may be difficult for people to show incapability of answering a question if it appears linguistically and logically clear. The present author knows only one person in our profession to refuse consciously to answer such a question. Davis and Maschler (1965), Sec.6 reported that when a number of game theorists/economists were asked about their predictions of choices in a specific example in a cooperative game theory, Martin Shubik refused to answer a questionnaire. He mentioned the reason that the specification in terms of

Table 4 Ratios of choices *c* and *d*

	<i>A</i>	<i>B</i>
$\rho = 2$	$c : d = 2 : 1$	$c : d = 2 : 1$
$\rho \geq 3$	$c : d = 1 : 0$	$c : d = 0 : 1$

Table 5 $r_2 : r_{+3} = 9 : 1$

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

Table 6 $r_2 : r_{+3} = 8 : 2$

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

Table 4 summarizes the above calculated results. To see the relationship between Tables 3 and 4, we specify the distribution of people over $\rho = 2, 3, \dots$. We consider two distributions of ρ

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

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

In the case $r_2 : r_{+3} = 9 : 1$, the percentage of the choices $a \wedge c$ is calculated as $100 \times \frac{2}{10} \times (\frac{9}{10} \times \frac{2}{3} + \frac{1}{10} \times 1) = 14\%$. The corresponding percentages $b \wedge c$ is calculated as $100 \times \frac{8}{10} \times (\frac{9}{10} \times \frac{2}{3} + \frac{2}{10} \times 0) = 48\%$. Thus, we obtain Table 5. Table 6 is based on $r_2 : r_{+3} = 8 : 2$.

The results in Tables 5 and 6 are quite compatible with Table 4. Perhaps, we should admit that this is based upon our specifications of parameter values as well as Postulate BH. To make stronger assertions, we need to think about more cases of parameter values and different forms of BH. Nevertheless, this study may lead to observations on new aspects on bounded rationality that ρ seems quite small.

Remark 8.1 (Common ratio effect). The anomaly mentioned in (51) is often called the ‘‘common ratio effect’’ (cf. Prelec 1998; van de Kuilen and Wakker 2006, and their references). It refers to the observation such as the fact that the opposite of the second inequality in (51) is obtained from the first with multiplication of $1/4 = 25/10^2$. In our theory for case *B* with $\rho = 2$, *b* is strictly preferred to *a*, but *c* and *d*, which are obtained by the multiplication, are incomparable, and the independence condition, NM2, is violated. We made the additional postulate BH to connect incomparability to the observed behavior in the experiment. The postulate shows a bigger tendency to

cooperative game is not enough to have a precise prediction for the question. Usually, people answer such a question, often unconsciously by filling up gaps.

choose c . In this sense, our result shows the “common ratio effect”. However, Postulate BH does not directly take depths for the choice behavior of agents. Perhaps, there are different postulates taking depths of lotteries to explain the “common ratio effect” more directly. This is an open problem (see also Sect. 9, [c]).

9 Conclusions

We have developed the EU theory with probability grids and preference formation. The permissible probabilities are restricted to the form of ℓ -ary fractions up to a given cognitive bound ρ . We divide the argument into the measurement step of preferences and utilities on pure alternatives in terms of the benchmark scale and the extension step to lotteries with more risks. We take the constructive view point of the decision maker for our theory. The development includes the approach in terms of vector-valued utilities with the interval order due to Fishburn (1970). The connections between the preference approach and the utility approach are shown to be equivalent in Sects. 3 to 5. These approaches are complementary; each may give better interpretations as well as some technical merits over the other.

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

In Sect. 8, we applied the incomparability results to the Allais paradox, specifically to an experimental example in Kahneman and Tversky (1979). Our theory is compatible with their experimental result; incomparabilities involved for $\rho = 2$ are crucial in interpreting their result.

We considered a few aspects of bounded rationality in terms of probability grids and cognitive bounds for EU theory; bounded rationality is more salient with shallow ρ . When, however, we consider a specific decision problem, other aspects of bounded rationality may manifest themselves. We should have more researches on the aspects of bounded rationality in various directions. Here, we give a few possible research agenda.

The first three are related to bounded rationality.

[a] Constructive method of particular preferences: We presented our theory following Table 1 to derive all the preferences in a layer from the previous layer. However, the decision maker may think about his preferences more locally focusing only on the target lotteries and involved pure alternatives and relevant probabilities. This question could enable us to think about complexities of preference formation. It may give a better understanding of how much bounded rationalities are involved when only target lotteries are concerned.

[b] Preference formation in inductive game theory (IGT): This theory studies experimental sources for individual knowledge/belief about the structure of the society

(cf. Kaneko and Matsui 1999). Our approach has some parallelism to the constructive approach to IGT, due to Kline et al. (2019). In particular, Kaneko and Kline (2015) study the other person's preferences from experiences of the other's position through role-switching. However, since it is assumed that experiences include numerical utility values, their treatment does not capture the partial understanding/non-understanding of the other's preferences/desires. Perhaps, lack of full experiences is closely related to incomparability in our theory.

This is also related to the case-based decision theory by Gilboa (2001) as well as to the frequentist interpretation of probability in the context of the classical EU theory (cf., Hu 2013). The former concerns evaluations of probabilities for causality (course-effect) from experiences, and the latter is about probability as frequency of an event. Bounded memory capacity of a person is relevant for both. Our theory with probability grids and cognitive bounds may give a suggestion to analyze such problems.

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

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

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

One is a *vertical extension*: we take another pair of benchmarks $\bar{\bar{y}}$ and $\underline{\underline{y}}$ such as $\bar{\bar{y}} \succeq_0 \bar{y} \succeq_0 \underline{y} \succeq_0 \underline{\underline{y}}$. The relation between the original system and the new system is not simple. In the case of measurement of temperatures, the grids for the Celsius system do not exactly correspond to those in the Fahrenheit system. We need multiple bases ℓ for probability grids and may have multiple preference systems even for similar target problems.

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

[e]: Extensions of the probability grids Π_ρ : The above extensions may require more subtle treatments of probability grids. A possibility is to extend Π_ρ to $\cup_{\ell=2}^{\bar{\ell}} \Pi_\ell$; that is, probability grids having the denominators $\ell \leq \bar{\ell}$ are permissible. Then, the

Celsius and Fahrenheit systems of measuring temperatures are converted from each to the other. A question is how large $\bar{\ell}$ is required for such classes of problems.

[f]: Subjective probability: Our theory is almost directly applied to Anscombe and Aumann’s (1963) theory of subjective probability and subjective utility. An event E such as tomorrow’s weather is evaluated asking an essentially the same question as (1) in Sect. 1. We could have an extension of our theory including the subjective probability theory. It could be difficult to have an extension corresponding to Savage (1954), since no benchmark scale is assumed; perhaps, Savage’s theory is not in our scope.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

We prepare the extension δ^* of the depth measure δ to $\Pi_k^* = \{\pi : \pi = v/\ell^k \text{ and } v \text{ is a nonnegative integer}\}$, where v may be larger than ℓ^k . This Π_k^* is closed with respect to $+$. For $\pi \in \Pi_k^*$, we define $\delta^*(\pi) = k$ iff $k \in \Pi_k^* - \Pi_{k-1}^*$. Then, $\delta^*(\pi) = \delta(\pi)$ if $\pi \in \Pi_k$. The following facts will be used in the proof of Lemma 2.1:

$$\begin{aligned} \text{if } \pi = s + v'/\ell^k \text{ for an integer } s \text{ and } v' < \ell^k, \text{ then } \delta^*(\pi) &= \delta^*(v'/\ell^k); & (53) \\ \text{if } \delta^*(\pi), \delta^*(\pi') \leq k, \text{ then } \delta^*(\pi + \pi') &\leq k. & (54) \end{aligned}$$

Proof of Lemma 2.1 ¹² Let $k \geq 1$. We show that if $f \in L_k(X)$, then $f = \widehat{e} * \widehat{f}$ for some $\widehat{f} \in L_{k-1}(X)^\ell$ with the depth constraint $\delta(f(x)) > \delta(f_t(x))$ for all $t \leq \ell$ and $x \in X$ with $\delta(f(x)) > 0$. We assume $\delta(f) = k$. Let $\{x_1, \dots, x_m\}$ be the support of f with $f(x_t) > 0$ for $t = 1, \dots, m$. Since $\delta(f) = k \geq 1$, we have $m \geq 2$.

Notice that f is regarded as the list $(x_1, v_1/\ell^k, \dots, x_m, v_m/\ell^k)$ with $0 \leq v_t < \ell^k$ for all $t \leq m$ and $\sum_{t=1}^m v_t = \ell^k$. This is expressed as:

$$f = \overbrace{(x_1, 1/\ell^k, \dots, x_1, 1/\ell^k)}^{v_1 \text{ times}}, \dots, \overbrace{(x_m, 1/\ell^k, \dots, x_m, 1/\ell^k)}^{v_m \text{ times}}, \quad (55)$$

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

$$[[y_1^1, \dots, y_{\ell^{k-1}}^1], \dots, [y_1^\ell, \dots, y_{\ell^{k-1}}^\ell]]. \quad (56)$$

¹² The author is indebted to a referee for improving this proof significantly.

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

$$\begin{aligned}
 f_1 &= [y_1^1, 1/\ell^{k-1}, \dots, y_{\ell^{k-1}}^1, 1/\ell^{k-1}], \\
 &\vdots \\
 f_\ell &= [y_1^\ell, 1/\ell^{k-1}, \dots, y_{\ell^{k-1}}^\ell, 1/\ell^{k-1}].
 \end{aligned}
 \tag{57}$$

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

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

$$[z_\xi : \bar{v}_{t-1} < \xi \leq \bar{v}_t] \text{ with weight } 1/\ell^k.
 \tag{58}$$

This corresponds to the fragment with x_t in (55). Thus, z is partitioned in two ways: $[[y_1^1, \dots, y_{\ell^{k-1}}^1], \dots, [y_1^\ell, \dots, y_{\ell^{k-1}}^\ell]]$ of (56) and $[[z_\xi : \bar{v}_{t-1} < \xi \leq \bar{v}_t] : t = 1, \dots, m]$ of (58); we call $[y_1^{t'}, \dots, y_{\ell^{k-1}}^{t'}]$ and $[z_\xi : \bar{v}_{t-1} < \xi \leq \bar{v}_t]$ *fragments*. The fragments of these partitions may have three types of non-empty intersections:

- (a): $[y_1^{t'}, \dots, y_{\ell^{k-1}}^{t'}]$ is a subfragment of $[z_\xi : \bar{v}_{t-1} < \xi \leq \bar{v}_t]$;
- (b): $[y_1^{t'}, \dots, y_{\ell^{k-1}}^{t'}]$ is a superfragment of $[z_\xi : \bar{v}_{t-1} < \xi \leq \bar{v}_t]$;
- (c): one of them starts inside the other but does not stop in it.

In (a), $f_{t'}(x_t) = 1$, since $y_1^{t'} = \dots = y_{\ell^{k-1}}^{t'} = x_t$. Hence, $\delta(f_{t'}(x_t)) = 0 < \delta(f(x_t))$. In (b), since the length of $[z_\xi : \bar{v}_{t-1} < \xi \leq \bar{v}_t]$ is v_t , we have $f_{t'}(x_t) = v_t/\ell^{k-1}$ and $f(x_t) = v_t/\ell^k$. Since $\delta(v_t/\ell^{k-1}) < \delta(v_t/\ell^k)$, we have $\delta(f_{t'}(x_t)) < \delta(f(x_t))$.

Consider (c), which is neither (a) nor (b). Suppose that $[z_\xi : \bar{v}_{t-1} < \xi \leq \bar{v}_t]$ ends inside $[y_1^{t'}, \dots, y_{\ell^{k-1}}^{t'}]$ but it starts in a previous fragment. Then, $f_{t'}(x_t) = v'_t/\ell^{k-1}$ for some v'_t with $0 < v'_t < \ell^{k-1}$. In fact, $\sum_{s \leq t} v_s = (t' - 1)\ell^{k-1} + v'_t$. By (53) and (54),

$$\delta^*(v_t/\ell^{k-1}) \geq \delta^*(\sum_{s \leq t} v_s/\ell^{k-1}) = \delta^*((t' - 1) + v'_t/\ell^{k-1}) = \delta^*(v'_t/\ell^{k-1}).$$

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

References

Anscombe, F.J., Aumann, R.J.: A definition of subjective probability. *Ann. Math. Stat.* **34**(1), 199–205 (1963)

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

- Aumann, R.J.: Utility theory without the completeness axiom. *Econometrica* **30**, 445–462 (1962)
- Blume, L., Easley, D., Halpern, J.Y.: Constructive decision theory, Working Paper (2013)
- Davis, M., Maschler, M.: The kernel of a cooperative game. *Naval Res. Logist. Quart.* **12**(3), 223–259 (1965)
- Dubra, J., Ok, E.A.: A model of procedural decision making in the presence of risk. *Int. Econ. Rev.* **43**, 1053–1080 (2002)
- Dubra, J., Maccheroni, F., Ok, E.A.: Expected utility theory without the completeness axiom. *J. Econ. Theory* **115**(1), 118–133 (2004)
- Fishburn, P.: Intransitive indifference with unequal indifference intervals. *J. Math. Psychol.* **7**, 144–149 (1970)
- Fishburn, P.: One-way expected utility with finite consequence space. *Ann. Math. Stat.* **42**(2), 572–577 (1971)
- Fishburn, P.: Alternative axiomatizations of one-way expected utility. *Ann. Math. Stat.* **43**(5), 1648–1651 (1972)
- Fishburn, P.: *The Foundations of Expected Utility*. D. Reidel Publishing Co., London (1982)
- Gilboa, I., Schmeidler, D.: *A Theory of Case-Based Decisions*. Cambridge University Press, Cambridge (2001)
- Herstein, I.B., Minor, J.: An axiomatic approach to measurable utility. *Econometrica* **21**(2), 291–297 (1953)
- Hu, T.: Expected utility theory from the frequentist perspective. *Econ. Theory* **53**(1), 9–25 (2013). <https://doi.org/10.1007/s00199-009-0482-9>
- Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. *Econometrica* **47**(1), 263–292 (1979)
- Kaneko, M.: The ordered field property and a finite algorithm for the Nash bargaining solution. *Int. J. Game Theory* **20**(3), 227–236 (1992)
- Kaneko, M., Kline, J.J.: Understanding the other through social roles. *Int. Game Theory Rev.* **17**(1), 1540005–1–31 (2015). <https://doi.org/10.1142/S0219198915400058>
- Kaneko, M., Matsui, A.: Inductive game theory: discrimination and prejudices. *J. Public Econ. Theory* **1**(1), 101–137 (1999)
- Kaneko, M., Nakamura, K.: The Nash social welfare function. *Econometrica* **47**(2), 423–435 (1979)
- Kline, J.J., Lavendhomme, T., Waltener, S.: From memories to inductively derived views: a constructive approach. *Econ. Theory* **68**(1), 403–420 (2019). <https://doi.org/10.1007/s00199-018-1129-5>
- Kontek, K., Lewandowski, M.: Range-dependent utility. *Manag. Sci.* (2018). <https://doi.org/10.1287/mnsc.2017.2744>
- Mendelson, E.: *Numbers Systems and Foundations of Analysis*. Krieger, Malabar (1973)
- Nash, J.F.: The bargaining problem. *Econometrica* **18**(2), 155–162 (1950)
- Prelec, D.: The probability weighting function. *Econometrica* **66**(3), 497–528 (1998)
- Rubinstein, A.: Similarity and decision making under risk (Is there a utility theory resolution to the Allais paradox?). *J. Econ. Theory* **46**(1), 145–153 (1988)
- Savage, L.J.: *The Foundations of Statistics*. Wiley, New York (1954)
- Shafer, G.: Savage revisited. *Stat. Sci.* **1**(4), 463–501 (1986)
- Shafer, G.: Constructive decision theory. *Int. J. Approx. Reason.* **79**(C), 45–62 (2016)
- Simon, H.A.: Rational choice and the structure of the environment. *Psychol. Rev.* **63**(2), 129–138 (1956)
- Simon, H.A.: *Reason in Human Affairs*. Stanford University Press, Stanford (1983)
- Turing, A.: On computable numbers, with an application to the Entscheidungsproblem. *Proc. Lond. Math. Soc.* **s2-42**(1), 230–265 (1937)
- van de Kuilen, G., Wakker, P.: Learning in the Allais Paradox. *J. Risk Uncertain.* **33**(3), 155–164 (2006)
- Von Neumann, J., Morgenstern, O.: *Theory of Games and Economic Behavior*, 3rd edition (1953), Princeton University Press, Princeton 1st edn. (1944)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.