Decision Making Under Multidimensional Risk

Shaowei Ke* Mu Zhang†‡

May 19, 2024

Abstract

Choice alternatives are often multidimensional and risky, but how to model a decision maker's evaluation of them is unclear. Three popular approaches are in sharp contrast: One aggregates all dimensions before evaluating risk, one does it reversely, and one evaluates each dimension's risk recursively. We characterize a hierarchical expected utility model that generalizes these approaches. The decision maker's preference reveals how she brackets and orders the dimensions, based on which she evaluates risk recursively. We analyze the model's uniqueness properties and characterize several special cases. We study the model's implications in contexts of group inequality, multisource income, and time lotteries.

*Department of Economics, China Europe International Business School. Email: shaoweike@ceibs.edu.

†Department of Economics, University of Michigan. Email: muzhang@umich.edu.

‡This paper subsumes an earlier paper titled "Multidimensional Choices under Uncertainty" by the same authors, and partially subsumes Zhang (2023). We are grateful to Tilman Börgers, David Dillenberger, Amanda Friedenberg, Faruk Gul, David Kreps, Fabio Maccheroni, David Miller, Xiaosheng Mu, Efe Ok, Pietro Ortoleva, Wolfgang Pesendorfer, Debraj Ray, Ran Spiegler, and seminar participants at numerous seminars and conferences for helpful comments.

1 Introduction

Decision makers often face a variety of alternatives that are complex and uncertain. The evaluation of such an alternative must take into account multiple dimensions of the alternative, as well as the associated risk. For instance, the decision maker may need to evaluate a product that has uncertain attributes; she may evaluate a job opportunity that generates an uncertain sequence of future payoffs; or she may assess a policy that induces an uncertain income distribution for multiple individuals. Although evaluating risky multidimensional alternatives is a fundamental and ubiquitous task in economics, there are ongoing debates regarding how to do it, and no principles that can reliably elucidate a general solution.

To illustrate, consider the following example. Let (x_1, x_2) denote the incomes of individuals 1 and 2, respectively. A policymaker is assessing a policy that will lead to $(0, 1)$ and $(1, 0)$ with equal probability. She does not like inequality. Let $u(\cdot, \cdot)$ be a concave and symmetric function. To evaluate the policy, she may first use u to evaluate each possible income profile, and then take the expectation: $\frac{1}{2}u(0, 1) + \frac{1}{2}u(1, 0)$. It is well known that this approach in general captures ex post inequality aversion.[1] Alternatively, she may first compute each individual's expected income, which is $1/2$, and then use u to evaluate the profile of expected incomes: $u(1/2, 1/2)$. This approach captures ex ante inequality aversion.[2] These two approaches to evaluating multidimensional risk are both desirable but also incompatible.[3]

One might assume that what we derive from this example is specific to inequality aversion, but the same issue occurs in many different contexts, including in the evaluation of risky consumption bundles, dynamic

[1]This formula captures ex post inequality aversion because $(1/2, 1/2)$ is better than having $(0, 1)$ and $(1, 0)$ with equal probability. Ex post inequality is also called inequality of outcome.

[2]This formula captures ex ante inequality aversion because having $(0, 1)$ and $(1, 0)$ with equal probability is better than having either $(0, 1)$ or $(0, 1)$ with certainty. Ex ante inequality is also called inequality of opportunity. This formula does not capture ex post inequality because it does not distinguish between having $(0, 1)$ and $(1, 0)$ with equal probability and having exactly $(1/2, 1/2)$ with certainty, but the former has ex post inequality while the latter does not. Conversely, the previous formula is linear in probability, so having $(0, 1)$ and $(1, 0)$ with equal probability is as good as having either $(0, 1)$ or $(1, 0)$ with certainty, which renders it unable to capture ex ante inequality.

[3]See, among others, Fleurbaey (2010), Fudenberg and Levine (2012), Grant et al. (2010, 2012), and Saito (2013).

choice, and ambiguity aversion.[4] The common theme is that there are two opposite approaches to evaluating a risky multidimensional alternative: One first aggregates across dimensions for each possible realization and then takes expectation (called the *first-aggregation-then-expectation* (FATE) approach henceforth), and the other first takes expectation for each dimension and then aggregates across dimensions (called the *first-expectation-then-aggregation* (FETA) approach henceforth). They both often seem reasonable, yet have behavioral implications that stand in sharp contrast.

However, the evaluation of a risky multidimensional alternative is not limited to those two approaches. They evaluate all dimensions of the alternative simultaneously, whether before or after taking expectation. Yet the decision maker does not necessarily need to do so. Consider a decision maker who is facing uncertainty over the ratings of a car's safety and driving experience. She may want to consider every possible realization of the safety rating, conditioning on which she then takes into account the (conditional) uncertainty over the driving experience.[5]

This third approach is similar to how, in Kreps and Porteus (1978) and Epstein and Zin (1989), the decision maker evaluates risk recursively. Consider a two-period example. For every realization of her period-1 consumption x_1, the decision maker evaluates the conditional expected utility of period-2 consumption given x_1, denoted by U_{x_1}. She aggregates x_1 and U_{x_1} in a possibly nonadditive way. Then, she takes the expectation with respect to x_1.

Clearly, the decision maker may also evaluate multidimensional risk in ways that differ from these three approaches. For instance, if she is facing uncertainty over the ratings of a car's safety, driving experience, and cost, she may consider every realization of the safety rating, conditioning on which she may adopt the FETA approach to evaluate the uncertainty over the driving experience and cost.

In this paper, we introduce the *hierarchical expected utility* (HEU) representation of the decision maker's preference, which offers a versatile framework capable of accommodating a wide range of behavioral patterns about

[4] See Online Appendix I for a detailed discussion of these examples.

[5] See Keeney (1973), Zhang (2023), and Li et al. (2023) for similar ideas.

the evaluation of multidimensional risk. It also serves as a unifying framework that encompasses the three approaches as special cases. From the decision maker's choice behavior, we can identify how she brackets and orders the dimensions when evaluating a risky multidimensional alternative. This is described by a *hierarchy*, which is a collection of sets of dimensions such that any two of them are either nested or disjoint. The HEU representation has one main equation that will be applied iteratively to derive an evaluation of a risky multidimensional alternative.

We discuss some choice behavior that can be captured by the HEU representation. We illustrate how we can elicit the hierarchy from choice behavior, and provide an axiomatic characterization of the HEU representation. We analyze in what sense the HEU representation can be uniquely identified: Roughly speaking, the Bernoulli indices in the representation are unique, as in expected utility theory, and a *canonical* hierarchy, a hierarchy that (i) highlights the separability (across dimensions) properties of the decision maker's preference and (ii) avoids labeling redundant conditioning evaluation of risk, can be uniquely identified.

Within the HEU framework, we characterize the three commonly used extreme cases: the FATE representation, the FETA representation, and the recursive representation. We then derive two generalizations of those special cases. First, the generalized bracketing representation features a hierarchy that is essentially a partition of the dimensions. Each cell of the partition is a bracket. The decision maker jointly evaluates the risk over all dimensions within each bracket, and then aggregates across brackets.[6] We discuss how this representation relates to group inequality. Second, the generalized recursive representation allows the decision maker to evaluate risk recursively according to a weak order on the dimensions. By comparison, the recursive representation implies a linear order on the dimensions.

We apply the HEU representation to two problems. In the first application, different dimensions represent different income sources. We assume that the decision maker's preference has a generalized bracketing representation. Within each bracket, she rationally evaluates the risk of the total income and computes a certainty equivalent. Across different brackets,

[6] The trivial partition and the finest partition (every dimension is a cell of the partition) correspond to the FATE representation and the FETA representation, respectively.

however, she simply sums the certainty equivalents of all brackets, failing to consider the interdependence of income risk. We call such behavior narrow bracketing. We characterize the circumstances under which narrow bracketing leads the decision maker to choose a distribution of total income that is stochastically dominated, and when the decision maker prefers to avoid multidimensional risk. The second application features a setting with uncertainty over what and when prizes will be delivered. As emphasized by DeJarnette et al. (2020), the widely used exponentially discounted expected utility model exhibits risk-seeking behavior in the time dimension. This is neither natural nor consistent with experimental findings. We examine how the HEU representation can address various issues pointed out by DeJarnette et al. (2020).

1.1 Related Literature

Many papers have studied multivariate risk, but most stay within expected utility theory and focus on analyzing measures of risk attitude (see Duncan (1977), Eeckhoudt et al. (2007), Grant (1995), Karni (1979), Keeney (1973), Kihlstrom and Mirman (1974, 1981), Levy and Levy (1991), Richard (1975), and Schlee (1990)). Some papers deviate from expected utility theory, but in a way that is more in line with classic non-expected utility analyses (see, for example, Karni (1989)). Our approach is complementary to the above. The HEU representation does not necessarily satisfy independence, and when it violates independence, it is because of the interaction between the evaluation of risk and how the decision maker brackets and orders dimensions—rather than, for example, the Allais paradox.

The FETA representation can capture full narrow bracketing and correlation neglect,[7] and the HEU representation allows for more general bracketing and correlation preferences. Our paper generalizes Zhang (2023) from two dimensions to multiple dimensions and allows history-dependent risk attitudes in the recursive model. By studying computationally tractable choices, Camara (2021) characterizes a dynamic choice bracketing model that generalizes full narrow bracketing in a way different from ours. Com-

[7]See Camara (2021), Ellis and Freeman (2021), Enke and Zimmermann (2019), Levy and Razin (2015), Read et al. (1999), Thaler (1985), Tversky and Kahneman (1981), Vorjohann (2023), and Zhang (2023).

pared with our analysis, independence is maintained in Camara (2021).

Our paper offers a novel solution to two other long-standing problems in the literature of inequality aversion and the literature of ambiguity aversion. In both cases, there are two well-known incompatible approaches to model inequality aversion and ambiguity aversion: the ex ante approach and the ex post.[8] The two approaches have rather different behavioral implications. In the context of inequality, they capture inequality of opportunity (ex ante inequality) and inequality of outcome (ex post inequality) respectively. In the context of ambiguity, the ex ante approach predicts that randomization cannot hedge away the effect of ambiguity, while the ex post approach predicts the opposite. Some papers have attempted to address such incompatibility. For example, Saito (2013, 2015) characterizes representations that are weighted averages of the two approaches, and Ke and Zhang (2020) generalize Saito (2015). The HEU representation offers a novel way to resolve the conflict between the ex ante and the ex post approaches. Under a special case of the HEU representation—the generalized bracketing representation—within each bracket, it is as if the decision maker takes the ex ante approach, but across different brackets, it is as if the decision maker takes the ex post approach.

Our paper is related to the literature on dynamic preferences. As stated previously, DeJarnette et al. (2020) point out that the exponentially discounted expected utility model implies risk-seeking behavior in the time dimension, which is neither natural nor consistent with experimental findings. If we cast their solution (see also Kihlstrom and Mirman (1981) and Dillenberger et al. (2020)) to this problem to our setting, the solution first aggregates across dimensions (different time periods) via exponential discounting. Then, a Bernoulli index is applied to the aggregation before the decision maker takes expectation to evaluate the risk. This is similar to a special case of our FATE representation. Other papers have proposed opposite approaches, such as Selden (1978) and Selden and Stux (1978), that are similar to special cases of the FETA representation. Our paper provides a unifying framework that nests these approaches and the recursive approach (Epstein and Zin, 1989; Kreps and Porteus, 1978) as special

[8]For ambiguity aversion, see Baillon et al. (2022), Dominiak and Schnedler (2011), Ke and Zhang (2020), Oechssler et al. (2019), Raiffa (1961), and Saito (2015).

cases. Notably, in Section 6.2 we show that the recursive approach can generate a novel solution.

If we view different dimensions in our setup as potentially different sources of uncertainty, our paper is also related to Ergin and Gul (2009) and Cappelli et al. (2021). In both papers, the decision maker's risk attitude may be source-dependent, and she may evaluate risk source-wise before across sources, which conceptually is similar to our generalized bracketing representation. However, different from our representation, how the decision maker brackets the states into sources in those papers is exogenous. Chew et al. (2023) also study a source-dependent extension of expected utility theory. However, different sources are captured by different mixture operators, rather than dimensions.

2 Setup and Representation

For an arbitrary set Z, let $\Delta(Z)$ denote the set of all simple lotteries (probability measures with a finite support) on Z. Let $I = \{1, \dots, N\}$ be a finite set of integers with $N > 1$. For every $i \in I$, let $X_i = [\underline{x}_i, \overline{x}_i]$ be a nondegenerate bounded closed interval in $\mathbb{R}$. Let $X = \bigtimes_{i \in I} X_i$. Generic elements of X are called consequences. Generic elements of $\Delta(X)$ are called lotteries.

Fix any $A \subseteq I$. Let $X_A = \bigtimes_{i \in A} X_i$. We use x, y, z to denote generic elements of X_A and p, q, r, s to denote generic elements of $\Delta(X_A)$.[9] We denote $p \in \Delta(X_A)$ that yields $x \in X_A$ with certainty by δ_x. When there is no risk of confusion, we identify δ_x with x, and identify a subscript or a superscript $A \subseteq I$ with i if $A = \{i\}$ and with $-i$ if $A = \{i\}^{\mathsf{c}}$.[10] For any $p, q \in \Delta(X_A)$ and $\alpha \in [0, 1]$, we write $p\alpha q$ as shorthand for the convex combination $\alpha p + (1 - \alpha)q \in \Delta(X_A)$.

Marginal and conditional distributions. For any $A \subseteq B \subseteq I$ and $p \in \Delta(X_B)$, we use $p_A \in \Delta(X_A)$ to denote p's marginal distribution on A, and use $x_A \in X_A$ to denote the restriction of $x \in X_B$ to A. For any $A, B \subseteq C \subseteq I$ such that $A \cap B = \emptyset$, $p \in \Delta(X_C)$, and $x \in X_B$, let $p_{A|x} \in \Delta(X_A)$ denote p's conditional marginal distribution on A given x, and let $\mathbb{E}^p_{A|x}$ denote the expectation operator under distribution $p_{A|x}$. We

[9] However, only elements in X and $\Delta(X)$ are called consequences and lotteries.
[10] For any $A \subseteq I$, the set A^{c} is its complement in I.

write $\mathbb{E}^p_A$ if $B = \emptyset$ and $\mathbb{E}^p$ if $A = C$ and $B = \emptyset$. For any disjoint subsets of I, $A_1, \ldots, A_n$, with $\bigcup_{i=1}^n A_i = A \subseteq I$, and any $p_i \in \Delta(X_{A_i})$ for every $i \in \{1, \ldots, n\}$, we use $(p_1, \ldots, p_n)$ to denote the unique $q \in \Delta(X_A)$ such that $q(x) = p_1(x_{A_1}) \times \cdots \times p_n(x_{A_n})$ for every $x \in X_A$.

Preference. The decision maker has a preference $\succsim$ over $\Delta(X)$. Its asymmetric and symmetric parts are denoted by $\succ$ and $\sim$, respectively. For any nonempty $A \subseteq I$ and $x \in X_{A^c}$, we define the conditional preference $\succsim_x$ on $\Delta(X_A)$ such that for any $p, q \in \Delta(X_A)$, we have $p \succsim_x q \iff (p, x) \succsim (q, x)$. We define $\succ_x$ and $\sim_x$ similarly.

2.1 Examples: Two Dimensions and Beyond

Before formally defining the HEU representation of $\succsim$, we introduce some examples to elucidate its structure. Consider a father who is deciding whether to take his child on a Disney ride. We begin with two dimensions, the first measuring how thrilling the ride is, and the second measuring how entertaining the ride is for two people together rather than one alone. If the ride is too thrilling, the father will feel nauseated during or after the ride but the child will be unaffected. The three commonly used approaches to evaluate multidimensional risk are as follows: For any lottery p,

- FATE: $U(p) = \mathbb{E}^p\, u(x_1, x_2)$.
- FETA: $U(p) = u(\mathbb{E}^p_1\, v(x_1),\ \mathbb{E}^p_2\, w(x_2))$.
- Recursive: $U(p) = \mathbb{E}^p_2\, u(x_2, \mathbb{E}^p_{1|x_2} v_{x_2}(x_1))$.[11]

The FATE approach corresponds to expected utility, which is well understood. The FETA approach corresponds to a utility function that aggregates separate evaluation of each dimension's risk. It describes a father who evaluates the risk of each measure independently of the other, which is an example of full narrow bracketing and correlation neglect.

The formula for the recursive approach is well understood in the context of dynamic choice, but less so in a static setting. In this particular example, it captures a father who considers every possible realization of his and his child's experience x_2, and given any x_2 he dislikes (if v_{x_2} is concave)

[11] Alternatively, we can have $U(p) = \mathbb{E}^p_1\, u(x_1, \mathbb{E}^p_{2|x_1} v_{x_1}(x_2))$.

facing additional uncertainty about whether he will feel nauseated or not. More specifically, suppose $p = \frac{1}{2}\delta_{(0,0)} + \frac{1}{2}\delta_{(1,1)}$, $q = \frac{1}{2}\delta_{(1,0)} + \frac{1}{2}\delta_{(0,1)}$. Then, $p\frac{1}{2}q = \frac{1}{4}\delta_{(0,0)} + \frac{1}{4}\delta_{(0,1)} + \frac{1}{4}\delta_{(1,0)} + \frac{1}{4}\delta_{(1,1)}$. With this formula, it is possible that

$$p \sim q \succ p\frac{1}{2}q.$$

This is because with r, given any realization of dimension 2, there is still conditional uncertainty over dimension 1—but with p, q, this is not the case.

Figure 1 naturally represents how the father evaluates the two dimensions in the three examples above, respectively. As will become clear later, they correspond to the hierarchies of the three approaches.[12]

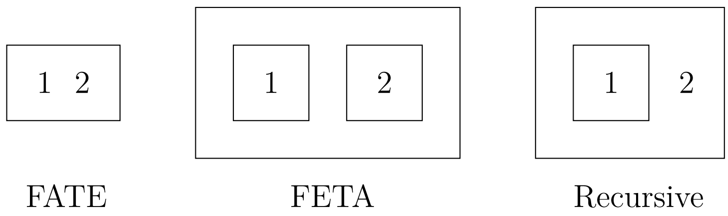

Figure 1: The three approaches.

When we have more than two dimensions, it becomes clear that these three approaches can only capture some rather special cases of a decision maker's evaluation of multidimensional risk. For example, suppose that in addition to the two dimensions described above, a third dimension of the Disney ride measures how well the ride connects to the stories. In this case, we may encounter a father who wants to separately consider dimension 1 and dimensions 2 and 3, because these two subsets of dimensions describe properties of different natures; a father who considers the conditional uncertainty of dimension 1 given each joint realization of dimensions 2 and 3, following the same logic of the recursive approach; or a father who considers every realization of dimension 1 and given each realization, he applies the FETA approach to evaluate dimensions 2 and 3. Figure 2 provides a graphic illustration of these examples.

[12]It may appear that the boxes in Figure 1 have different meanings. For example, the outer box for FETA does not involve an expectation operator, while the others do. In fact, they all correspond to equation (1), except that the expectation of the outer box for FETA is degenerate.

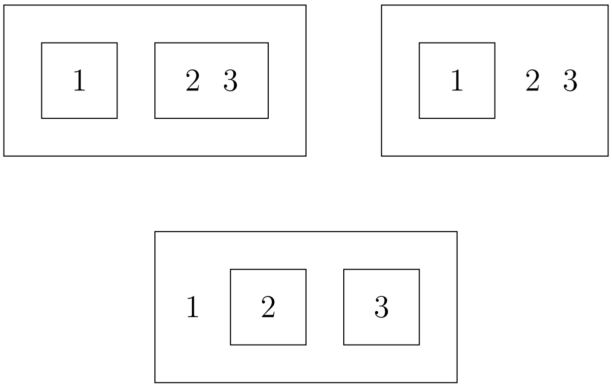

Figure 2: The top-left figure captures separate evaluations of dimension 1 and dimensions 2 and 3. The top-right figure captures the evaluation of the risk of dimensions 2 and 3 and the conditional risk of dimension 1 given each realization of dimensions 2 and 3. The bottom figure captures the evaluation of dimension 1's risk and the conditional risk of dimensions 2 and 3 separately given each realization of dimension 1.

2.2 Representation

The HEU representation of the decision maker's preference can capture the behavior described in all figures of Section 2.1. Those figures are graphical illustrations of hierarchies. A hierarchy is an important ingredient of the HEU representation. A collection of nonempty subsets of I, denoted by $\mathcal{H}$, is a hierarchy if (i) $I \in \mathcal{H}$ and (ii) for any $A, B \in \mathcal{H}$, we have $A \subseteq B$, $B \subseteq A$, or $A \cap B = \emptyset$.[13] We call the elements of a hierarchy *components*.

To define the HEU representation, we need a few operators for components. For any $i \in I$, let $H(i)$ denote the smallest component in $\mathcal{H}$ that contains i. Clearly, $H(i)$ is uniquely defined for every $i \in I$. For any $A \in \mathcal{H}$, let $\tau(A) = \{i \in A : A = H(i)\}$; $\eta(A) = \{i \in I : A \subsetneq H(i)\}$; and $\Phi(A) = \{B \in \mathcal{H} : B \subsetneq A$, and there does not exist any $B' \in \mathcal{H}$ such that $B \subsetneq B' \subsetneq A\}$. It is immediate from the definitions that $\tau(A) = A \setminus \bigcup_{B \in \Phi(A)} B$.

To understand the interpretation of τ, η, and Φ, consider dynamic choice as an analogy and envision the dimensions as time periods. Then, it is as if τ and η identify dimensions that are currently evaluated (present) and dimensions whose risk has been resolved previously (history), respectively, and Φ identifies the components whose risk will be resolved later (future). To see this more concretely, we apply τ, η, and Φ to the top-left and bottom

[13] Hierarchies can be equivalently represented using trees.

examples in Figure 2.

Remark 1. *From here on, to simplify the discussion, we may use the terms "present/current," "past/history," and "future" to refer to* τ, η*, and* Φ*, respectively, even when the context is not dynamic choice.*

Example 1. *In the top-left example in Figure 2, let* $A = \{2, 3\}$ *and* $B = \{1\}$*. Then,* $\mathcal{H} = \{I, A, B\}$ *is a hierarchy. We have* $\tau(I) = \emptyset$*;* $\tau(A) = \{2, 3\}$*; and* $\tau(B) = \{1\}$*. In this example, the decision maker does not evaluate any dimension conditioning on another, so* $\eta(I) = \eta(A) = \eta(B) = \emptyset$*. Finally,* $\Phi(I) = \{A, B\}$ *and* $\Phi(A) = \Phi(B) = \emptyset$*.*

Example 2. *In the bottom example in Figure 2, let* $A = \{2\}$ *and* $B = \{3\}$*. Then* $\mathcal{H} = \{I, A, B\}$ *is a hierarchy. For component* I*, the present dimension is* $\tau(I) = \{1\}$*,* I*'s history* $\eta(I)$ *is empty, and* I*'s future is* $\Phi(I) = \{A, B\}$*. For component* A*, the present dimension is* $\tau(A) = \{2\}$*, the history is* $\eta(A) = \{1\}$*, and the future* $\Phi(A)$ *is empty. Similarly, for component* B*,* $\tau(B) = \{3\}$*,* $\eta(B) = \{1\}$*, and* $\Phi(B)$ *is empty.*

Several notational conventions are useful to understand the definition of the HEU representation. First, if we encounter $x \in X_A$ in which $A = \emptyset$, then x will be ignored in the expression. We illustrate what this convention implies in the following examples:

- For any $p_{A|x}$ in which $x \in X_B$ and $B = \emptyset$, we identify $p_{A|x}$ with p_A.
- For any f_x in which $x \in X_A$ and $A = \emptyset$, we identify f_x with f.
- For any $A \in \mathcal{H}$, $\mathcal{A} \subseteq \mathcal{H}$, and $f : X_A \times \mathbb{R}^{\mathcal{A}} \to \mathbb{R}$, if $A = \emptyset$, then f's domain is identified with $\mathbb{R}^{\mathcal{A}}$.

Second, in the last example, if $\mathcal{A} = \emptyset$ instead, then f's domain is identified with X_A. Last, if we have $p_{A|x}$ in which $A = \emptyset$, the conditional expectation operator $\mathbb{E}^p_{A|x}$ will be ignored in the expression.

Definition 1. *The preference has an HEU representation if there exist a hierarchy* $\mathcal{H}$ *and functions* $u^A : X_{\eta(A)} \times X_{\tau(A)} \times \mathbb{R}^{\Phi(A)} \to \mathbb{R}$ *for every* $A \in \mathcal{H}$ *such that, recursively defining for all* $A \in \mathcal{H}$ *and* $x \in X_{\eta(A)}$ *the function* $U^A_x : \Delta(X_A) \to \mathbb{R}$ *by*

$$U^A_x(p) = \mathbb{E}^p_{\tau(A)}\, u^A(\, x,\ y,\ (U^B_{(x,y)}(p_{B|y}))_{B \in \Phi(A)}\,), \tag{1}$$

the following statements hold for all $A \in \mathcal{H}$ and $x \in X_{\eta(A)}$:

1. *For any $p, q \in \Delta(X_A)$ and $z \in X_{(A \cup \eta(A))^c}$, we have $p \succsim_{(x,z)} q \iff U^A_x(p) \geqslant U^A_x(q)$.*
2. *The function $U^A_x(\delta_z)$ is continuous and strictly increasing in $z \in X_A$.*

We denote the HEU representation by $(\mathcal{H}, (u^A)_{A \in \mathcal{H}})$.

The first condition in the above definition implies the usual representation condition: For any lotteries $p, q \in \Delta(X)$, we have $p \succsim q$ if and only if $U^I(p) \geqslant U^I(q)$. Hence, the function U^I (U^A_x when $A = I$) represents $\succsim$ and is derived recursively, similar to Kreps and Porteus (1978). The first condition also implies representation conditions for conditional preferences. The second condition requires that the representations of conditional preferences are continuous and monotone in the absence of risk. This condition is useful in the proofs and may be relaxed. It ensures that the set of consequences is sufficiently rich so that for any lottery, we can construct lotteries with different supports that are indifferent to it.

Although the HEU representation appears complex, its main idea is simple and similar to Kreps and Porteus (1978). Take any component A and $x \in X_{\eta(A)}$ on which evaluation of the dimensions in A conditions. We can understand equation (1) that evaluates $p \in \Delta(X_A)$ as follows:

$$U^A_x(p) = \underbrace{\mathbb{E}^p_{\tau(A)}}_{\substack{\text{expectation with} \\ \text{respect to risk in} \\ \text{the present dimensions}}} u^A(x, \underbrace{y}_{\substack{y \in X_{\tau(A)}, \\ \text{the present} \\ \text{dimensions}}}, \underbrace{(U^B_{(x,y)}(p_{B|y}))_{B \in \Phi(A)}}_{\substack{\text{conditional expected} \\ \text{utility of futures} \\ \text{given } (x,y)}}),$$

in which x as an argument of u^A plays the role of allowing for history dependence. Figure 3 illustrates how the HEU representation is applied recursively to the top-left and top-right examples in Figure 2.

One way to interpret the hierarchy is that it captures how the decision maker analyzes risk. Risk is fully described by the joint distribution over X, but the decision maker has her own way of understanding risk: She subjectively decomposes the joint distribution into conditional distributions and/or marginal distributions iteratively. The rationale for the decomposition depends on what the specific dimensions represent. For example, sometimes it may capture narrow bracketing (see Sections 2.1 and 6.1), and sometimes it may capture the decision maker's subjective perception of time periods (see Section 5). Based on the decomposition, the decision maker then evaluates risk recursively according to equation (1). Her risk attitudes are described by the Bernoulli indices of the HEU representation.

One might wonder whether decomposing risk as described above and then analyzing risk recursively will be computationally more complex than the expected utility model. First, for instance, if the decision maker divides a subset of dimensions into several brackets for evaluation, she will ignore the lottery's correlation across different brackets (see the discussion of FETA in Section 2.1), which may simplify the computation.[14] It is also not clear whether conditioning necessarily complicates the computation. In terms of the formula, it appears more complex, but a more formal conclusion will require some analysis of computational complexity, which is beyond the scope of this paper.

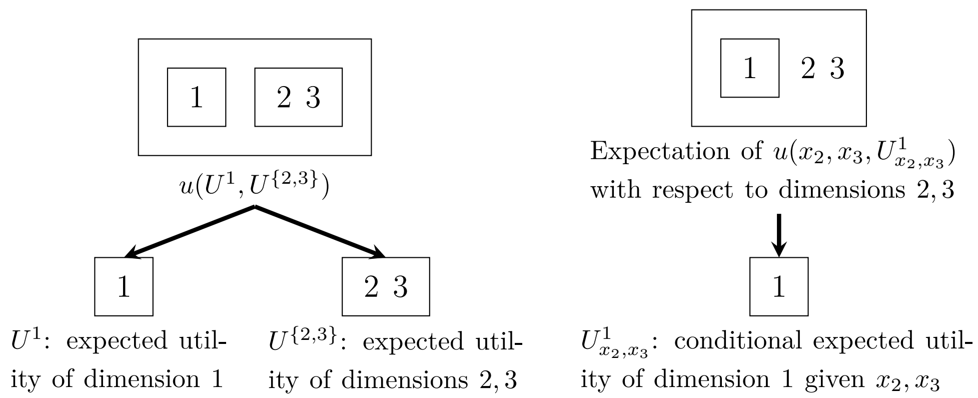

Figure 3: A graphical illustration of how the HEU representation is applied recursively to the top-left and top-right examples in Figure 2.

On the one hand, equation (1) can be viewed as an extension of the recursive expected utility formula in the dynamic choice literature (Kreps and Porteus, 1978; Epstein and Zin, 1989), modified for our setting. On the other hand, there are several important differences between our approach and the typical approach used in dynamic choice.

First, in the dynamic choice literature, the choice alternatives them-

[14]See Camara (2021) for a formal discussion of the connection between narrow bracketing and computational tractability.

selves are often constructed recursively, as in Kreps and Porteus (1978), following an exogenous linear order on the dimensions such as time. This structure simplifies the characterization of recursive models. As will be seen later, identifying the subjective recursive evaluation of risk is more challenging in our case. By contrast, because our theory applies to many other settings in which the dimensions are not naturally ordered and we seek to learn from her behavior how the decision maker structures the dimensions, we choose not to define our choice domain $\Delta(\bigtimes_{i \in I} X_i)$ recursively. Although the alternatives are not recursively constructed, our approach still allows the decision maker's behavior to follow recursive models.

Second, unlike standard recursive models in dynamic choice, the HEU representation is not necessarily continuous on $\Delta(X)$. The reason for the lack of continuity will be clear when we introduce the continuity axiom in Section 3.

3 Axiomatic Foundation

Before presenting the axiomatic characterization of the HEU representation, we first use some examples to illustrate how we can elicit the decision maker's hierarchy from her choices. These examples will help us understand the motivation for our main axiom.

3.1 Elicitation of the Hierarchy

Consider the simplest example in which $N = 2$ and suppose the decision maker's preference has an HEU representation. There are four possible hierarchies: $\mathcal{H}^1 = \{\{1,2\}\}; \mathcal{H}^2 = \{\{1,2\},\{1\}\}; \mathcal{H}^3 = \{\{1,2\},\{2\}\}; \mathcal{H}^4 = \{\{1,2\},\{1\},\{2\}\}$. As discussed in Section 2.1, the hierarchy $\mathcal{H}^1$ corresponds to expected utility, which means that the decision maker's choices are consistent with independence. It turns out that if the decision maker's hierarchy is either $\mathcal{H}^2$, $\mathcal{H}^3$, or $\mathcal{H}^4$, her choices may violate independence, and different hierarchies lead to different types of violations. This enables us to infer the decision maker's hierarchy by observing how her choices violate independence.

3.1.1 Violations of Independence Due to Conditioning

If the decision maker decides that dimension j's risk should be evaluated conditioning on dimension i's realization, independence applied to dimensions i and j may not hold, as shown in the example below.

Example 3. *Suppose the decision maker's utility function is*

$$U(p) = \sum_{x_1} p_1(x_1) u(x_1, p_{2|x_1}),$$

in which $u(x_1, p_2) = x_1 + (\sum_{x_2} p_2(x_2)\sqrt{x_2})^2$*. That is, the utility of* (x_1, p_2) *is the sum of* x_1 *and the certainty equivalent of* p_2 *under the Bernoulli index* $w(z) = \sqrt{z}$ *for* $z \geqslant 0$*. The corresponding hierarchy is* $\mathcal{H}^3 = \{\{1,2\},\{2\}\}$*.*

Consider three degenerate lotteries that yield $x = (0,0)$*,* $y = (-1/2, 1/4)$ *and* $z = (0,4)$*, respectively. We can verify that* $U(\delta_x) = 0 > U(\delta_y) = -0.25$*, but* $U(\delta_y \frac{1}{2} \delta_z) = 1.875 > U(\delta_x \frac{1}{2} \delta_z) = 1$*, which violates independence.*

The reason behind the violation of independence shown in Example 3 is as follows. Because $y = (-1/2, 1/4)$ and $z = (0,4)$ do not share the same value in dimension 1, when the decision maker evaluates $\delta_y \frac{1}{2} \delta_z$, she thinks of it as follows: with 50% chance, $-1/2$ in dimension 1, conditioning on which $1/4$ in dimension 2; and with 50% chance, 0 in dimension 1, conditioning on which 4 in dimension 2. However, because $x = (0,0)$ and $z = (0,4)$ share the same value in dimension 1, when it comes to $\delta_x \frac{1}{2} \delta_z$, the decision maker thinks of it as follows: with certainty 0 in dimension 1, conditioning on which $\delta_0 \frac{1}{2} \delta_4$ in dimension 2. For $\delta_y \frac{1}{2} \delta_z$, the decision maker's risk attitude for dimension 2 is never factored in, but for $\delta_x \frac{1}{2} \delta_z$, the decision maker will compute the certainty equivalent of $\delta_0 \frac{1}{2} \delta_4$ in dimension 2.

In this example, the violation of independence is caused by the fact that one mixture mixes lotteries that do not have overlapping supports in dimension 1, while the other mixes lotteries that do. Put differently, when we only mix lotteries that do not have overlapping supports in dimension 1, we should not expect independence to be violated. This observation leads to the following definition that *tentatively* identifies which set of dimensions is evaluated conditioning on which dimension. For any $i \in I$ and $p_i, r_i \in \Delta(X_i)$, we write $p_i \perp r_i$ if they have disjoint supports.

Definition 2. *For any* $i \in I$ *and* $A \subseteq I$*, we write* $i \rightharpoonup A$ *if for all* $\alpha \in (0,1)$*,* $x \in X_{(\{i\}\cup A)^c}$*, and* $p,q,r,s \in \Delta(X_{\{i\}\cup A})$ *such that* $p_i \perp r_i$ *and* $q_i \perp s_i$*, we have* $p \succ_x q$ *and* $r \sim_x s \implies p\alpha r \succ_x q\alpha s$*. We write* $i \rightharpoonup j$ *if* $A = \{j\}$*.*

In the above definition, the requirements $p_i \perp r_i$ and $q_i \perp s_i$ ensure that the mixtures $p\alpha r$ and $q\alpha s$ do not have overlapping supports in dimension i. The idea is that if, by avoiding mixing alternatives that have overlapping supports in dimension i, independence holds on dimensions $\{i\} \cup A$, then it is possible that the decision maker evaluates dimensions in A conditioning on dimension i.

Several observations should be noted. First, if the decision maker's conditional preference on $\Delta(X_A)$ always has an expected utility representation, then $i \rightharpoonup A$ for all $i \in A$. Second, if $i \rightharpoonup A$, then $i \rightharpoonup A \cup \{i\}$ and $i \rightharpoonup B$ for all $B \subseteq A$. Third, if we focus on the case in which A is a singleton set, then $\rightharpoonup$ induces a binary relation on I.

In Example 3, we can verify that $1 \rightharpoonup 2$ but $2 \not\rightharpoonup 1$, which is consistent with its hierarchy $\mathcal{H}^3 = \{\{1,2\},\{2\}\}$.[15] Hence, it appears that $\rightharpoonup$ precisely identifies how the decision maker conducts conditional risk evaluation. This is not true, however, as shown in the following example in which $N = 3$.

Example 4. *Let* $N = 3$*. Suppose the decision maker's utility function is*

$$U(p) = \sum_{x_1} v_1(x_1)p_1(x_1) \cdot \sqrt{\sum_{x_2} v_2(x_2)^2 p_{2|x_1}(x_2) \cdot \left[\sum_{x_3} v_3(x_3)p_{3|x_1,x_2}(x_3)\right]^2}.$$

The corresponding hierarchy is $\{I,\{2,3\},\{3\}\}$*. Intuitively, we should have* $1 \rightharpoonup 2$*,* $1 \rightharpoonup 3$*, and* $2 \rightharpoonup 3$*, and not conversely. However,* $3 \rightharpoonup 1$ *also holds. To see this, fixing any* $x_2 \in X_2$*, for every* $p_{\{1,3\}} \in \Delta(X_{\{1,3\}})$*, we have*

$$\begin{aligned} U((p_{\{1,3\}}, x_2)) &= \sum_{x_1} v_1(x_1)p_1(x_1) \cdot v_2(x_2) \cdot \sum_{x_3} v_3(x_3)p_{3|x_1}(x_3) \\ &= v_2(x_2) \cdot \mathbb{E}^p_{\{1,3\}}(v_1(x_1) \cdot v_3(x_3)), \end{aligned}$$

which means that focusing on dimensions 1 *and* 3*, the conditional preference has an expected utility representation. Therefore, we have* $1 \rightharpoonup 3$

[15] When verifying whether $2 \rightharpoonup 1$ holds, one is allowed to consider mixtures that have overlapping supports in dimension 1, which causes the property in Definition 2 to fail.

and $3 \rightharpoonup 1$, *although the decision maker does not evaluate dimension* 1 *conditioning on dimension* 3.

If $\rightharpoonup$ precisely identifies which dimension is evaluated conditioning on which other dimension, $3 \rightharpoonup 1$ should not hold, but it does because in the definition of $\rightharpoonup$, when we examine the relation between dimensions 1 and 3, we require that all of the other dimensions have no risk. Therefore, $\rightharpoonup$ may identify conditioning behavior that does not exist.

Our solution to this issue is as follows. In Example 4, we have $1 \rightharpoonup \{1,2,3\}$ but not $2 \rightharpoonup \{1,2,3\}$ or $3 \rightharpoonup \{1,2,3\}$. We conclude that the decision maker evaluates dimensions 2 and 3 conditioning on dimension 1. Next, remove dimension 1 from I and we have $2 \rightharpoonup \{2,3\}$ but not $3 \rightharpoonup \{2,3\}$. We conclude that given the realization of dimension 1, the decision maker evaluates dimension 3 conditioning on dimension 2. Comparing what this procedure tells us with the utility function, we find that we have identified the decision maker's conditioning behavior correctly.

3.1.2 Violations of Independence Due to Bracketing

The procedure at the end of the previous subsection relies on—given a component A under consideration—the existence of a dimension $i \in A$ such that $i \rightharpoonup A$. What if we cannot find such a dimension? In that case, the decision maker must have partitioned the dimensions in A into several parts, and evaluated them separately without conditioning between different parts. What happens to independence in this case? Return to the $N = 2$ case and consider the following example.

Example 5. *Suppose the decision maker evaluates a lottery by summing the certainty equivalents of marginal lotteries in both dimensions:*

$$u^{-1}(\mathbb{E}_1^p\, u(x_1)) + u^{-1}(\mathbb{E}_2^p\, u(x_2)),$$

in which $u(z) = \sqrt{z}$ *for* $z \geqslant 0$. *Consider* $x = (0,1)$ *and* $y = (1,0)$. *They both have utility* 1, *but the utility of* $\delta_x \frac{1}{2} \delta_y$ *is* $1/2$.

The above violation of independence differs from that in Section 3.1.1, because δ_x and δ_y do not have overlapping supports in either dimension. Rather, the violation is caused by the fact that the decision maker neglects correlation and evaluates each dimension's risk separately. This observation leads to the definition below that identifies which dimensions are evaluated separately from some other dimensions. When there is no risk, it is similar to the notion of separability in Debreu (1960) and Gorman (1968).

Definition 3. *We say that a proper subset $B \subsetneq A$ is isolated in A, denoted by $B \rhd A$, if for all $x \in X_{A^c}$, $r \in \Delta(X_{A \setminus B})$, and $p, q \in \Delta(X_A)$ such that $p_{A \setminus B} = q_{A \setminus B}$, we have $p \succsim_x q \iff (p_B, r) \succsim_x (q_B, r)$.*

Motivated by Example 5, Definition 3 captures two forms of separability if $B \rhd A$. First, the decision maker neglects the correlation between risk in B and risk in $A \setminus B$. To see this, let $r = p_{A \setminus B} = q_{A \setminus B}$ and $p_B = q_B$. Since $(p_B, r) = (q_B, r)$, we must have $p \sim_x q$, which implies that the correlation between risk in B and in $A \setminus B$ does not matter. Second, the decision maker's evaluation of risk in B is independent of what she faces in $A \setminus B$. To see this, for any $s, r \in \Delta(X_{A \setminus B})$, let $p = (p_B, s)$ and $q = (q_B, s)$. Since $B \rhd A$, we have $(p_B, s) \succsim_x (q_B, s) \iff (p_B, r) \succsim_x (q_B, r)$.

In Example 5, we can verify that $\{1\} \rhd \{1, 2\}$ and $\{2\} \rhd \{1, 2\}$, which is consistent with the hierarchy $\mathcal{H}^4 = \{\{1, 2\}, \{1\}, \{2\}\}$.

3.2 Characterization

We impose the following axioms on the decision maker's preference. We begin with two standard axioms.

Axiom 1. *(Weak Order) The preference $\succsim$ is complete and transitive.*

Axiom 2. *(Outcome Monotonicity) For all $x, y \in X$, if $x \geqslant y$ and $x \neq y$, then $x \succ y$.*

We have seen that independence may be violated due to either bracketing or conditioning of different dimensions (Section 3.1). However, when we focus on a single dimension, we require that independence holds.

Axiom 3. *(Unidimensional Independence) For all $i \in I$, $p, q, r \in \Delta(X_i)$, $x \in X_{-i}$, and $\alpha \in (0, 1)$, we have $p \succ_x q \implies p\alpha r \succ_x q\alpha r$.*

Our main axiom relaxes independence across dimensions by drawing a connection between conditioning and bracketing, which can be revealed from choice behavior using $\rightharpoonup$ (Definition 2) and $\triangleright$ (Definition 3). Its idea is simple: If no conditioning, then there must be bracketing. More specifically, given any set of dimensions A, if no dimension in A is conditioned on when the decision maker evaluates all the other dimensions in A, then the decision maker must have partitioned A into several subsets to evaluate separately. Consequently, every dimension in A must belong to some bracket that is isolated in A, and the union of those brackets should be A.

Axiom 4. *(Separability Under Bracketing) If $i \not\rightharpoonup A$ for all $i \in A$, then $\bigcup_{B \triangleright A} B = A$.*

Axiom 4 is satisfied in Example 3, since $1 \rightharpoonup \{1, 2\}$, and in Example 5, since $\{1\} \triangleright \{1, 2\}$ and $\{2\} \triangleright \{1, 2\}$. We also provide examples that will show how the above two axioms may be violated in Online Appendix I.2.

The last axiom is continuity. The standard continuity axiom requires that for every lottery p, the set of lotteries that are weakly better than p and the set of lotteries that are weakly worse than p are closed. This notion of continuity may be too demanding in our theory because of conditioning. Consider the following example.

Example 6. *Let $N = 2$. Again suppose the decision maker's utility of p is given by $\sum_{x_1} p_1(x_1)u(x_1, p_{2|x_1})$, in which $u(x_1, p_2) = x_1 + (\sum_{x_2} p_2(x_2)\sqrt{x_2})^2$. Consider a lottery that yields $(\varepsilon, 0)$ and $(0, 4)$ with equal probability. As ε converges to 0, its utility converges to 2, but the lottery converges in distribution to (δ_0, q_2) with utility $u(0, q_2) = 1$, in which $q_2 = \delta_0 \frac{1}{2} \delta_4$.*

Nonetheless, the following weaker notion of continuity is orthogonal to the observation behind Example 6 and should remain valid in our theory.

Axiom 5. *(Continuity) For all $p, q, r \in \Delta(X)$, the sets $\{\alpha \in [0, 1] : p\alpha q \succsim r\}$ and $\{\alpha \in [0, 1] : r \succsim p\alpha q\}$ are closed in $[0, 1]$, and the sets $\{x \in X : x \succsim p\}$ and $\{x \in X : p \succsim x\}$ are closed in X.*

The main representation theorem is below. The main idea behind its proof is essentially the elicitation process described in Section 3.1.

Theorem 1. *The preference satisfies weak order, outcome monotonicity, unidimensional independence, separability under bracketing, and continuity if and only it it has an HEU representation.*

4 Uniqueness

To what extent can we uniquely identify the hierarchy and the Bernoulli indices in an HEU representation? To answer this question, we first introduce a useful definition. We say that a hierarchy is *tight* if for every $A \in \mathcal{H}$ that is not I, we have $\tau(A) \neq \emptyset$. That is, given any component A that is not I, there must be some dimension in A that is evaluated currently. The hierarchies in Figures 1 and 2 are tight.

The first observation is that the decision maker's preference has an HEU representation if and only if it has an HEU representation in which the hierarchy is tight. This observation is not trivial. It is implied by Lemma 1 in the Appendix.

The second observation is that, fixing a tight hierarchy $\mathcal{H}$, the uniqueness of the corresponding $(u^A)_{A\in\mathcal{H}}$ is similar to that in expected utility theory: Roughly speaking, for all $x \in X_{\eta(A)}$ and $a \in \mathbb{R}^{\Phi(A)}$, the function $u^A(x, \cdot, a)$ is unique up to a positive affine transformation.[16] The arguments are standard. We leave the details to Online Appendix III.

Note that it is important to the second observation that the hierarchy is tight. Consider the example in Figure 4. Because $\tau(\{2, 3, 4\}) = \emptyset$, this hierarchy is not tight. It is easy to see that it is without loss of generality to remove the component $\{2, 3, 4\}$ from the hierarchy. The reason is simple: A function that takes the form $f(x, g(y, z))$ is more restrictive than a function that takes the form $h(x, y, z)$, and moreover, in the former case, g is redundant and cannot be jointly identified with f.

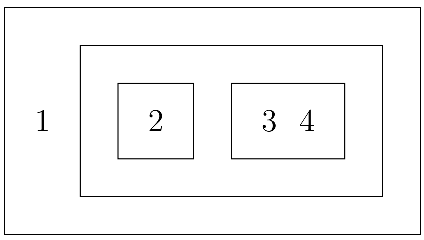

Figure 4: A hierarchy on $I = \{1, 2, 3, 4\}$ that is not tight.

[16] We have referred to all $(u^A)_{A\in\mathcal{H}}$ as Bernoulli indices by an abuse of terminology. A function u^A is a Bernoulli index if and only if $\tau(A) \neq \emptyset$. When $\tau(A) = \emptyset$, the expectation operator in equation (1) is degenerate and hence u^A is not a Bernoulli index. Given a tight hierarchy, it is possible when $A = I$ that $\tau(A) = \emptyset$. In that case, u^I is unique up to a monotone transformation.

More importantly, we want to analyze the uniqueness of the hierarchy. We say that the hierarchy is *unique* if for any HEU representations of $\succsim$, $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$ and $(\tilde{\mathcal{H}}, (v^A)_{A\in\mathcal{H}})$, we have $\mathcal{H} = \tilde{\mathcal{H}}$. In general, the hierarchy in the HEU representation is not unique, even if we focus on tight hierarchies. For instance, if $\succsim$ has an expected utility representation with an additively separable Bernoulli index, then we can verify that for any tight hierarchy $\mathcal{H}$, there exists an HEU representation of $\succsim$ with hierarchy $\mathcal{H}$.

However, it is possible to identify a unique *canonical* hierarchy. The idea is simple: The canonical hierarchy highlights bracketing and avoids redundant conditioning. To define it, we need some additional notations. Let $\mathbb{H}$ be the set of all hierarchies $\mathcal{H}$ such that $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$ is an HEU representation of $\succsim$ for some $(u^A)_{A\in\mathcal{H}}$. We add a superscript to functions $H^{\mathcal{H}}$, $\tau^{\mathcal{H}}$, $\eta^{\mathcal{H}}$, and $\Phi^{\mathcal{H}}$ defined in Section 2 to emphasize their dependence on $\mathcal{H}$.

Definition 4. *We say that a tight hierarchy $\mathcal{H}^* \in \mathbb{H}$ is a canonical hierarchy for $\succsim$ if for every $A \in \mathcal{H}^*$, the following statements hold:*

1. *If there exists an $\mathcal{H} \in \mathbb{H}$ such that $A \in \mathcal{H}$ and $\tau^{\mathcal{H}}(A) = \emptyset$, then $\tau^{\mathcal{H}^*}(A) = \emptyset$.*

2. *If $\tau^{\mathcal{H}}(A) \neq \emptyset$ for all $\mathcal{H} \in \mathbb{H}$ such that $A \in \mathcal{H}$, then $\tau^{\mathcal{H}}(A) \subseteq \tau^{\mathcal{H}^*}(A)$ for all $\mathcal{H} \in \mathbb{H}$ such that $A \in \mathcal{H}$.*

The motivation for the first statement can be understood using the previous example in which $\succsim$ has an expected utility representation with an additively separable Bernoulli index. Recall that in that case, any tight hierarchy $\mathcal{H}$ can be used to form an HEU representation of $\succsim$. However, all but one hierarchies fail to convey the most crucial feature of $\succsim$. When the Bernoulli index is additively separable, the decision maker's choice behavior exhibits full narrow bracketing and correlation neglect. The hierarchy $\{I, \{1\}, \ldots, \{N\}\}$ precisely highlights the fact that every dimension has its own bracket from the decision maker's point of view—whereas, for instance, the hierarchy $\{I\}$ treats this example no differently from a generic expected utility representation. Following our discussion in Section 3.1, we know that when $\tau^{\mathcal{H}}(A) = \emptyset$, the decision maker must have divided the dimensions in A into groups for separate evaluation. Therefore, the first statement says that whenever the evaluation of A is performed in groups and A is contained in the canonical hierarchy, the canonical hierarchy must capture the separate evaluation of A.

To understand the second statement, consider an example in which $N = 2$ and $\succsim$ has an expected utility representation whose Bernoulli index is not additively separable. It can be shown that $\succsim$ is represented by HEU representations with any of the following hierarchies: $\{\{1,2\}\}$, $\{\{1,2\},\{2\}\}$, or $\{\{1,2\},\{1\}\}$. However, the conditioning exhibited in the second and third hierarchies is redundant. The second statement says that the canonical hierarchy should avoid exhibiting such redundant conditioning.

Our second theorem establishes the existence and uniqueness of the canonical hierarchy. Its proof is constructive.

Theorem 2. *If $\succsim$ has an HEU representation, then there exists a unique canonical hierarchy for $\succsim$.*

5 Special Cases of the HEU Representation

Our framework allows us to characterize several well-known utility representations in a novel way. Moreover, it helps us derive useful generalizations of those representations. We first formally define the three extreme cases of the HEU representation discussed in the Introduction.

Definition 5. *The preference $\succsim$ has a FATE representation if there exists a continuous and strictly increasing function $u : X \to \mathbb{R}$ such that for all $p, q \in \Delta(X)$,*

$$p \succsim q \iff \mathbb{E}^p\, u(x) \geqslant \mathbb{E}^q\, u(x).$$

The preference $\succsim$ has a FETA representation if there exist continuous and strictly increasing functions $v : \mathbb{R}^I \to \mathbb{R}$ and $u^i : X_i \to \mathbb{R}$, $i = 1, \ldots, N$, such that for all $p, q \in \Delta(X)$,

$$p \succsim q \iff v(\,\mathbb{E}^p_1\, u^1(x_1), \ldots, \mathbb{E}^p_N\, u^N(x_N)\,) \geqslant v(\,\mathbb{E}^q_1\, u^1(x_1), \ldots, \mathbb{E}^q_N\, u^N(x_N)\,).$$

The preference $\succsim$ has a recursive representation if it has an HEU representation whose hierarchy $\mathcal{H}$ is unique and satisfies $|\mathcal{H}| = N$ and, for all

$A, B \in \mathcal{H}$, we have $A \subseteq B$ or $B \subseteq A$.

Clearly, the FATE representation is an HEU representation with hierarchy $\mathcal{H} = \{I\}$, and the FETA representation is an HEU representation with hierarchy $\mathcal{H} = \{I, \{1\}, \ldots, \{N\}\}$. As stated in Section 2.1, the former is an expected utility representation, and the latter captures full narrow bracketing. They correspond to the two commonly used opposite approaches to evaluate lotteries.

Given a recursive representation with hierarchy $\mathcal{H}$, for any component $A \in \mathcal{H}$, there is only one present dimension in $\tau(A)$ and one future component in $\Phi(A)$. Hence, the recursive representation is analogous to those of Kreps and Porteus (1978) and Epstein and Zin (1989). The order of the dimensions, however, is subjective. It is as if dimension $\tau(I)$ predates the remaining dimensions $I_1 = I \setminus \tau(I)$—which must be a component itself—and dimension $\tau(I_1)$ predates the remaining dimensions $I_2 = I_1 \setminus \tau(I_1)$—which again must be a component itself—and so on.

The next theorem provides a characterization of these three representations. The FATE representation obviously can be characterized using independence from expected utility theory, but we will provide an alternative characterization, making use of our framework. Our new characterization of the FATE representation will help us derive a natural generalization of it. For any $i, j \in I$, denote $i \rightharpoonup\!\!\!\!\rightharpoonup j$ if $i \rightharpoonup j$ and $j \not\rightharpoonup i$.

Theorem 3. *Suppose the preference $\succsim$ has an HEU representation. It has a FATE representation if and only if $i \rightharpoonup I$ for every $i \in I$. It has a FETA representation if and only if $\{i\} \rhd I$ for every $i \in I$. It has a recursive representation if and only if there exists a bijective function $\pi : I \to I$ such that $\pi(i) \rightharpoonup\!\!\!\!\rightharpoonup \pi(i+1)$, for all $i = 1, \ldots, N-1$.*

The characterization of the FATE representations is intuitive. One might wonder whether the following condition characterizes the FETA representation: For all $i, j \in I$, we have $i \rightharpoonup j \iff i = j$. The issue with this idea can be seen by noticing that the expected utility function with an additively separable Bernoulli index is a special case of the FETA representation, and in that case, $i \rightharpoonup j$ for all $i, j \in I$.

For the recursive representation, the permutation function π and the relation $\rightharpoonup$ indicate the subjective order, following which the decision maker conducts the recursive evaluation of a lottery. We provide additional results in Online Appendix III.2 for the case in which the permutation function is the identity function. This case will be relevant when, for example, the dimensions represent (ordered) time periods.

One might conjecture that to characterize the recursive representation, we simply need $\rightharpoonup$ to induce a linear order on I. This is incorrect. As discussed in Section 3.1.1 and Example 4, the relation $\rightharpoonup$ does not precisely identify how the decision maker orders different dimensions. However, it can be shown that if we observe $i \rightharpoonup\!\!\!\!\!\!- j$, then it must be true that in any HEU representation of the decision maker's preference, it is as if dimension i predates dimension j strictly. This observation is the key to our characterization above.

We relax the conditions that characterize the FATE, FETA, and recursive representations to derive useful generalizations of them. Consider the following two representations, whose examples are illustrated in Figure 2.

Definition 6. *The preference has a generalized bracketing representation if there exist a partition $\{A_i\}_{i=1}^n$ of I and continuous and strictly increasing functions $v : \mathbb{R}^{\{A_i\}_{i=1}^n} \to \mathbb{R}$ and $u^{A_i} : X_{A_i} \to \mathbb{R}$, $i = 1, \dots, n$, such that for all $p, q \in \Delta(X)$, we have $p \succsim q$ if and only if*

$$v(\mathbb{E}^p_{A_1} u^{A_1}(x_{A_1}), \dots, \mathbb{E}^p_{A_n} u^{A_n}(x_{A_n})) \geqslant v(\mathbb{E}^q_{A_1} u^{A_1}(x_{A_1}), \dots, \mathbb{E}^q_{A_n} u^{A_n}(x_{A_n})).$$

The preference has a generalized recursive representation if it has an HEU representation with hierarchy $\mathcal{H}$ such that for all $A, B \in \mathcal{H}$, we have $A \subseteq B$ or $B \subseteq A$.

The FATE representation can be viewed as a generalized bracketing representation with the trivial partition: The decision maker puts all dimensions into one bracket and evaluates the uncertainty over that bracket jointly. The FETA representation can be viewed as a generalized bracketing representation with the finest partition, in which every partition element is a singleton. Every dimension has its own bracket, and the decision maker evaluates the uncertainty over each bracket separately before aggregating across brackets. The generalized bracketing representation allows for more general brackets, such as the top-left example in Figure 2.

For a recursive representation, there is only one present dimension in any component. By contrast, for a generalized recursive representation, there may be multiple present dimensions at any component. This generalization certainly enables us to capture richer behavioral patterns of the recursive evaluation of risk in static settings, but it also introduces new ways to model the recursive evaluation of risk in dynamic choice. For example, we usually divide the timeline into multiple periods evenly. Under the recursive representation, how the decision maker brackets and orders the time periods to evaluate risk recursively coincides with the exogenous arrangement of time periods: a linear order on the dimensions. Under the generalized recursive representation, the decision maker may evaluate risk recursively following her own subjective arrangement of time periods: a (subjective) weak order on the dimensions. For example, a student may think of a whole semester as one period but treat each day of a vacation as one period. In other words, the uncertainty over the semester will be evaluated jointly, conditioning on which the daily uncertainty during the vacation will be evaluated recursively.

The next result characterizes the generalized bracketing representation and the generalized recursive representation.

Theorem 4. *Suppose the preference $\succsim$ has an HEU representation. It has a generalized bracketing representation if and only if for every nonempty $A \subseteq I$, if $\bigcup_{B \rhd A} B \neq A$, then $i \rightharpoonup A$ for every $i \in A$. It has a generalized recursive representation if and only if for every nonempty $A \subseteq I$, there exists $i \in A$ such that $i \rightharpoonup A$.*

One might conjecture that the characterization of a generalized recursive representation is that $\rightharpoonup$ induces a weak order on I. Again, this is incorrect, because of the reason discussed in Section 3.1.1 and Example 4. We provide additional results in Online Appendix III.2 for the case in which the dimensions are exogenously ordered, as in the setting of dynamic choice.

5.1 Examples: Group Inequality

At the beginning of the Introduction, we discuss the conflict between the FATE and the FETA approaches; more examples of such conflict are discussed in Online Appendix I.1. The generalized bracketing representation offers a new solution to such conflict. We explain it in the context of inequality.

Over the last few decades, economists have documented significant changes in inequality within and between social groups, and proposed measures that can capture both within-group and between-group inequality to study them.[17] As discussed in the Introduction, the FATE representation captures ex post inequality aversion, while the FETA representation captures ex ante inequality aversion, if the relevant functions in those representations are concave. The former's hierarchy is $\{I\}$ and the latter's is $\mathcal{H} = \{I, \{1\}, \dots, \{N\}\}$.

Suppose $I = \{1, 2, 3, 4\}$, and interpret the i^{th} component of $x \in X$ as individual i's income. Let $A = \{1, 4\}$ and $B = \{2, 3\}$ represent two social groups. Suppose the decision maker's preference has generalized bracketing representation such that the hierarchy is $\{I, A, B\}$ and functions u^I, u^A, and u^B are concave. Suppose $\delta_{(0,1,1,0)} \sim \delta_{(1,0,0,1)}$. It can be seen that the decision maker is averse to ex post inequality for individuals in the same social group (components A and B):

$$\delta_{(1/2,1/2)} \succ_{(x_2,x_3)} \frac{1}{2}\delta_{(0,1)} + \frac{1}{2}\delta_{(1,0)} \quad \text{and} \quad \delta_{(1/2,1/2)} \succ_{(x_1,x_4)} \frac{1}{2}\delta_{(0,1)} + \frac{1}{2}\delta_{(1,0)}.$$

Moreover, the decision maker is averse to ex ante inequality between social groups:

$$\frac{1}{2}\delta_{(0,1,1,0)} + \frac{1}{2}\delta_{(1,0,0,1)} \succ \delta_{(0,1,1,0)} \sim \delta_{(1,0,0,1)}.$$

Therefore, the generalized bracketing representation becomes a natural intermediate case between the FATE and FETA representations, and can capture the following notion of group inequality: The decision maker cares about ex post inequality (inequality of outcome) for individuals within the same group, and cares about ex ante inequality (inequality of opportunity) across groups.

Note that we may also take the generalized bracketing representation to a setting with ambiguity. In that case, it will become a natural intermediate case between the ex ante approach to evaluate randomization under

[17] See, among others, Burstein et al. (2019), Darity Jr (2022), Elbers et al. (2008), La Ferrara (2002), Formby et al. (1989), Gottschalk (1997), and Lemieux (2006).

ambiguity and the ex post approach. See more discussion in Section 1.1.

6 Applications

6.1 Multisource Income

In this section, we consider a decision maker who receives income from multiple sources and show how narrow bracketing can induce (i) stochastically dominated choices and (ii) avoidance of multidimensional risk.

Suppose for some $b > 0$, $X_i = Z = [-b, b]$ for every $i \in I$. Denote $Z_+ = [0, b]$ and $Z_- = [-b, 0]$. Each element in Z is a monetary prize and can indicate either a gain or a loss, depending on its sign. For any $A \subseteq I$, we interpret $p \in \Delta(X_A)$ as a joint distribution of incomes from sources in A, which induces a distribution of total income denoted by $f[p]$. That is, the probability of total income $z \in \mathbb{R}$ is $f[p](z) = \sum_{x \in X_A} p(x)\mathbb{1}\{\sum_{i\in A} x_i = z\}$, in which $\mathbb{1}$ is the indicator function. To keep notation simple, we only work with p's such that $f[p] \in \Delta(Z)$ throughout this subsection.

Suppose $\succsim$ has the following generalized bracketing representation:

$$U(p) = \sum_{i=1}^{n} c(f[p_{A_i}], u), \tag{2}$$

in which $\{A_i\}_{i=1}^n$ is a partition of I, $u : Z \to \mathbb{R}$ is a continuous and strictly increasing function, and $c(q, u) = u^{-1}(\mathbb{E}^q\, u(x))$ is the certainty equivalent of $q \in \Delta(Z)$ under the Bernoulli index u.[18] Under this representation, the decision maker classifies income sources into several brackets and evaluates a lottery by summing the certainty equivalents of the distribution of total income in each bracket. As a result, her behavior exhibits narrow bracketing across different brackets of income sources. We further assume that u is twice continuously differentiable, and hence the Arrow–Pratt measure of absolute risk aversion $A(x) = -u''(x)/u'(x)$ is well defined.[19]

[18] In Online Appendix II.1, we characterize the behavioral implications of (2). We also discuss an alternative formulation and explain why we choose (2).

[19] Our results remain valid without this assumption, once we replace conditions on A with alternative definitions of increasing/constant/decreasing absolute risk aversion.

6.1.1 Violations of Stochastic Dominance

First, we study the connection between narrow bracketing and dominated choices. For any $p, q \in \Delta(Z)$, we say that p (first-order) stochastically dominates q, denoted by $p \succ_{FOSD} q$, if $p \neq q$ and $\sum_{x \leqslant z} q(x) \geqslant \sum_{x \leqslant z} p(x)$ for all $z \in Z$. We say that $\succsim$ satisfies *dominance* if $f[p] \succ_{FOSD} f[q]$ implies $p \succ q$ for all $p, q \in \Delta(Z^N)$ (such that $f[p], f[q] \in \Delta(Z)$). This property, although commonly assumed in economic models, faces challenges from experimental evidence.[20] Consider the following example.

Example 7. *Consider the following pair of decisions. Each decision's risk will be resolved independent of the other's, and both choices will impact your overall payment. Examine both decisions and indicate your preferred choices.*

Decision 1: Choose between

A. A sure gain of \$2.40.

B. A 25% chance to gain \$10.00, and a 75% chance to gain \$0.

Decision 2: Choose between

C. A sure loss of \$7.50.

D. A 75% chance to lose \$10.00, and a 25% chance to lose \$0.

Rabin and Weizsäcker (2009) and Tversky and Kahneman (1981) show a significant proportion of subjects—at least 28%—choose A in decision 1 and D in decision 2. However, the resulting distribution of total income is stochastically dominated by that obtained by the combination of B and C:

$$\frac{3}{4}\delta_{-7.50} + \frac{1}{4}\delta_{2.50} \succ_{FOSD} \frac{3}{4}\delta_{-7.60} + \frac{1}{4}\delta_{2.40},$$

in which the left-hand mixture results from choosing B and C and the right-hand from choosing A and D. This violation of dominance is notable because the combination of B and C is equal to the combination of A and D plus a payoff of \$0.10 with certainty. Such violations are inconsistent with models that consider only the distribution of total income, including those that allow violations of dominance.[21] By contrast, when the subject

[20] In the context of dynamic choice, Bommier et al. (2017) also show that many commonly used recursive models fail to satisfy a version of dominance.

[21] See, among others, Bell (1985), Kőszegi and Rabin (2007), Loomes and Sugden (1986), Mononen (2022), and Puri (2022).

evaluates the two choice problems separately, her choice of A in decision 1 can be rationalized by risk aversion over gains and her choice of D in decision 2 can be rationalized by risk seekingness over losses, both of which are commonly observed patterns.[22] The next result illustrates the tension between bracketing and dominance.

Proposition 1. *Suppose the preference $\succsim$ is represented by (2). It satisfies dominance if and only if it is represented by $U(p) = \mathbb{E}^{f[p]}u(x)$.*

Proposition 1 states that the decision maker's behavior satisfies dominance for all lotteries if and only if she is a broad bracketer. That is, she puts all income sources into one bracket. By comparison, if we only focus on lotteries with independent marginals like those in Example 7, dominance can be maintained under narrow bracketing if u exhibits constant absolute risk aversion (CARA)—i.e., if $A(x)$ is a constant function.[23] We say that $\succsim$ satisfies *dominance without correlation* if $f[p] \succ_{FOSD} f[q]$ implies $p \succ q$ for all $p, q \in \Delta(Z^N)$ such that $p = (p_1, \ldots, p_N)$ and $q = (q_1, \ldots, q_N)$.

Proposition 2. *Suppose the preference $\succsim$ is represented by (2). It satisfies dominance without correlation if and only if it is represented by $U(p) = \mathbb{E}^{f[p]}u(x)$ or u exhibits CARA.*

6.1.2 Avoidance of Multidimensional Risk

Another implication of narrow bracketing is that the decision maker may possess a strict preference regarding whether her income comes from a single source or multiple ones. For instance, due to the difficulty of integration across sources, she may be predisposed to avoid issues that feature multidimensional risk (Heo, 2021). To capture such behavior, we define p as a *single-source* lottery if there exist $i \in I$ and $r \in \Delta(Z)$ such that $p_i = r$ and $p_j = \delta_0$ for all $j \neq i$. Since the utility of p is $U(p) = \mathbb{E}^r u(x)$ according to (2), we can use elements of $\Delta(Z)$ to denote single-source lotteries when there is no risk of confusion. We say that $\succsim$ satisfies *avoidance of multidimensional*

[22] As a concrete example, consider the utility function in (2), in which $A_1 = \{1\}, A_2 = \{2\}$, $u(x) = \sqrt{x}$ for $x \geqslant 0$, and $u(x) = -2\sqrt{-x}$ for $x < 0$.

[23] Mu et al. (2023b) and Rabin and Weizsäcker (2009) have made similar observations.

risk if $f[p] \succsim p$ for every $p \in \Delta(Z^N)$ such that $p = (p_1, \dots, p_N)$.[24] Similarly, $\succsim$ satisfies avoidance of multidimensional risk for gains/losses if $f[p] \succsim p$ for every $p \in (\Delta(Z_+^N)$/for every $p \in (\Delta(Z_-^N)$ such that $p = (p_1, \dots, p_N)$.

Proposition 3. *Suppose the preference $\succsim$ is represented by (2).*

1. *It satisfies avoidance of multidimensional risk if and only if $\succsim$ is represented by $U(p) = \mathbb{E}^{f[p]}u(x)$ or u exhibits CARA.*

2. *It satisfies avoidance of multidimensional risk for gains if and only if $\succsim$ is represented by $U(p) = \mathbb{E}^{f[p]}u(x)$ or $A(x)$ is decreasing in $x \in Z_+$.*

3. *It satisfies avoidance of multidimensional risk for losses if and only if $\succsim$ is represented by $U(p) = \mathbb{E}^{f[p]}u(x)$ or $A(x)$ is increasing in $x \in Z_-$.*

6.2 Time Lotteries

In this section, we study decisions involving risk about both which and when prizes will be delivered. Suppose $N = 2, X_1 = Z = [w, b] \subset \mathbb{R}_{++}$ and $X_2 = T = [0, \bar{t}\,] \subset \mathbb{R}_+$. Each lottery in $\Delta(Z \times T)$ denotes a distribution of dated prizes. In particular, a *time lottery* $(z, p) \in Z \times \Delta(T)$ is a lottery in which the prize z is fixed and the payment date follows the distribution p.

Suppose the preference $\succsim$ admits an HEU representation and it is a monotone transformation of the exponentially discounted utility function in the absence of risk:

$$U^I(\delta_{(z,t)}) = \phi\Big(v(z)e^{-rt}\Big), \tag{3}$$

in which $r > 0$ and both $v : Z \to \mathbb{R}_{++}$ and $\phi : [e^{-r\bar{t}}u(w), u(b)] \to \mathbb{R}$ are strictly increasing and continuous.[25] Note that we have not yet fully specified the HEU representation, and equation (3) differs slightly from

[24]If we adopt a more stringent notion of avoidance of multidimensional risk by requiring $f[p] \succsim p$ for every $p \in \Delta(Z^N)$, then the decision maker must be a broad bracketer as in Proposition 1.

[25]Fishburn and Rubinstein (1982) show that exponentially discounted utility can be characterized by the stationarity axiom: For any $z, z' \in Z$, $s, t \in T$, and $\tau \in \mathbb{R}$ with $s + \tau, t + \tau \in T$, if $(z, t) \sim (z', t + \tau)$, then $(z, s) \sim (z', s + \tau)$.

Definition 1 in that it is strictly decreasing in t in order to capture impatience of the decision maker.[26]

We say that a decision maker is *risk-averse over time lotteries* if she prefers receiving a prize on a sure date rather than on a random date with the same mean; that is, $(z, \mathbb{E}^p(t)) \succsim (z, p)$ for all $(z, p) \in Z \times \Delta(T)$. Analogously, she is *risk-seeking over time lotteries* if the opposite holds.

For now, assume ϕ is affine. If $\mathcal{H} = \{\{1, 2\}\}$, the HEU representation reduces to the widely used exponentially discounted expected utility model $V(p) = \mathbb{E}^p[v(z)e^{-rt}]$. Because e^{-rt} is convex in t, the decision maker must be risk-seeking over time lotteries. However, DeJarnette et al. (2020) find that the majority of their subjects are risk-averse over time lotteries in most questions in their experiment. To see how other HEU representations can resolve such inconsistency, suppose $\mathcal{H} = \{\{1, 2\}, \{1\}, \{2\}\}$. The utility of a time lottery (z, p) can be written as

$$U(z, p) = v(z)e^{-rc(p,u^2)}.$$

Because $(z, \mathbb{E}_p(t)) \succsim (z, p)$ if and only if $u^2(\mathbb{E}_p(t)) \leqslant \mathbb{E}^p u^2(t)$, risk aversion and risk seekingness over time lotteries are equivalent to the convexity and concavity of u^2, respectively. The same logic also applies for hierarchy $\mathcal{H} = \{\{1, 2\}, \{2\}\}$. The following result summarizes these observations.

Proposition 4. *Suppose $\succsim$ has an HEU representation $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$ and satisfies (3) in which ϕ is affine. It is risk-averse over time lotteries if and only if either (i) $\mathcal{H} = \{\{1, 2\}, \{2\}\}$ and $u^2(z, \cdot)$ is convex for every $z \in Z$, or (ii) $\mathcal{H} = \{\{1, 2\}, \{1\}, \{2\}\}$ and u^2 is convex.*

If we allow ϕ to be nonaffine, $\mathcal{H} = \{\{1, 2\}\}$ can also accommodate risk aversion over time lotteries. However, it entails violations of a risky counterpart of impatience (DeJarnette et al., 2020). We say that $\succsim$ satisfies *(nontrivial) stochastic impatience* if for all $t_1, t_2 \in T$, and $z_1, z_2 \in Z$ with $t_1 < t_2$ and $z_1 > z_2$, we have $\delta_{(z_1,t_1)}\frac{1}{2}\delta_{(z_2,t_2)} \succsim \delta_{(z_2,t_1)}\frac{1}{2}\delta_{(z_1,t_2)}$, and this ranking is not always indifference. Intuitively, it states that if the decision maker can pair monetary prizes with payment dates in the presence of risk, she will prefer to receive the highest prize at the earliest time. DeJarnette et al.

[26] It is easy to modify Definition 1 to accommodate this: We assume that u^I is strictly decreasing in the second argument and all other Bernoulli indices are strictly increasing.

(2020) show that the incompatibility between stochastic impatience and any violation of risk-seeking behavior over time lotteries persists in both expected utility theory with general discount functions and a broad class of non-expected utility models. The next result states that the HEU representation can address such incompatibility if and only if $\mathcal{H} = \{\{1,2\},\{2\}\}$. That is, the decision maker acts as if she evaluates risk in money and the conditional risk in time given each realization of money.

Proposition 5. *Suppose $\succsim$ has an HEU representation $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$ and satisfies (3). It satisfies stochastic impatience and is risk-averse over time lotteries if and only if $\mathcal{H} = \{\{1,2\},\{2\}\}$, $u^2(z,\cdot)$ is convex for all $z \in Z$, and ϕ is a nontrivial convex transformation of* ln.[27]

It is worth noting that the HEU representation with $\mathcal{H} = \{\{1,2\},\{2\}\}$ can also accommodate risk attitudes over the time dimension that are non-uniform or depend on the monetary prize (DeJarnette et al., 2020; Mu et al., 2023b). We conclude with a parametric example of Proposition 5.

Example 8. *Suppose the preference $\succsim$ is represented by*

$$U(p) = \mathbb{E}_1^p\Big[\frac{v(z)}{\mathbb{E}_{2|z}^p[e^{rt}]}\Big]. \tag{4}$$

This is an HEU representation with $\mathcal{H} = \{\{1,2\},\{2\}\}$, $u^2(z,t) = e^{rt}$, and $\phi(a) = a$. Because $u^2(z,\cdot)$ is convex and ϕ is a strictly convex transformation of ln, *Proposition 5 ensures that (4) is risk-averse over time lotteries and hence can serve as an alternative to the exponentially discounted expected utility model without compromising stochastic impatience or introducing additional free parameters.*

On the domain of time lotteries $Z \times \Delta(T)$, (4) can be rewritten as $U(z,p) = v(z)e^{-r\psi_r(p)}$, in which $\psi_r(p) = \frac{1}{r}\ln \mathbb{E}_p[e^{rt}]$. Note that ψ_r is a monotone additive statistic in Mu et al. (2023b).[28] Since stochastic impatience is not well defined on $Z \times \Delta(T)$, the representation (4) can be

[27] A function f defined on $B \subseteq \mathbb{R}_{++}$ is a nontrivial convex transformation of ln if there exists a convex and nonaffine function g such that $f(x) = g(\ln(x))$ for all $x \in B$.

[28] Mu et al. (2023b) characterize monotone additive statistics as weighted averages over ψ_r across different r. In our HEU model, ψ_r cannot be replaced with a general monotone additive statistic because of the axiom of unidimensional independence.

interpreted as a generalization of a monotone stationary time preference in Mu et al. (2023b) to the set $\Delta(Z \times T)$, which allows exploration of the interaction between stochastic impatience and risk attitudes toward time.

7 Conclusion

This paper proposes and axiomatizes a flexible framework to study preferences over risky multidimensional alternatives. The interaction between risk evaluation and how the decision maker brackets and orders dimensions in the evaluation process is encapsulated by a hierarchy, which can be revealed from the decision maker's choice behavior. We discuss in what sense the HEU representation is unique, characterize several special cases of it, and study applications of multisource income and time lotteries.

In Online Appendix II, we discuss several extensions of our applications. For multisource income, we consider an alternative model of narrow bracketing (Camara, 2021; Vorjohann, 2023) that involves adding expected utilities instead of certainty equivalents, as in (2). We also propose a notion of comparative avoidance of multidimensional risk and study a generalization with background risk (Freeman, 2015, 2017; Mu et al., 2023a). For time lotteries, we discuss how the recursive representation in Proposition 5 can be extended to a setting in which the decision maker may receive multiple prizes over time. In Online Appendix I.2, we discuss two examples that also generalize the FATE, FETA, and recursive approaches, but in ways that differ from the HEU representation. We discuss which of our axioms are not satisfied by these examples.

A Proofs

A.1 Proof of Theorem 1

Checking the necessity of the axioms is routine (yet nontrivial in our case). Below, we show that assuming tightness of hierarchies is without loss of generality and leave the rest of the proof of the necessity of the axioms to Online Appendix IV.

Lemma 1. *The preference $\succsim$ has an HEU representation if and only if it has an HEU representation in which the hierarchy is tight.*

Proof of Lemma 1. It suffices to prove the "only if" part. Suppose $\succsim$ has an HEU representation $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$. Define $\mathcal{H}' = \{A \in \mathcal{H} : \tau^{\mathcal{H}}(A) \neq \emptyset\} \cup \{I\}$. It is easy to check that $\mathcal{H}'$ is a tight hierarchy and $H^{\mathcal{H}}(A) = H^{\mathcal{H}'}(A)$, $\tau^{\mathcal{H}}(A) = \tau^{\mathcal{H}'}(A)$, and $\eta^{\mathcal{H}}(A) = \eta^{\mathcal{H}'}(A)$ for every $A \in \mathcal{H}'$. Hence, we can omit the superscripts for these three functions. For each $A \in \mathcal{H}'$, we define $\hat{u}^A : X_{\eta(A)} \times X_{\tau(A)} \times \mathbb{R}^{\Phi^{\mathcal{H}'}(A)} \to \mathbb{R}$ as follows: For any $x \in X_{\eta(A)}, y \in X_{\tau(A)}$, and $a \in \bigtimes_{B\in\Phi^{\mathcal{H}'}(A)} U^B_{x,y}(X_B)$ ($U^B_{x,y}$'s come from the HEU representation $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$), we can find $z \in \bigtimes_{B\in\Phi^{\mathcal{H}'}(A)} X_B$ such that $a_B = U^B_{x,y}(\delta_{z_B})$ for every $B \in \Phi^{\mathcal{H}'}(A)$. Let $\hat{u}^A(x, y, a) = U^A_x(\delta_y, \delta_z)$ and define $\hat{U}^A_x$ as in Definition 1. Clearly, $\hat{U}^A_x = U^A_x$. Extending $\hat{u}^A$ to its full domain trivially, we obtain $(\mathcal{H}', (\hat{u}^A)_{A\in\mathcal{H}})$ as an HEU representation of $\succsim$ in which $\mathcal{H}'$ is tight. □

Next, we focus on the proof of the sufficiency of the axioms. For every $A \subseteq I$, denote $M(A) = \{i \in A : i \rightharpoonup A\}$.

Step 1: Preliminary results.

We present several lemmas that will be useful in later steps. Most proofs of these lemmas are deferred to Online Appendix IV.2. First, the conditional preference on each dimension has an expected utility (EU) representation.

Lemma 2. *For any $i \in I$ and $x \in X_{-i}$, the conditional preference $\succsim_x$ on $\Delta(X_i)$ has an EU representation with a continuous and strictly increasing Bernoulli index $v_{i|x}$, which is unique up to a positive affine transformation.*

The second lemma strengthens Axiom 2 (outcome monotonicity). For any $A \subseteq I$ and $p \in \Delta(X_A)$, we denote by supp(p) the support of p—i.e., supp$(p) = \{x \in X_A : p(x) > 0\}$. For any $p \in \Delta(X_A)$ and $x \in X_A$, we say that p *dominates* x if $p \neq \delta_x$ and $y_i \geqslant x_i$ for all $i \in A$ and $y_i \in$ supp(p_i). Similarly, x *dominates* p if $p \neq \delta_x$ and $x_i \geqslant y_i$ for all $i \in A$ and $y_i \in$ supp(p_i). The dominance relation is weak if we allow for the possibility that $\delta_x = p$.

Lemma 3. (i) *For any $A \subseteq I$, $p \in \Delta(X_A)$, and $x' \in X_{A^c}$, if p dominates x, then $p \succ_{x'} x$, and if x dominates p, then $x \succ_{x'} p$.* (ii) *For any $A \subseteq I$,*

$p \in \Delta(X_A)$, $x' \in X_{A^c}$, and $x, y \in X_A$ such that p dominates y and is dominated by x, there exists some $z \in X_A$ such that $p \sim_{x'} z$ and $x \geqslant z \geqslant y$.

An immediate corollary of Lemma 3 is that for any $p \neq \overline{x} = (\overline{x}_i)_{i\in A}, \underline{x} = (\underline{x}_i)_{i\in A}$, we have $\overline{x} \succ p \succ \underline{x}$ and the set $\{y_i : y \sim p\}$ is uncountable for each i. The next result studies properties of the binary relation $\rhd$.

Lemma 4. (i) *If $B \rhd A$, $B' \rhd A$, and $B \cap B', B \setminus B', B' \setminus B$ are nonempty, then $B \cap B' \rhd A, B \setminus B' \rhd A$, and $B' \setminus B \rhd A$.* (ii) *If $A = \bigcup_{B \rhd A} B$, then there exists a nontrivial partition $\{B_k\}_{k=1}^n$ of A such that $B_k \rhd A$ for every $k = 1, \ldots, n$.*

We say that A is *bracket separable* if there exists a nontrivial partition $\{B_k\}_{k=1}^n$ of A such that $B_k \rhd A$ for all $k = 1, \ldots, n$. In this case, we call $\{B_k\}_{k=1}^n$ a bracket partition of A. Lemma 4 establishes the equivalence between bracket separability and the condition $A = \bigcup_{B \rhd A} B$. The next result guarantees that a bracket separable set has a "finest" bracket partition.

Lemma 5. *If A is bracket separable, then A must have a bracket partition $\{A_k\}_{k=1}^n$ in which A_k is not bracket separable for all $k = 1, \ldots, n$. Moreover, such $\{A_k\}_{k=1}^n$ is unique and is finer than any other bracket partition of A.*[29]

Proof of Lemma 5. Suppose that A has a bracket partition $\{A_k\}_{k=1}^n$ in which A_1 is bracket separable with bracket partition $\{B_k\}_{k=1}^m$. We claim that $\{A_k\}_{k=2}^n \cup \{B_k\}_{k=1}^m$ is a bracket partition of A. To see this, note that for all $p \in \Delta(X_A)$ and $x \in X_{A^c}$, we have $p \sim_x (p_{A_1}, \ldots, p_{A_n})$.

According to the definition of bracket partition and Lemma 3, there exists $x_{A_k} \in X_{A_k}$ for every $k \geqslant 2$ such that $(p_{A_1}, \ldots, p_{A_n}) \sim_x (p_{A_1}, x_{A_2}, \ldots, x_{A_n})$. Then, given $(x, (x_{A_k})_{k=2}^n) \in X_{A_1^c}$, since $\{B_k\}_{k=1}^m$ is a bracket partition of A_1,

$$\begin{aligned} p &\sim_x (p_{A_1}, x_{A_2}, \ldots, x_{A_n}) \\ &\sim_x (p_{B_1}, \ldots, p_{B_m}, x_{A_2}, \ldots, x_{A_n}) \\ &\sim_x (p_{B_1}, \ldots, p_{B_m}, p_{A_2}, \ldots, p_{A_n}). \end{aligned}$$

The last indifference relation holds, since $\{A_k\}_{k=1}^n$ is a bracket partition.

[29]For two partitions $\{A_k\}_{k=1}^n$ and $\{B_l\}_{l=1}^m$ of the same set A, we say that $\{A_k\}_{k=1}^n$ is finer than $\{B_l\}_{l=1}^m$ if for every A_k there exists some B_l such that $A_k \subseteq B_l$.

Fix any $k \in \{1, \ldots, m\}$. We want to show that $B_k \rhd A$. For any $x \in X_{A^c}$, $r \in \Delta(X_{A \setminus B_k})$, and $p, q \in \Delta(X_A)$ such that $p_{A \setminus B_k} = q_{A \setminus B_k}$, we have

$$\begin{aligned}
& p \succsim_x q \\
\Longleftrightarrow\ & (p_{B_k}, p_{A_1 \setminus B_k}, p_{A \setminus A_1}) \succsim_x (q_{B_1}, p_{A_1 \setminus B_k}, p_{A \setminus A_1}) \\
\overset{A_1 \rhd A}{\Longleftrightarrow}\ & (p_{B_k}, p_{A_1 \setminus B_k}, x_{A \setminus A_1}) \succsim_x (q_{B_k}, p_{A_1 \setminus B_k}, x_{A \setminus A_1}) \\
\overset{\text{by definition}}{\Longleftrightarrow}\ & (p_{B_k}, p_{A_1 \setminus B_k}) \succsim_{(x, x_{A \setminus A_1})} (q_{B_k}, p_{A_1 \setminus B_k}) \\
\overset{B_k \rhd A_1}{\Longleftrightarrow}\ & (p_{B_k}, r_{A_1 \setminus B_k}) \succsim_{(x, x_{A \setminus A_1})} (q_{B_k}, r_{A_1 \setminus B_k}) \\
\overset{\text{by definition}}{\Longleftrightarrow}\ & (p_{B_k}, r_{A_1 \setminus B_k}, x_{A \setminus A_1}) \succsim_x (q_{B_k}, r_{A_1 \setminus B_k}, x_{A \setminus A_1}) \\
\overset{A_1 \rhd A}{\Longleftrightarrow}\ & (p_{B_k}, r_{A_1 \setminus B_k}, r_{A \setminus A_1}) \succsim_x (q_{B_k}, r_{A_1 \setminus B_k}, r_{A \setminus A_1}) \\
\overset{A_1 \rhd A}{\Longleftrightarrow}\ & (p_{B_k}, r) \succsim_x (q_{B_k}, r).
\end{aligned}$$

Hence, $B_k \rhd A$ for every k and $\{A_k\}_{k=2}^n \cup \{B_k\}_{k=1}^m$ is a bracket partition of A. Continue this process until all elements in the bracket partition of A_1 are not bracket separable, and then repeat this procedure for other A_k's. After finitely many steps, we will end up with a bracket partition of A in which none of its elements is bracket separable, since singleton sets are not bracket separable and A is finite. For simplicity, we still denote it by $\{A_k\}_{k=1}^n$.

Consider a different bracket partition $\{B_l\}_{l=1}^m$. If $\{A_k\}_{k=1}^n$ is not finer than $\{B_l\}_{l=1}^m$, then there exist A_k and $B_{l_1}, \ldots, B_{l_t}$ with $t \geqslant 2$ such that $A_k \cap B_{l_i} \neq \emptyset$ for all $i = 1, \ldots, t$ and $A_k \subseteq \bigcup_{i=1}^t B_{l_i}$. By Lemma 4, $(A_k \cap B_{l_i}) \rhd A$ for all $i = 1, \ldots, t$, which implies that A_k has a bracket partition $\{A_k \cap B_{l_i}\}_{i=1}^t$, a contradiction. Hence, $\{A_k\}_{k=1}^n$ is finer than any other bracket partition $\{B_l\}_{l=1}^m$. Moreover, there must exist some bracket separable B_1 with a bracket partition being a subset of $\{A_k\}_{k=1}^n$. This ensures the uniqueness of the bracket partition in which all elements are not bracket separable. □

We call $\{A_k\}_{k=1}^n$ in Lemma 5 the *finest bracket partition* of A. The final lemma states that if $i \rightharpoonup A$, then we can find a representation of the conditional preference on $\{i\} \cup A$ with certain linearity properties.

Lemma 6. *Let* $B = \{i\} \cup A$. *If* $i \rightharpoonup A$, *then for every* $x \in X_{B^c}$, *there exists a function* $U_x : \Delta(X_B) \to \mathbb{R}$ *such that* (i) $p \succsim_x q$ *if and only if* $U_x(p) \geqslant U_x(q)$ *for all* $p, q \in \Delta(X_B)$; (ii) $U_x(p\alpha q) = \alpha U_x(p) + (1-\alpha)U_x(q)$ *for all* $\alpha \in (0,1)$ *and* $p, q \in \Delta(X_B)$ *with* $p_i \perp q_i$; (iii) *the function* $w_x : X_B \to \mathbb{R}$ *defined by* $w_x(y) = U_x(\delta_y)$ *for every* $y \in X_B$ *is continuous and strictly increasing; and* (iv) U_x *is unique up to a positive affine transformation.*

Step 2: Construct a (tight) hierarchy $\mathcal{H}$.

Consider the following procedure to construct a tight hierarchy $\mathcal{H}$:

Stage 0. We start with $\mathcal{H}_0 = \{I\}$.

Stage 1. Consider the following cases:

(1) If I is bracket separable, then denote by $\{A_k\}_{k=1}^n$ the finest bracket partition of I as defined in Lemma 5. Note that A_k is not bracket separable for every k. Let $\mathcal{H}_1 = \{A_k\}_{k=1}^n$ and move to the Stage 2.

(2) If I is not bracket separable, then $I \neq \bigcup_{B \rhd I} B$ (Lemma 4) and the contrapositive of Axiom 4 (separability under bracketing) implies $M(I) = \{i \in I : i \rightharpoonup I\} \neq \emptyset$. Write $M(I) = \{l_1, \ldots, l_n\}$. Denote $A_1 = I \setminus \{l_1\}$ and $A_i = A_{i-1} \setminus \{l_i\}$ for $i = 2, \ldots, n$. Note that $A_n = I \setminus M(I)$.

  (i) If $A_n = \emptyset$, then $n \geqslant 2$ and denote $\mathcal{H}_1 = \{A_1, \ldots, A_{n-1}\}$. The procedure terminates.

  (ii) If $A_n \neq \emptyset$ is not bracket separable, then let $\mathcal{H}_1 = \{A_1, \ldots, A_n\}$ and move to Stage 2.

  (iii) If $A_n \neq \emptyset$ is bracket separable, then denote by $\{B_k\}_{k=1}^m$ the finest bracket partition of A_n as defined in Lemma 5. Let $\mathcal{H}_1 = \{A_1, \ldots, A_{n-1}, B_1, \ldots, B_m\}$ and move to Stage 2.

Stage $t \geqslant 2$. Consider any $A \in \mathcal{H}_{t-1}$ such that $|A| \geqslant 2$ and A is *smallest*. That is, there is no $A' \in \mathcal{H}_{t-1}$ such that $A' \subsetneq A$. For instance, in case (2.ii) and case (2.iii) of Stage 1, $A_1, \ldots, A_{n-1}$ will not be smallest. There may be multiple smallest elements in $\mathcal{H}_{t-1}$. By construction, A is not bracket separable. Again by Axiom 4 (separability under bracketing), with an abuse of notation, $M(A) = \{l_1, \ldots, l_n\}$ for some $n \geqslant 1$. Again with an abuse of notation, denote $A_1 = A \setminus \{l_1\}$ and $A_i = A_{i-1} \setminus \{l_i\}$ for $i = 2, \ldots, n$. Note that $A_n = A \setminus M(A)$.

(i) If $A_n = \emptyset$, then $n \geqslant 2$ and let $\{A_1, \ldots, A_{n-1}\} \subseteq \mathcal{H}_t$.

(ii) If $A_n \neq \emptyset$ is not bracket separable, then let $\{A_1, \ldots, A_n\} \subseteq \mathcal{H}_t$.

(iii) If $A_n \neq \emptyset$ is bracket separable, then let $\{A_1, \ldots, A_{n-1}, B_1, \ldots, B_m\} \subseteq \mathcal{H}_t$, in which $\{B_k\}_{k=1}^m$ is the finest bracket partition of A_n as defined in Lemma 5.

Repeat the process for all smallest $A \in \mathcal{H}_{t-1}$ such that $|A| \geqslant 2$. Then, we obtain all elements of $\mathcal{H}_t$. Move on to Stage $t+1$.

This procedure terminates in finitely many stages n when all smallest sets in $\bigcup_{t=0}^n \mathcal{H}_t$ are singleton. Define $\mathcal{H} = \bigcup_{t=0}^n \mathcal{H}_t$. We claim that $\mathcal{H}$ is a tight hierarchy. First, $I \in \mathcal{H}_0 \in \mathcal{H}$. Second, by construction, any two elements A and B in the same $\mathcal{H}_k$ are either disjoint or satisfy $A \subseteq B$ or $B \subseteq A$. Now consider any $A \in \mathcal{H}_k$ and $B \in \mathcal{H}_{k'}$ with $k < k'$. Then we can find a unique smallest $B' \in \mathcal{H}_k$ such that $B \subsetneq B'$. Either $B' \subseteq A$, in which case $B \subsetneq A$, or $B' \cap A = \emptyset$, in which case $B \cap A = \emptyset$. Thus $\mathcal{H}$ is a hierarchy. Finally, observe that for all $A \in \mathcal{H} \setminus \{I\}$, we have $A = H(i)$ for some $i \in I$. Hence, $\mathcal{H}$ is a tight hierarchy.

Step 3: Construct the Bernoulli indices associated with each $A \in \mathcal{H}$.

By construction, $\tau(A)$ is a singleton set for every $A \in \mathcal{H}$ with $A \neq I$. Denote $\tau(A) = \{i_A\}$. We start with any smallest $A \in \mathcal{H}_n$, which must be a singleton set $A = \tau(A) = \{i_A\}$. By the construction of $\mathcal{H}$ and the definition of a bracket partition, $\succsim_z = \succsim_{z'}$ on $\Delta(X_A)$ for any $z, z' \in X_{A^c}$ such that $z_{\eta(A)} = z'_{\eta(A)}$. In other words, the conditional preference $\succsim_z$ only depends on $z_{\eta(A)}$. By Lemma 2, there exists $u^A : X_{\eta(A)} \times X_{i_A} \to [0,1]$ such that for all $z \in X_{\eta(A)}$ and $z' \in X_{(A \cup \eta(A))^c}$, (i) $u^A(z, \cdot\,) : X_{i_A} \to [0,1]$ is continuous and strictly increasing, and satisfies $u^A(z, \overline{x}_{i_A}) = 1$, $u^A(z, \underline{x}_{i_A}) = 0$; and (ii) if one defines $U_z^A : \Delta(X_{i_A}) \to [0,1]$ by $U_z^A(p) = \mathbb{E}^p(u^A(z,y))$, then for all $p, q \in \Delta(X_{i_A})$, $p \succsim_{(z,z')} q$ if and only if $U_z^A(p) \geqslant U_z^A(q)$. The normalization $u^A(z, \overline{x}_{i_A}) = 1$ and $u^A(z, \underline{x}_{i_A}) = 0$ works since $u^A(z, \cdot\,)$ is unique up to a positive affine transformation by Lemma 2. Clearly, $U_z^A(\delta_x) = u^A(z,x)$ is continuous and strictly increasing in $x \in X_A$.

Next, we define the Bernoulli index for every $A \in \mathcal{H}_n$ that is not the smallest recursively. Suppose all components in $\Phi(A)$ are smallest in $\mathcal{H}_n$. Again by the construction of $\mathcal{H}$ and the definition of a bracket partition, $\succsim_z = \succsim_{z'}$ on $\Delta(X_A)$ for any $z, z' \in X_{A^c}$ such that $z_{\eta(A)} = z'_{\eta(A)}$. Since $i_A \rightharpoonup A$, for any $z \in X_{\eta(A)}$ and $z' \in X_{(A \cup \eta(A))^c}$, Lemma 6 ensures the existence of $U_z^A : \Delta(X_A) \to [0,1]$ such that (i) $p \succsim_{(z,z')} q$ if and only if $U_z^A(p) \geqslant U_z^A(q)$ for all $p, q \in \Delta(X_A)$; (ii) $U_z^A(p \alpha q) = \alpha U_z^A(p) + (1 - \alpha) U_z^A(q)$ for all $\alpha \in (0,1)$ and $p, q \in \Delta(X_A)$ with $p_{i_A} \perp q_{i_A}$; and (iii) the function $w_z^A : X_A \to [0,1]$ defined by $w_z^A(y) = U_z^A(\delta_y)$ is continuous and strictly increasing, and satisfies $w_z^A(\overline{x}_A) = 1$ and $w_z^A(\underline{x}_A) = 0$. Define $u^A : X_{\eta(A)} \times X_{\tau(A)} \times [0,1]^{\Phi(A)} \to [0,1]$ by

$$u^A(\, z,\, y,\, (U^B_{(z,y)}(p_B))_{B \in \Phi(A)}\,) = U_z^A(\delta_y, p)$$

for all $z \in X_{\eta(A)}$, $y \in X_{\tau(A)}$, and $p \in \Delta_{X_{A \setminus \tau(A)}}$. The function u^A is well defined because $U^B_{(z,y)}(\delta_{\overline{x}_B}) = 1$, $U^B_{(z,y)}(\delta_{\underline{x}_B}) = 0$ for every $B \in \Phi(A)$, and $(\delta_y, p) \sim_z (\delta_y, (p_B)_{B \in \Phi(A)})$ following the definition of a bracket partition. Since $U_z^A(\delta_{\overline{x}_A}) = w_z^A(\overline{x}_A) = 1$ and $U_z^A(\delta_{\underline{x}_A}) = w_z^A(\underline{x}_A) = 0$, we know that $u^A(z, \overline{x}_{\tau(A)}, a) = 1$ if $a_B = 1$ for every $B \in \Phi(A)$ and $u^A(z, \underline{x}_{\tau(A)}, a) = 0$ if $a_B = 0$ for every $B \in \Phi(A)$. It is easy to see that U_z^A is onto. Also, for any $p \in \Delta(X_A)$, we have the recursive equation

$$U_z^A(p) = \mathbb{E}^p_{\tau(A)} U_z^A(\delta_y, p_{A \setminus \tau(A)|y}) = \mathbb{E}^p_{\tau(A)}\, u^A(\, z,\, y,\, (U^B_{(z,y)}(p_{B|y}))_{B \in \Phi(A)}\,). \quad (5)$$

Now consider any $A \in \mathcal{H}_n$ such that u^B and U_z^B have been defined for all $B \in \Phi(A)$ and $z \in X_{\eta(B)}$. Repeating the previous procedure, we can construct $u^A : X_{\eta(A)} \times X_{\tau(A)} \times [0,1]^{\Phi(A)} \to [0,1]$ and $U_z^A : \Delta(X_A) \to [0,1]$ for every $z \in X_{\eta(A)}$ such that (i) $p \succsim_{(z,z')} q$ if and only if $U_z^A(p) \geqslant U_z^A(q)$ for all $p, q \in \Delta(X_A)$ and $z' \in X_{(A \cup \eta(A))^c}$; (ii) $U_z^A(\delta_x)$ is continuous and strictly increasing in $x \in X_A$; (iii) U_z^A and $u^A(z, \cdot)$ are onto; and (iv) for any $p \in \Delta(X_A)$, the recursive equation (5) holds. Following this procedure, we can define u^A and U_z^A for all $A \in \mathcal{H}_n$ and $z \in X_{\eta(A)}$.

By induction, suppose for some $t \geqslant 2$ we have defined u^A and U_z^A for all $A \in \bigcup_{i=t}^n \mathcal{H}_i$ and $z \in X_{\eta(A)}$ that satisfy conditions (i)–(iv) above. Consider $A \in \mathcal{H}_{t-1}$. Since $t \geqslant 2$, we know $A \neq I$ and $M(A) = \{i_A\}$. Begin with A

being smallest in $\mathcal{H}_{t-1}$. Either $|A| = 1$ or $\Phi(A) \subseteq \bigcup_{i=t}^{n} \mathcal{H}_i$. In both cases we can repeat the previous construction for components in $\mathcal{H}_n$. Hence, we can define u^A and U_z^A for arbitrary $A \in \mathcal{H}_{t-1}$ and $z \in X_{\eta(A)}$ recursively that satisfy conditions (i)–(iv) above.

The induction works for $A \in \bigcup_{i=1}^{n} \mathcal{H}_i$ and I if $\tau(I) \neq \emptyset$. If instead $\tau(I) = \emptyset$, then I is bracket separable and $\Phi(I) = \{A_1, \ldots, A_m\}$ is the finest bracket partition of I. For any $p \in \Delta(X)$, we have $p \sim (p_{A_1}, \ldots, p_{A_m}) \sim (x_{A_1}, \ldots, x_{A_m})$, in which $U^{A_i}(\delta_{x_{A_i}}) = U^{A_i}(p_{A_i})$ for all $i = 1, \ldots, m$. Since $\succsim$ restricted to degenerate lotteries X is continuous and monotone, Debreu's Theorem implies that there exists a continuous and strictly increasing function $w^I : X \to [0,1]$ such that $w^I(\overline{x}) = 1, w^I(\underline{x}) = 0$, and $x \succsim y$ if and only if $w^I(x) \geqslant w^I(y)$ for all $x, y \in X$. Define $u^I : [0,1]^{\Phi(I)} \to [0,1]$ by

$$u^I(U^{A_1}(\delta_{x_{A_1}}), \ldots, U^{A_m}(\delta_{x_{A_m}})) = w^I(x)$$

for all $x \in X$. Because $U^{A_i}(\delta_{\overline{x}_{A_i}}) = 1, U^{A_i}(\delta_{\underline{x}_{A_i}}) = 0$, and $U^{A_i}(\delta_{x_{A_i}})$ is continuous and strictly increasing in $x_{B_i} \in X_{B_i}$ for all $i = 1, \ldots, K$, the function u^I is well defined, continuous, strictly increasing, and satisfies $u^I(1, \ldots, 1) = 1, u^I(0, \ldots, 0) = 0$. Define $U^I : \Delta(X) \to [0,1]$ by

$$U^I(p) = u^I(U^{A_1}(p_{A_1}), \ldots, U^{A_m}(p_{A_m}))$$

for every $p \in \Delta(X)$ and we know that U^I represents $\succsim$.

To conclude, $(\mathcal{H}, (u^A)_{A \in \mathcal{H}})$ is an HEU representation of $\succsim$. Indeed, it is a normalized HEU representation defined in Online Appendix III.

References

Baillon, A., Y. Halevy, and C. Li (2022): "Randomize at Your Own Risk: On the Observability of Ambiguity Aversion," *Econometrica*, 90, 1085–1107.

Bell, D. E. (1985): "Disappointment in Decision Making under Uncertainty," *Operations Research*, 33, 1–27.

Bommier, A., A. Kochov, and F. Le Grand (2017): "On Monotone Recursive Preferences," *Econometrica*, 85, 1433–1466.

Burstein, A., E. Morales, and J. Vogel (2019): "Changes in

Between-Group Inequality: Computers, Occupations, and International Trade," *American Economic Journal: Macroeconomics*, 11, 348–400.

Camara, M. K. (2021): "Computationally Tractable Choice," Working Paper.

Cappelli, V., S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and S. Minardi (2021): "Sources of Uncertainty and Subjective Prices," *Journal of the European Economic Association*, 19, 872–912.

Chew, S. H., G. Kader, and W. Wang (2023): "Rich Mixture Set, Process Preference, and Home Bias," Working Paper.

Darity Jr, W. A. (2022): "Position and Possessions: Stratification Economics and Intergroup Inequality," *Journal of Economic Literature*, 60, 400–426.

Debreu, G. (1960): "Topological Methods in Cardinal Utility Theory," in *Mathematial Methods in the Social Sciences*, Stanford Univ. Press, 16–26.

DeJarnette, P., D. Dillenberger, D. Gottlieb, and P. Ortoleva (2020): "Time Lotteries and Stochastic Impatience," *Econometrica*, 88, 619–656.

Dillenberger, D., D. Gottlieb, and P. Ortoleva (2020): "Stochastic Impatience and the Separation of Time and Risk Preferences," Working Paper.

Dominiak, A. and W. Schnedler (2011): "Attitudes Toward Uncertainty and Randomization: An Experimental Study," *Economic Theory*, 48, 289–312.

Duncan, G. T. (1977): "A Matrix Measure of Multivariate Local Risk Aversion," *Econometrica*, 895–903.

Eeckhoudt, L., B. Rey, and H. Schlesinger (2007): "A Good Sign For Multivariate Risk Taking," *Management Science*, 53, 117–124.

Elbers, C., P. Lanjouw, J. A. Mistiaen, and B. Özler (2008): "Reinterpreting Between-Group Inequality," *Journal of Economic Inequality*, 6, 231–245.

Ellis, A. and D. J. Freeman (2021): "Revealing Choice Bracketing," Working Paper.

Enke, B. and F. Zimmermann (2019): "Correlation Neglect in Belief Formation," *Review of Economic Studies*, 86, 313–332.

Epstein, L. G. and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Ergin, H. and F. Gul (2009): "A Theory of Subjective Compound Lotteries," *Journal of Economic Theory*, 144, 899–929.

Fishburn, P. C. and A. Rubinstein (1982): "Time Preference," *International Economic Review*, 677–694.

Fleurbaey, M. (2010): "Assessing Risky Social Situations," *Journal of Political Economy*, 118, 649–680.

Formby, J. P., T. G. Seaks, and W. J. Smith (1989): "On the Measurement and Trend of Inequality: A Reconsideration," *American Economic Review*, 79, 256–264.

Freeman, D. (2015): "Calibration Without Reduction for Non-Expected Utility," *Journal of Economic Theory*, 158, 21–32.

——— (2017): "Risk Taking With Background Risk Under Recursive Rank-Dependent Utility," *Mathematical Social Sciences*, 87, 72–74.

Fudenberg, D. and D. K. Levine (2012): "Fairness, Risk Preferences and Independence: Impossibility Theorems," *Journal of Economic Behavior & Organization*, 81, 606–612.

Gorman, W. M. (1968): "The Structure of Utility Functions," *Review of Economic Studies*, 35, 367–390.

Gottschalk, P. (1997): "Inequality, Income Growth, and Mobility: The Basic Facts," *Journal of Economic Perspectives*, 11, 21–40.

Grant, S. (1995): "A Strong (Ross) Characterization of Multivariate Risk Aversion," *Theory and Decision*, 38, 131–152.

Grant, S., A. Kajii, B. Polak, and Z. Safra (2010): "Generalized Utilitarianism and Harsanyi's Impartial Observer Theorem," *Econometrica*, 78, 1939–1971.

——— (2012): "Equally-Distributed Equivalent Utility, Ex Post Egalitarianism and Utilitarianism," *Journal of Economic Theory*, 147, 1545–1571.

Heo, Y. (2021): "Uncertainty Aversion with Multiple Issues," Working Paper.

Karni, E. (1979): "On Multivariate Risk Aversion," *Econometrica*, 1391–1401.

——— (1989): "Generalized Expected Utility Analysis of Multivariate Risk Aversion," *International Economic Review*, 297–305.

Ke, S. and Q. Zhang (2020): "Randomization and Ambiguity Aversion," *Econometrica*, 88, 1159–1195.

Keeney, R. L. (1973): "Risk Independence and Multiattributed Utility Functions," *Econometrica*, 27–34.

Kihlstrom, R. E. and L. J. Mirman (1974): "Risk Aversion with Many Commodities," *Journal of Economic Theory*, 8, 361–388.

——— (1981): "Constant, Increasing and Decreasing Risk Aversion with Many Commodities," *Review of Economic Studies*, 48, 271–280.

Kőszegi, B. and M. Rabin (2007): "Reference-Dependent Risk Attitudes," *American Economic Review*, 97, 1047–1073.

Kreps, D. M. and E. L. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200.

La Ferrara, E. (2002): "Inequality and Group Participation: Theory and Evidence From Rural Tanzania," *Journal of Public Economics*, 85, 235–273.

Lemieux, T. (2006): "Postsecondary Education and Increasing Wage Inequality," *American Economic Review*, 96, 195–199.

Levy, G. and R. Razin (2015): "Correlation Neglect, Voting Behavior, and Information Aggregation," *American Economic Review*, 105, 1634–1645.

Levy, H. and A. Levy (1991): "Arrow-Pratt Measures of Risk Aversion: The Multivariate Case," *International Economic Review*, 891–898.

Li, C., K. Rohde, and P. P. Wakker (2023): "The Deceptive Beauty of Monotonicity, and the Million-Dollar Question: Row-First or Column-First Aggregation," Working Paper.

Loomes, G. and R. Sugden (1986): "Disappointment and Dynamic Consistency in Choice under Uncertainty," *Review of Economic Studies*, 53, 271–282.

Mononen, L. (2022): "On Preference for Simplicity and Probability Weighting," Working Paper.

Mu, X., L. Pomatto, P. Strack, and O. Tamuz (2023a): "Background Risk and Small-Stakes Risk Aversion," Working Paper.

——— (2023b): "Monotone Additive Statistics," Working Paper.

Oechssler, J., H. Rau, and A. Roomets (2019): "Hedging, Ambiguity, and the Reversal of Order Axiom," *Games and Economic Behavior*, 117, 380–387.

Puri, I. (2022): "Simplicity and Risk," Working Paper.

Rabin, M. and G. Weizsäcker (2009): "Narrow Bracketing and Dominated Choices," *American Economic Review*, 99, 1508–1543.

Raiffa, H. (1961): "Risk, Ambiguity, and the Savage Axioms: Comment," *Quarterly Journal of Economics*, 75, 690–694.

Read, D., G. Loewenstein, and M. Rabin (1999): "Choice Bracketing," *Journal of Risk and Uncertainty*, 19, 171–97.

Richard, S. F. (1975): "Multivariate Risk Aversion, Utility Independence and Separable Utility Functions," *Management Science*, 22, 12–21.

Saito, K. (2013): "Social Preferences Under Risk: Equality of Opportunity Versus Equality of Outcome," *American Economic Review*, 103, 3084–3101.

——— (2015): "Preferences for Flexibility and Randomization Under Uncertainty," *American Economic Review*, 105, 1246–1271.

Schlee, E. E. (1990): "Multivariate Risk Aversion and Consumer Choice," *International Economic Review*, 737–745.

Selden, L. (1978): "A New Representation of Preferences Over 'Certain X Uncertain' Consumption Pairs: The 'Ordinal Certainty Equivalent Hypothesis'," *Econometrica*, 1045–1060.

Selden, L. and I. Stux (1978): "Consumption Trees, OCE Utility and the Consumption/Savings Decision," Working Paper.

Thaler, R. (1985): "Mental Accounting and Consumer Choice," *Marketing Science*, 4, 199–214.

Tversky, A. and D. Kahneman (1981): "The Framing of Decisions and the Psychology of Choice," *Science*, 211, 453–458.

Vorjohann, P. (2023): "Reference-Dependent Choice Bracketing," Working Paper.

Zhang, M. (2023): "Procedural Expected Utility," Working Paper.

Online Appendix

(For Online Publication Only)

This online appendix to "*Decision Making Under Multidimensional Risk*" is organized as follows. Section I includes additional examples. Section II provides a detailed discussion of brief remarks on the applications in Section 6. Section III provides additional results on the uniqueness of Bernoulli indices and the characterization of (generalized) recursive preferences with an exogenous order on dimensions. Section IV contains omitted proofs.

Online Appendix I: Additional Examples

I.1: More on the FATE and FETA Approaches

The two commonly used and yet opposite approaches to evaluate a risky multidimensional alternative, the FATE and the FETA approaches, have appeared in many different contexts, in addition to the inequality aversion example discussed in the Introduction. Below are a few others.

1. Suppose the decision maker is evaluating a risky consumption bundle that yields $(0, 1)$ and $(1, 0)$ with equal probability. She wants to use a constant-elasticity-of-substitution function u to aggregate the quantities of different goods. Should she use the FATE approach $\frac{1}{2}u(0, 1) + \frac{1}{2}u(1, 0)$ or the FETA approach $u(1/2, 1/2)$? The first approach may seem more natural, but the second may capture narrow bracketing and correlation neglect, which are often observed in people's choice behavior (Ellis and Freeman, 2021).

2. Consider a risky consumption sequence. If we simply compute the exponentially discounted expected utility, the decision maker will exhibit risk-seeking behavior in the time dimension, and hence more general evaluation approaches that can avoid such risk-seeking behavior have been proposed. Some first evaluate risk within each period and then aggregate across periods, and others first aggregate across periods and then take expectation. They are incompatible with each other, and it is not clear which approach is more appropriate.

3. Choice models under subjective uncertainty face the same dilemma. Interpret (x, y) as the decision maker's utility in states 1 and 2, respectively. When the decision maker is ambiguity-averse, whether to first compute expected utility for each state and then aggregate (using the maxmin aggregator introduced by Gilboa and Schmeidler (1989), for example) across states, or to first aggregate across states and then take expectation leads to opposite predictions about people's preference for randomization/hedging. The prediction of both approaches have been observed in different experiments.

I.2: Other Ways to Generalize the Three Approaches

We introduce two examples that also generalize the three approaches—FATE, FETA, and recursive—but in ways that differ from the HEU representation. We discuss which of our axioms are not satisfied by these examples. The first example features a convex combination of the FATE and FETA approaches.

Example 9. *Let $N = 2$. Suppose the decision maker's utility function is*

$$U(p) = \alpha\, \mathbb{E}^p \sqrt{x_1 + x_2} \;+\; (1-\alpha)\, \sqrt{(\mathbb{E}^p_1 \sqrt{x_1})^2 + (\mathbb{E}^p_2 \sqrt{x_2})^2},$$

in which $\alpha \in (0, 1)$. The idea of this utility function is simple, but the decision maker's preference represented by this utility function violates unidimensional independence and separability under bracketing. To see why unidimensional independence is violated, fix $x_2 = 1$ and note that the conditional preference $\succsim_{x_2}$ is represented by

$$U(p_1, 1) = \alpha\, \mathbb{E}^{p_1} \sqrt{x_1 + 1} \;+\; (1-\alpha)\, \sqrt{(\mathbb{E}^{p_1} \sqrt{x_1})^2 + 1},$$

which is not a monotone transformation of any expected utility function.

Recall that the FETA approach entails correlation neglect, since the decision maker ignores the interdependence of risk in different brackets. The next example combines expected utility with correlation neglect.

Example 10. *Let $N = 2$. Suppose the decision maker's utility function is*

$$U(p) = \sum_{x_1,x_2} \sqrt{x_1 + x_2}\, p_1(x_1)\, p_2(x_2).$$

When $p = (p_1, p_2)$, the above equation agrees with an expected utility function whose Bernoulli index is $u(x_1, x_2) = \sqrt{x_1 + x_2}$. For more general lotteries p, the decision maker is indifferent between p and (p_1, p_2), which reflects correlation neglect. Her preference satisfies all axioms in Section 3.2 except for separability under bracketing. To see this, let $x = (0, 1)$ and $y = (1, 0)$. Then $U(x) = U(y) = 1 > U(\delta_x \frac{1}{2} \delta_y) = U(\delta_1 \frac{1}{2} \delta_0, \delta_1 \frac{1}{2} \delta_0) = \frac{2+\sqrt{2}}{4}$. Since $x_1 \neq y_1$ and $x_2 \neq y_2$, we know $1 \not\vartriangleright \{1, 2\}$ and $2 \not\vartriangleright \{1, 2\}$. Moreover, the conditional preference $\succsim_{x_i}$ on dimension $-i$ clearly depends on x_i for both $i = 1, 2$, implying that $\{1\} \not\trianglerighteq \{1, 2\}$ and $\{2\} \not\trianglerighteq \{1, 2\}$. Hence, the preference violates separability under bracketing.

Online Appendix II: More on Applications

II.1: Multisource Income

In this section, we discuss several remarks related to Section 6.1.

First, if the decision maker strictly avoids multidimensional risk for some $p \in \Delta(Z^N)$ such that $p = (p_1, \ldots, p_N)$, then by Proposition 3, her preference cannot not be represented by $U(p) = \mathbb{E}^{f[p]} u(x)$ and u does not exhibit CARA. Hence, Propositions 1 and 2 imply that the decision maker violates both dominance and dominance without correlation. This observation reveals a connection between these two implications of narrow bracketing.

Second, in contrast to the representation in (2), Vorjohann (2023) and Camara (2021) study an alternative utility function of narrow bracketing:

$$\hat{U}(p) = \sum_{i=1}^{n} \mathbb{E}^{f[p_{A_i}]}\, u(x). \tag{6}$$

Indeed, (6) is an expected utility function, and hence an HEU representation with either hierarchy $\{I\}$ or hierarchy $\{I, A_1, \ldots, A_n\}$.

Fixing the Bernoulli index u and the brackets $\{A_i\}_{i=1}^{n}$, the decision maker with utility function (6) evaluates and adds up the *expected utilities* of risky income across different brackets, while (2) features the summation of their *certainty equivalents*.

When a decision maker faces a tuple of choice problems as in Example 7, in which choices in one dimension do not affect the availability of options in other dimensions, both (2) and (6) lead to the same predictions. However, in other choice scenarios, the decision maker whose preference is represented by (6) may violate dominance even when there is no risk, while (2) will not. To illustrate, consider two portfolios with two assets. Portfolio p delivers \$1 in both assets for sure, and portfolio q delivers \$2 in asset 1 and \$0 in asset 2 for sure. If the decision maker is risk-averse, which means u is strictly concave, she will strictly prefer P to Q because $2u(1) > u(0) + u(2)$, even though both portfolios deliver a total payoff of \$2 with certainty. Building such extreme departures from rationality into agents' behavior may result in a theory that explains certain anomalies in data at the expense of creating others that are unlikely to be present. By contrast, under representation (2), the decision maker will always choose more total money over less total money in the absence of risk.

Now we axiomatize the narrow bracketing representation (2). For a generalized bracketing representation characterized in Theorem 4 to take the functional form (2), two more axioms are necessary. The first is dominance over deterministic prospects: For any $x, y \in Z^N$, we have $x \succ y$ if and only if $\sum x_i > \sum y_i$. This property distinguishes representation (2) from representation (6). The second is symmetry: $\succsim_{i|0}=\succsim_{j|0}$ for all $i, j \in I$, in which $\succsim_{i|0}$ is the conditional preference for marginal lotteries in dimension i given δ_0 in all other dimensions. It guarantees that the risk attitudes are the same for all sources of income risk. The following proposition states that these two properties are also sufficient.

Proposition 6. *Suppose the preference $\succsim$ has a generalized bracketing representation. It admits representation (2) if and only if it satisfies dominance over deterministic prospects and symmetry.*

Proof of Proposition 6. The necessity of axioms is trivial, because $U(\delta_x) = \sum x_i$ and $\succsim_{i|0}$ is represented by an expected utility function with Bernoulli index u for all i. For sufficiency, suppose $\succsim$ is represented by

$$U(p) = v(\mathbb{E}^p_{A_1} u^{A_1}(x_{A_1}), \ldots, \mathbb{E}^p_{A_n} u^{A_n}(x_{A_n})). \tag{7}$$

By dominance over deterministic prospects, for every $i = 1, \ldots, n$ and $x_{A_i}, y_{A_i} \in X_{A_i}$ with $\sum_{l \in A_i} x_l = \sum_{l \in A_i} y_l$, we have $u^{A_i}(x_{A_i}) = u^{A_i}(y_{A_i})$. Then there exists a continuous and strictly increasing function u^i such that $u^{A_i}(x_{A_i}) = u^i(\sum_{l \in A_i} x_l)$ and hence $\mathbb{E}^p_{A_i} u^{A_i}(x_{A_i}) = \mathbb{E}^{f[p_{A_i}]} u^i(x)$. Again by dominance over deterministic prospects, (7) can be rewritten as

$$U(p) = \sum_{i=1}^{n} c(f[p_{A_i}], u^i). \tag{8}$$

Symmetry implies that we can choose u^i to be the same across all i, and hence (8) reduces to (2). This completes the proof. □

Third, we propose a notion of comparative avoidance of multidimensional risk. Since nontrivial avoidance of multidimensional risk for the entire domain is impossible by Proposition 3, we focus on gains. The analysis for losses is symmetric. Consider two decision makers whose preferences $\succsim_1$ and $\succsim_2$ can be represented by (2). Index decision maker 1's utility function by $(\{A^1_i\}^n_{i=1}, u^1)$ and decision maker 2's utility function by $(\{A^2_j\}^m_{j=1}, u^2)$. We say that $\succsim_1$ exhibits stronger avoidance of multidimensional risk for gains than $\succsim_2$ if $f[p] \succsim_2 q \Longrightarrow f[p] \succsim_1 q$ for all $p, q \in \Delta(Z^N_+)$ such that $p = (p_1, \ldots, p_N)$ and $q = (q_1, \ldots, q_N)$. When $\succsim_2$ satisfies strict avoidance of multidimensional risk for gains, the above comparison can be characterized by the coarseness of brackets of the two decision makers.

Proposition 7. *Suppose preferences $\succsim_1$ and $\succsim_2$ are represented by (2) with parameters $(\{A^1_i\}^n_{i=1}, u^1)$ and $(\{A^2_j\}^m_{j=1}, u^2)$, respectively, and the Arrow–Pratt measure of u^2 is strictly decreasing for $x \geqslant 0$. Then $\succsim_1$ exhibits stronger avoidance of multidimensional risk for gains than $\succsim_2$ if and only if (i) u^1 is a positive affine transformation of u^2 and (ii) $\{A^1_i\}^n_{i=1}$ is a finer partition of I than $\{A^2_j\}^m_{j=1}$.*

Proof of Proposition 7. For necessity, suppose $\succsim_1$ exhibits stronger avoidance of multidimensional risk for gains than $\succsim_2$. Normalize that $u^1(0) = u^2(0) = 0$ and $u^1(b) = u^2(b) = 1$. For any $z \in Z_+$, if $u^1(z) = \alpha > u^2(z) = \beta$, then choose p, q such that $p_1 = \delta_b \beta \delta_0$, $q_1 = \delta_z$, and $p_i = q_i = \delta_0$ for all $i > 1$. We have $f[p] \succsim_2 q \Longrightarrow f[p] \succsim_1 q$. That is, $\mathbb{E}_1^p u^1(x) = \beta \geqslant u^1(z) = \alpha$, a contradiction. A similar contradiction can be derived if $u^1(z) = \alpha < u^2(z) = \beta$. Hence, $u^1(z) = u^2(z)$ for all $z \in Z_+$. Denote the common Bernoulli index by u and the Arrow–Pratt measure by A. This proves (i). For (ii), suppose there exists A_i^1 such that $A_i^1 \cap A_j^2 \neq \emptyset$ and $A_i^1 \cap A_{j'}^2 \neq \emptyset$ for some $j \neq j'$. Fix $l \in A_i^1 \cap A_j^2$ and $l' \in A_i^1 \cap A_{j'}^2$. Since $A(x)$ is strictly decreasing, we can find $r \in \Delta(Z_+)$, $a \in Z_+$, and $\varepsilon > 0$ such that $f[r, \delta_a] \in \Delta(Z_+)$ and $c(r, u) + a < c(f[r, \delta_{a-\varepsilon}], u)$. Let $p, q \in \Delta(Z_+^N)$ such that $p_l = f[r, \delta_{a-\varepsilon}], p_{l'} = \delta_0$, $q_l = r, q_{l'} = \delta_a$, and $p_i = q_i = \delta_0$ for all $i \neq l, l'$. For decision maker 2, the utility of $f[p]$ is $c(f[r, \delta_{a-\varepsilon}], u)$, which is larger than $c(r, u) + a$, the utility of q. However, for decision maker 1, since l, l' are in the same bracket A_i^1, the utility of $f[p]$, $c(f[r, \delta_{a-\varepsilon}], u)$, is smaller than the utility of q, $c(f[r, \delta_a], u)$. That is, $f[p] \succ_2 q$ and $q \succ_1 f[p]$, a contradiction.

For sufficiency, we can normalize $u_1 = u_2 = u$. Since $\{A_i^1\}_{i=1}^n$ is finer than $\{A_j^2\}_{j=1}^m$, for each j, the collection $\{A_i^1 : A_i^1 \cap A_j^2 \neq \emptyset\}$ is a partition of A_j^2. Following the same proof of Proposition 3 establishes

$$U^1(q) = \sum_{i=1}^{n} c(f[q_{A_i^1}], u) \leqslant \sum_{j=1}^{m} c(f[q_{A_j^2}], u) = U^2(q).$$

Hence, $f[p] \succsim_2 q \Longrightarrow f[p] \succsim_1 q$ and $\succsim_1$ exhibits stronger avoidance of multidimensional risk for gains than $\succsim_2$. $\square$

Finally, we study an example in which the decision maker evaluates some sources of income recursively. Let dimension 1 represent the background risk and the other dimensions represent different income sources. Suppose the decision maker's utility function is

$$U(p) = \mathbb{E}_1^p \, v\Big(x + \sum_{i=1}^{n} c(f[p_{A_i|x}], u_x)\Big),$$

in which $\{A_i\}_{i=1}^n$ is a partition of $I \backslash \{1\}$, and v and u_x for each $x \in Z$ are continuous and strictly increasing. This is an HEU representation with hierarchy $\{I, I \backslash \{1\}, A_1, \ldots, A_n\}$. Under this representation, the decision maker evaluates the background risk, and conditioning on each realization of the background wealth, she takes the summation of certainty equivalents of the conditional distribution of total income in each bracket. We allow the decision maker's attitude toward income risk to depend on the background wealth and to differ from the attitude toward background risk. As a result, the decision maker cares about background risk and can be risk-averse over small gambles. Note that this is not contradictory to the impossibility result in Mu et al. (2023), because the decision maker with the above utility function may violate dominance, as discussed in Section 6.1.

II.2: Time Lotteries

In Section 6.2, we focused on random dated prizes and time lotteries, in which the decision maker receives a prize in at most one period for each realization. By Proposition 5, to accommodate stochastic impatience and risk aversion over time lotteries, the HEU representation should have hierarchy $\mathcal{H} = \{\{1,2\},\{2\}\}$—that is, the decision maker evaluates the risk in consumption and conditioning on its realizations, she evaluates the conditional risk in time. The utility of $p \in \Delta(X \times T)$ is

$$U(p) = \mathbb{E}_1^p \, \phi\Big(v(z)\, e^{-r\, c(p_{2|z}, u^2(z,\cdot))}\Big).$$

Denote by $\psi_z(a) = u^2(z, -\frac{1}{r}\ln a)$. The above function can be rewritten as

$$U(p) = \mathbb{E}^{p_1} \, \phi\Big[v(z)\, \psi_z^{-1}\Big(\mathbb{E}^{p_{2|z}}\, \psi_z(e^{-rt})\Big)\Big]. \tag{9}$$

In this section, we consider an extension of representation (9) that allows the delivery of multiple prizes over time. Unlike Section 6.2, we consider an intertemporal setting in which zero consumption is allowed and there are finitely many periods. That is, $Z = [0, b]$ with $b > 0$ and $T = \{1, \ldots, N\}$. A lottery $p \in \Delta(Z^T)$ represents a distribution over consumption streams.

We argue that the set of random dated prizes $\Delta(Z \times T)$ can be identified with a subset of $\Delta(Z^T)$. To see this, note that each $(z,t) \in Z \times T$ means receiving z in period t. If we interpret no consumption in some period as having consumption 0 in that period, then (z,t) (uniquely) corresponds to $x^{(z,t)} \in Z^T$, in which $x_t^{(z,t)} = z$ and $x_{t'}^{(z,t)} = 0$ for all $t' \neq t$. In this way, we can identify $p \in \Delta(Z \times T)$ with $\hat{p} \in \Delta(Z^T)$ such that $\hat{p} = \sum_{(z,t)} p(z,t)\delta_{x^{(z,t)}}$.

Hence, a lottery over dated prizes can be interpreted as a lottery over consumption streams such that each of the lottery's realizations only has nonzero consumption in at most one period. When there is no risk of confusion, we write $\Delta(Z \times T) \subsetneq \Delta(Z^T)$.

To extend representation (9) from $\Delta(Z \times T)$ to $\Delta(Z^T)$, the main challenge is that the same consumption may be received in multiple periods in one realization of the lottery. For instance, $(1, 2, 1) \in Z^3$ represents a consumption stream in which the decision maker receives 2 in period 2 and 1 in both periods 1 and 3. The following example illustrates how our generalization of (9) works, which will be formally introduced later.

Example 11. *Suppose $N = 2$ and consider a lottery p over consumption streams such that $p(1, 2) = p(1, 1) = 1/2$. Suppose $\phi(a) = a$ and $\psi_z(a) = 1/a$ in representation (9). That is, the parameterization is the same as Example 8. Our generalization of (9) evaluates p as follows:*

- *There are two possible prizes in p: 1 and 2. In the case of a realized prize 2, the decision maker receives it in period 2 with certainty, and the utility given prize 2 is $w(2) = v(2)e^{-2r}$.*

- *In the case of a realized prize 1, the decision maker receives it in only period 1 with probability 1/2 and in both periods 1 and 2 with probability 1/2. Therefore, the utility should be $v(1)e^{-r}$ with probability 1/2 and $v(1)(e^{-r} + e^{-2r})$ with probability 1/2. We then apply the risk preference over time ψ_z to the distribution over the summation of discount factors as in equation (9) and obtain the utility given prize 1:*

$$w(1) = v(1)\psi_z^{-1}\Big(\frac{1}{2}\psi_z(e^{-r}) + \frac{1}{2}\psi_z(e^{-r} + e^{-2r})\Big)\Big) = v(1)\frac{1}{\frac{1}{2}e^r + \frac{1}{2}\frac{1}{(e^{-r}+e^{-2r})}}.$$

- *Finally, the utility of p is the weighted sum of utilities given prize 2 and prize 1. To determine the weights, note that prize 2 will show up with probability 1/2 and prize 1 will show up regardless of the realization of p. Therefore, the weight for prize 2 is 1/2, the weight for prize 1 is 1, and the utility of p is $U(p) = \frac{1}{2} \times w(2) + 1 \times w(1)$.*

Example 11 illustrates how we decompose a lottery over consumption streams into risk in consumption and risk in time. Compared with the case before the extension (equation (9) in the domain of $\Delta(X \times T)$), there are two main differences. First, we may need to sum up multiple discount factors before multiplying it by v (the second bullet point above), because now a prize may show up multiple times in one realization. Second, the sum of the weights for the prizes may not be equal to 1 (the last bullet point above).

To emphasize these two differences more formally, consider an arbitrary $p \in \Delta(Z^T)$. For any $z \in Z$, denote by μ_1 the aggregate distribution of consumption. That is, the probability of receiving consumption z in at least one of the T periods is

$$\mu_1(z) = \sum_{\substack{y \in Z^T:\, \exists t \\ s.t.\ y_t = z}} p(y).$$

The distribution μ_1 is a measure of mass at most N that captures the risk in consumption. In Example 11, $\mu_1(1) = 1$ and $\mu_1(2) = 1/2$.

Conditioning on some consumption $z \in Z$ with $\mu_1(z) > 0$, the decision maker faces uncertainty regarding the delivery date(s) of z, which can be described as a probability distribution $\mu_{2|z}$ over nonempty subsets of T. That is, $\mu_{2|z}(B)$ is the probability for the decision maker to consume z exactly in periods $t \in B$ but not in periods $t \in B^{\mathsf{c}}$. In Example 11, $\mu_{2|1}(\{1\}) = \mu_{2|1}(\{1,2\}) = 1/2$ and $\mu_{2|2}(\{2\}) = 1$. To further illustrate this construction, consider more examples with $N = 2$.

Example 12. *(i) Let $p \in \Delta(Z^2)$ such that $p(1,1) = p(0,0) = 1/2$. The aggregation distribution over consumption is $\mu_1(1) = p(1,1) = 1/2$ and $\mu_1(0) = p(0,0) = 1/2$. Conditioning on either consumption 1 or 0, the decision maker receives it for sure in both periods: $\mu_{2|1} = \mu_{2|0} = \delta_{\{1,2\}}$.*

(ii) Let $q \in \Delta(Z^2)$ such that $q(1,0) = q(0,1) = 1/2$. Since the decision maker receives 1 and 0 in all realizations of consumption streams, $\mu_1(1) = \mu_1(0) = 1$. Conditioning on either consumption 1 or 0, the decision maker receives it in either period 1 or period 2 with equal probability 1/2. Hence, $\mu_{2|0} = \mu_{2|1} = \delta_{\{1\}} \frac{1}{2} \delta_{\{2\}}$.

(iii) Let $r \in \Delta(Z^2)$ such that $r(1,1) = r(0,0) = r(1,0) = r(0,1) = 1/4$. The aggregation distribution over consumption is $\mu_1(1) = r(1,1) + r(1,0) +$

$r(0,1) = 3/4$ *and* $\mu_1(0) = r(0,0) + r(0,1) + r(1,0) = 3/4$. *Conditioning on either consumption* 1 *or* 0, *the decision maker receives it in either period* 1, *period* 2, *or both periods, with equal probability* 1/3. *Hence,* $\mu_{2|0} = \mu_{2|1} = \frac{1}{3}\delta_{\{1\}} + \frac{1}{3}\delta_{\{2\}} + \frac{1}{3}\delta_{\{1,2\}}$.

Now we are ready to formally introduce our extension of (9) to $\Delta(X^T)$ in light of Example 11. Consider the following utility function for $p \in \Delta(X^T)$:

$$U(p) = \mathbb{E}^{\mu_1}\, \phi\Big[v(z)\, \psi_z^{-1}\Big(\mathbb{E}^{\mu_{2|z}}\, \psi_z(\sum_{t\in B} e^{-rt})\Big)\Big]. \tag{10}$$

Compared with (9), representation (10) features two differences. First, the aggregate distribution of consumption μ_1 is not necessarily a probability distribution. Second, $\mu_{2|z}$ is a probability distribution over nonempty subsets of T, instead of one over T. Because of the second difference, we also generalize e^{-rt} to $\sum_{t\in B} e^{-rt}$ in order to capture the aggregation over time.

To see why (10) is a natural extension of (9), consider a deterministic consumption stream $x = (x_1, \dots, x_N)$. For each $t = 1, \dots, N$, we have $\mu_1(x_t) = 1$ and $\mu_{2|x_t}(B_t) = 1$ in which $B_t = \{t' \in T : x_{t'} = x_t\}$. Hence, when ϕ is affine function, the utility of x is

$$U(x) = \mathbb{E}^{\mu_1} v(z) \sum_{t\in B_t} e^{-rt} = \sum_t v(x_t)\, e^{-rt},$$

which is exactly the exponentially discounted utility of a consumption sequence. Indeed, when ψ_z and ϕ are both affine functions, (10) reduces to the standard exponentially discounted expected utility model over $\Delta(X^T)$:

$$U(p) = \mathbb{E}^{\mu_1}\, v(z)\, \Big(\mathbb{E}^{\mu_{2|z}} \sum_{t\in B} e^{-rt}\Big) = \sum_{t=1}^{N} e^{-rt}\, \mathbb{E}^p_t\, v(x_t).$$

One might have noticed that there is another natural way to apply our HEU representation to $\Delta(Z^T)$, by interpreting different periods as different dimensions $I = T$. As in the discussion of dynamic choice in Section 1.1, our FATE representation corresponds to the generalized expected discounted utility functions of Kihlstrom and Mirman (1981) and Dillenberger et al. (2020); our FETA representation corresponds to the dynamic ordinal certainty equivalent model of Selden (1978) and Selden and Stux (1978);

and our recursive representation corresponds to the recursive models of Epstein and Zin (1989) and Kreps and Porteus (1978).

Under these two different methods of framing of dimensions, the HEU representation yields two different preferences. We consider this as a distinctive feature rather than a limitation of our theory, recognizing that the framing of dimensions may play a pivotal role in the evaluation process.

Online Appendix III: Additional Results

III.1: Uniqueness Properties of Bernoulli Indices

Suppose that $\succsim$ has an HEU representation $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$. We now explore the uniqueness properties of Bernoulli indices $(u^A)_{A\in\mathcal{H}}$ by fixing a tight hierarchy $\mathcal{H}$. For each $A \in \mathcal{H}$, we say that the function u^A is *normalized* if u^A is a mapping from $X_{\eta(A)} \times X_{\tau(A)} \times [0,1]^{\Phi(A)}$ to $[0,1]$ and $u^A(x,\cdot)$ is onto for all $x \in X_{\eta(A)}$. An HEU representation $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$ is *normalized* if u_A is normalized for all $A \in \mathcal{H}$. The following result shows that a normalized HEU representation exists and u^A is unique if $\tau(A) \neq \emptyset$.

Proposition 8. *If $\succsim$ has an HEU representation with a tight hierarchy $\mathcal{H}$, then it has a normalized HEU representation with hierarchy $\mathcal{H}$. If $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$ and $(\mathcal{H}, (\hat{u}^A)_{A\in\mathcal{H}})$ are both normalized HEU representations of $\succsim$, then $u^A = \hat{u}^A$ for all $A \in \mathcal{H} \setminus \{I\}$ and $u^I = \hat{u}^I$ if $\tau(I) \neq \emptyset$.*

Proof of Proposition 8. Suppose that $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$ is an HEU representation for $\succsim$. For each $A \in \mathcal{H}$ with $\tau(A) = A$ and $x \in X_{\eta(A)}$, we can apply a positive affine transformation to $u^A(x,\cdot)$ and derive $\hat{u}^A(x,\cdot)$ whose image is $[0,1]$. Clearly $\hat{u}^A$ is normalized. Then we adopt this procedure for other components inductively and derive a normalized $\hat{u}^A$ for other $A \in \mathcal{H}$. Since $\mathcal{H}$ is tight and $\hat{u}^A(x,\cdot)$ is unique up to a positive affine transformation for all $A \in \mathcal{H}$ with $\tau(A) \neq \emptyset$, we can derive the uniqueness properties of the normalized HEU representation as stated in the proposition. □

III.2: Exogenous Order on Dimensions

In the recursive and generalized recursive representations in Section 5, different decision makers are allowed to order dimensions in different ways. For instance, when $N = 2$, both hierarchies $\{\{1,2\},\{1\}\}$ and $\{\{1,2\},\{2\}\}$

are associated with some recursive representations. However, in some applications, a natural sequentiality is already built into the primitive. For example, in intertemporal settings, period t is before period $t+1$ for every t. In this case, it may be reasonable to sharpen predictions of the theory by imposing the following additional axiom.[30]

Axiom 6. *(Exogenous Order) For every $i < N$, we have $i \rightharpoonup \{i+1, \ldots, N\}$.*

The following corollary characterizes the implications of Axiom 6. Its proof is trivial and hence omitted.

Corollary 1. *Suppose the preference $\succsim$ has a generalized recursive representation. It satisfies exogenous order if and only if it has an HEU representation with hierarchy $\mathcal{H}$ such that $H(i+1) \subseteq H(i)$ for all $i = 1, \ldots, N-1$. If $\succsim$ has a recursive representation, then the above condition can be strengthened to $H(i+1) \subsetneq H(i)$ for all $i = 1, \ldots, N-1$.*

Online Appendix IV: Omitted Proofs

IV.1: Proof of Necessity of Axioms in Theorem 1

Necessity of axioms:

Suppose that $\succsim$ has an HEU representation $(\mathcal{H}, (u^A)_{A \in \mathcal{H}})$. Define U_x^A as in Definition 1 for all $A \in \mathcal{H}$ and $x \in X_{\eta(A)}$. Since $U^I(\delta_z)$ is continuous and strictly increasing in $z \in X$ and $p \succsim q$ if and only if $U^I(p) \geqslant U^I(q)$ for all $p, q \in \Delta(X)$, Axiom 1 (weak order) and Axiom 2 (outcome monotonicity) hold. The following lemma guarantees that $\succsim$ satisfies Axiom 5 (continuity).

Lemma 7. (i) *For any $A \in \mathcal{H}$ with $\tau(A) = A$, the function $u^A(x, y)$ is continuous and strictly increasing in $y \in X_A$ for all $x \in X_{\eta(A)}$.* (ii) *For any $A \in \mathcal{H}$ with $\tau(A) \neq A$, the function $u^A(x, y, a)$ is continuous and strictly increasing in $a \in \bigtimes_{B \in \Phi(A)} U_{x,y}^B(X_B)$ for all $x \in X_{\eta(A)}$ and $y \in X_A$.* (iii) *For any $A \in \mathcal{H}$, $x \in X_{A^c}$ and $p \in \Delta(X_A)$, there exists $z \in X_A$ such that $p \sim_x z$.* (iv) *The preference $\succsim$ satisfies Axiom 5 (continuity).*

[30] We can easily accommodate other exogenous orders by relabeling the dimensions.

Proof of Lemma 7. For any $A \in \mathcal{H}$ with $\tau(A) = A$ and $x \in X_{\eta(A)}$, we know $u^A(x, z) = U_x^A(\delta_z)$, which is continuous and strictly increasing by Definition 1. This proves (i).

For (ii), fix any $A \in \mathcal{H}$ with $\tau(A) \neq A$ and $x \in X_{\eta(A)}$ and $y \in X_A$. For any $a, a' \in \bigtimes_{B \in \Phi(A)} U_{x,y}^B(X_B)$ with $a \geqslant a'$ and $a \neq a'$, the strict monotonicity and continuity of $U_{x,y}^B$ on X_B guarantee the existence of $z_B, z'_B \in X_B$ for each $B \in \Phi(A)$ such that $z_B \geqslant z'_B$, $(z_B)_{B\in\Phi(A)} \neq (z'_B)_{B\in\Phi(A)}$, $a = (U_{x,y}^B(\delta_{z_B}))_{B\in\Phi(A)}$ and $a' = (U_{x,y}^B(\delta_{z'_B}))_{B\in\Phi(A)}$. Then by strict monotonicity of U_x^A on X_A, we have

$$u^A(x, y, a) = U_x^A(\delta_y, (\delta_{z_B})_{B\in\Phi(A)}) > U_x^A(\delta_y, (\delta_{z'_B})_{B\in\Phi(A)}) = u^A(x, y, a').$$

This implies that $u^A(x, y, a)$ is strictly increasing in a. For continuity, consider any sequence $(a^n)_{n\geqslant 1}$ with $a^n \in \bigtimes_{B\in\Phi(A)} U_{x,y}^B(X_B)$ and $a^n \to a \in \bigtimes_{B\in\Phi(A)} U_{x,y}^B(X_B)$. Again by strict monotonicity and continuity of $U_{x,y}^B$ on X_B, there exist $z_B^n, z_B \in X_B$ for each $B \in \Phi(A)$ and $n \geqslant 1$ such that $z_B^n \to z_B$, $a = (U_{x,y}^B(\delta_{z_B}))_{B\in\Phi(A)}$ and $a^n = (U_{x,y}^B(\delta_{z_B^n}))_{B\in\Phi(A)}$. Then by continuity of U_x^A on X_A, we have

$$u^A(x, y, a^n) = U_x^A(\delta_y, (\delta_{z_B^n})_{B\in\Phi(A)}) \to U_x^A(\delta_y, (\delta_{z_B})_{B\in\Phi(A)}) = u^A(x, y, a).$$

Hence, $u^A(x, y, a)$ is continuous in a.

Define a partition $(\mathcal{H}^k)_{k=1}^m$ of $\mathcal{H}$ as follows: $\mathcal{H}^1 = \{A \in \mathcal{H} \mid \tau(A) = A\}$ and for each $k \geqslant 1$, $\mathcal{H}^{k+1} = \{A \in \mathcal{H} \mid A \notin \bigcup_{i=1}^k \mathcal{H}^i \text{ and } \Phi(A) \subseteq \bigcup_{i=1}^k \mathcal{H}^i\}$. The iteration ends at $\mathcal{H}^m = \{I\}$ for some $m \geq 1$. We want to show (iii) by induction on k. For any $A \in \mathcal{H}^1$ and $x \in X_{\eta(A)}$, the utility of $p \in \Delta(X_A)$ is $U_x^A(p) = \mathbb{E}^p\, u^A(x, y)$. By (i), $u^A(x, y)$ is strictly increasing and continuous in y. Hence, there exists $z \in X_A$ such that $U_x^A(p) = U_x^A(\delta_z)$. Now assume that (iii) holds for all $A \in \bigcup_{i=1}^k \mathcal{H}^i$ for some $k \geqslant 1$. Consider any $A \in \mathcal{H}^{k+1}$ and $x \in X_{\eta(A)}$. The utility of $p \in \Delta(X_A)$ is

$$U_x^A(p) = \mathbb{E}^p_{\tau(A)}\, u^A(\, x,\ y,\ (U_{(x,y)}^B(p_{B|y}))_{B\in\Phi(A)}\,).$$

Since $\Phi(A) \subseteq \bigcup_{i=1}^k \mathcal{H}^i$, the inductive hypothesis implies the existence of $z_{B|y} \in X_B$ for all $y \in X_{\tau(A)}$ and $B \in \Phi(A)$ such that $U_{(x,y)}^B(p_{B|y}) =$

$U^B_{(x,y)}(z_{B|y})$. We can rewrite $U^A_x(p)$ as

$$U^A_x(p) = \mathbb{E}^p_{\tau(A)}\, U^A_x(\ \delta_y,\ (\delta_{z_{B|y}}))_{B\in\Phi(A)}\ ),$$

which is the expected utility of $p \in \Delta(X_A)$ under a continuous and strictly increasing Bernoulli index. Hence, we can find some $z \in X_A$ such that $U^A_x(p) = U^A_x(\delta_z)$. This proves the inductive hypothesis for $A \in \mathcal{H}^{k+1}$ and hence (iii) by induction.

For (iv), Axiom 5 (continuity) contains two components. For the second component, for any $p \in \Delta(X)$, by (iii), there exists $z \in X$ such that $p \sim z$. The sets $\{x \in X : x \succsim p\} = \{x \in X : U^I(\delta_x) \geqslant U^I(\delta_z)\}$ and $\{x \in X : p \succsim x\} = \{x \in X : U^I(\delta_x) \leqslant U^I(\delta_z)\}$ are closed because $U^I(\delta_y)$ is continuous in $y \in X$. To prove the first component, it suffices to show for any $p, q \in \Delta(X)$, the function $f : [0,1] \to \mathbb{R}$ defined by $f(\alpha) = U^I(p\alpha q)$ for all $\alpha \in [0,1]$ is continuous. Again we will prove it by induction on $(\mathcal{H}^k)^m_{k=1}$. First, for any $A \in \mathcal{H}^1$ and $x \in \text{supp}(p_{\eta(A)}) \cup \text{supp}(q_{\eta(A)})$,

$$U^A_x((p\alpha q)_{A|x}) = \mathbb{E}^{p\alpha q}_{A|x} u^A(x,y) = \lambda_x(\alpha) U^A_x(p_{A|x}) + (1-\lambda_x(\alpha)) U^A_x(q_{A|x}),$$

in which $\lambda_x(\alpha) = \frac{\alpha p_{\eta(A)}(x)}{\alpha p_{\eta(A)}(x) + (1-\alpha) q_{\eta(A)}(x)}$. Clearly, $U^A_x((p\alpha q)_{A|x})$ is continuous in α. Now assume $U^A_x((p\alpha q)_{A|x})$ is continuous in α for all $A \in \bigcup_{i=1}^k \mathcal{H}^i$ and $x \in \text{supp}(p_{\eta(A)}) \cup \text{supp}(q_{\eta(A)})$. Consider any $A \in \mathcal{H}^{k+1}$ and $x \in \text{supp}(p_{\eta(A)}) \cup \text{supp}(q_{\eta(A)})$,

$$\begin{aligned} U^A_x((p\alpha q)_{A|x}) =& \mathbb{E}^{p\alpha q}_{A|x} u^A(x, y,\ (U^B_{(x,y)}((p\alpha q)_{B|(x,y)}))_{B\in\Phi(A)}\ ) \\ =& \sum_y (p\alpha q)_{\tau(A)|x}(y) \cdot u^A(x, y,\ (U^B_{(x,y)}((p\alpha q)_{B|(x,y)}))_{B\in\Phi(A)}). \end{aligned}$$

Since $u^A(x,y,a)$ is continuous in a by (ii) and $\Phi(A) \subseteq \bigcup_{i=1}^k \mathcal{H}^i$, the inductive hypothesis implies that $u^A(x, y,\ (U^B_{(x,y)}((p\alpha q)_{B|(x,y)}))_{B\in\Phi(A)})$ is continuous in α for all $y \in X_{\tau(A)}$. In addition, $(p\alpha q)_{\tau(A)|x}(y) = \lambda_x(\alpha) p_{\tau(A)|x}(y) + (1-\lambda_x(\alpha)) q_{\tau(A)|x}(y)$ is continuous in α. Hence, $U^A_x((p\alpha q)_{A|x})$ is continuous in α for $A \in \mathcal{H}^{k+1}$. By induction, $f(\alpha) = U^I(p\alpha q)$ is continuous in α. □

By Lemma 1, for the rest of the proof, we assume $\mathcal{H}$ is tight.

To verify Axiom 3 (unidimensional independence), fix any $i \in I$ and let $A = H(i) \in \mathcal{H}$. For any $p, q \in \Delta(X_i)$ and $x \in X_{-i}$, we know $p \succ_x q$ if and only if $U^A_{x_{\eta(A)}}(p, \delta_{x_{A\setminus\{i\}}}) > U^A_{x_{\eta(A)}}(q, \delta_{x_{A\setminus\{i\}}})$. Since $U^A_{x_{\eta(A)}}(p\alpha r, \delta_{x_{A\setminus\{i\}}}) = \alpha U^A_{x_{\eta(A)}}(p, \delta_{x_{A\setminus\{i\}}}) + (1-\alpha) U^A_{x_{\eta(A)}}(r, \delta_{x_{A\setminus\{i\}}})$ for any $r \in \Delta(X_i)$ and $\alpha \in (0,1)$, we conclude that $p\alpha r \succ_x q\alpha r$.

To verify Axiom 4 (separability under bracketing), suppose $M(A) = \{i \in A : i \rightharpoonup A\} = \emptyset$. Identify B as the smallest element of $\mathcal{H}$ that includes A. Clearly $A \cap \tau(B) = \emptyset$. Then $\{A \cap C : C \in \Phi(B)\}$ (ignoring empty sets) is a partition of A. Since B is the smallest element of $\mathcal{H}$ that includes A, there exists at least two different (and hence disjoint) $C, C' \in \Phi(B)$ with $A \cap C \neq \emptyset$ and $A \cap C' \neq \emptyset$. Since $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$ is an HEU representation of $\succsim$, we have $C \rhd B \setminus \tau(B)$ for all $C \in \Phi(B)$ with $A \cap C \neq \emptyset$ and hence $A \cap C \rhd A$. Since $A \subseteq B \setminus \tau(B) = \bigcup_{C\in\Phi(B)} C$, we have $\bigcup_{B \rhd A} B = A$.

IV.2: Omitted Proofs of Lemmas in Appendix A

Proof of Lemma 2. By Axiom 1 (weak order), Axiom 2 (outcome monotonicity) and Axiom 3 (unidimensional independence), $\succsim_x$ admits an EU representation with a Bernoulli index $v_{i|x}$ defined on X_i, which is strictly increasing and unique up to a positive affine transformation. To see that $v_{i|x}$ is continuous, suppose by contradiction that there exists a sequence (y^n) in X_i such that $y^n \to y \in X_i$ and $v_{i|x}(y^n) \not\to v_{i|x}(y)$. Without loss of generality and passing to a subsequence if necessary, suppose $v_{i|x}(y^n) \to a < b = v_{i|x}(y)$ and $v_{i|x}(y^n) < (a+b)/2$ for every $n \geq 1$. Since $\succsim_{i|x}$ admits an EU representation, we can find $r \in \Delta(X_i)$ with $\mathbb{E}^r(v_{i|x}) = (a+b)/2$. That is, $y^n \prec_x r \prec_x y$ for every $n \geq 1$. Axiom 5 (continuity) implies that $y \precsim_x r \prec_x y$, a contradiction. Hence, $v_{i|x}$ is continuous for each $x \in X_{-i}$. □

Proof of Lemma 3. Without loss of generality, we can focus on the case in which $A = I$ and prove the results using induction on the cardinality of I. If $|I| = 1$, then both statements hold trivially by Lemma 2. Now suppose that both statements hold for $|I| \leqslant t$ for some $t \geqslant 1$. We need to show that they hold for $|I| = t + 1$.

If $\bigcup_{B \rhd I} B = I$, since I is finite, among those proper subsets of I, B's, we can find $A_1, \ldots, A_m$ such that $\bigcup_{i=1}^m A_i = I$. Note that by construction, $A_i \rhd I$ and $|A_i| \leqslant t$ for all $i = 1, \ldots, m$. Without loss of generality, we can let the union of any $m-1$ members among $A_1,\dots,A_m$ be a proper subset of I. Since $A_1 \rhd I$, we know that $p \sim (p_{A_1}, p_{A_1^c})$. Since $A_2 \rhd I$, we know that $p \sim (p_{A_1\setminus A_2}, p_{A_1\cap A_2}, p_{A_2\setminus A_1}, p_{(A_1\cup A_2)^c})$, and so on. Iteratively, we can derive a partition $\{B_k\}_{k=1}^n$ of I such that $p \sim (p_{B_1},\dots,p_{B_n}) := q$. Let $\{C_j^i\}_{j=1}^{n(i)}$ be a subset of $\{B_k\}_{k=1}^n$, for all $i = 1,\dots,m$, such that $A_i = \bigcup_{j=1}^{n(i)} C_j^i$, for all $i = 1,\dots,m$.

If p dominates x, then $p_{B_k} = q_{B_k}$ weakly dominates x_{B_k} for every k and the dominance is strict for some k^*. For every i, since $A_i \rhd I$, for any lottery $(r_{A_i}, r_{A_i^c})$, $\succsim$'s induced preference on $\Delta(X_{A_i})$ is independent of $r_{A_i^c}$. Hence, we can apply the inductive hypothesis to A_1 and know that $p \sim q \succsim (x_{C_1^1}, x_{C_2^1},\dots,x_{C_{n(1)}^1}, q_{A_1^c})$. Apply the inductive hypothesis to $A_2, A_3,\dots,A_m$ and follow the same argument iteratively. Then, we can conclude that $p \succsim (x_{B_1}, x_{B_2}, x_{B_3},\dots,x_{B_n}) = x$. Note that since at least one dominance relation is strict. We must have $p \succ x$. The proof for the case in which x dominates p is symmetric and omitted. This proves the first statement. The second statement can be proved similarly again by applying the inductive hypothesis to $A_1,\dots,A_m$ iteratively.

Next, suppose it cannot be the case that $\bigcup_{B\rhd I} B = I$. Then, by the contrapositive of Axiom 4 (separability under bracketing), there exists $i \in I$ with $i \rightharpoondown I$. Recall that we are using induction on the cardinality of I. Now for each cardinality of I, we will use another inductive argument based on the cardinality of $\text{supp}(p_i)$. If $|\text{supp}(p_i)| = 1$, then $p = (x, p_{-i})$ for some $x \in X_i$. By Axiom 2 (outcome monotonicity) and applying the inductive hypothesis to the conditional preference $\succsim_x$, the two statements hold trivially.

Assume that for some $n \geqslant 1$, the two statements hold if $|\text{supp}(p_i)| \leqslant n$. Suppose $|\text{supp}(p_i)| = n+1$ and p dominates x. Then, we can choose some $a_i \in \text{supp}(p_i) \setminus \{x_i\}$ with $a_i > x_i$ and write $p = p_i(a_i)(\delta_{a_i}, p_{-i|a_i}) + (1 - p_i(a_i))p'$, in which $|\text{supp}(p_i')| = n$. Note that $(\delta_{a_i}, p_{-i|a_i})$ dominates x and p' weakly dominates x, which implies that $p' \succsim x$ and $(\delta_{a_i}, p_{-i|a_i}) \succ x$. Clearly, $\delta_{a_i} \perp \delta_{x_i}$ and $\delta_{a_i} \perp p_i'$. Since $a_i > x_i$, by Axiom 5 (continuity) and Axiom 2 (outcome monotonicity), there exists some $x_i' > x_i$ and $x' = (x_i', x_{-i})$ such that $(\delta_{a_i}, p_{-i|a_i}) \succ x' \succ x$. Using the definition of $i \rightharpoondown I$, we obtain that $p = p_i(a_i)(\delta_{a_i}, p_{-i|a_i}) + (1 - p_i(a_i))p' \succsim p_i(a_i)(\delta_{a_i}, p_{-i|a_i}) + (1 - p_i(a_i))\delta_x \succ p_i(a_i)\delta_{x'} + (1 - p_i(a_i))\delta_x$. The first relation is $\sim$ if and only if $p' = x$.

Since x' and x agree in all dimensions other than i, Lemma 2 implies $p_i(a_i)\delta_{x'} + (1 - p_i(a_i))\delta_x \succ x$ and hence $p \succ x$. The case in which x dominates p is symmetric. This proves the first statement.

For the second statement, suppose that p dominates y and is dominated by x. By the first statement and Axiom 2 (outcome monotonicity), $\overline{x} \succ p \succ \underline{x}$. Denote $x^0 := \overline{x}, x^1 = (\underline{x}_1, \overline{x}_{-1}), x^2 = (\underline{x}_{\{1,2\}}, \overline{x}_{\{1,2\}^c}), \ldots,$ and $x^{t+1} = \overline{x}$. Then $x^0 \succ x^1 \succ \cdots \succ x^{t+1}$. We can find a unique $k = 0, \ldots, t$ such that $x^k \succsim p \succsim x^{k+1}$. By Axiom 5 (continuity), we can find $\alpha \in [0, 1]$ such that $p \sim \delta_{x^k}\ \alpha\ \delta_{x^{k+1}}$, which, by Lemma 2, is indifferent to some $z \in X$. Since $x \succ p \sim z \succ y$, again by Axiom 5 (continuity) and Axiom 2 (outcome monotonicity), there exists $z' \sim z$ with $x \geqslant z' \geqslant y$. By induction, we conclude that both statements hold for any finite cardinality of $\text{supp}(p)$ and I. This completes the proof. □

Proof of Lemma 4. For the first statement, suppose $B \rhd A$, $B' \rhd A$, and $B \cap B', B \setminus B', B' \setminus B$ are nonempty. We first verify that $B \cap B' \rhd A$. Take any $x \in X_{A^c}$, $r \in \Delta(X_{A \setminus (B \cap B')})$, and $p, q \in \Delta(X_A)$ such that $p_{A \setminus (B \cap B')} = q_{A \setminus (B \cap B')}$. Then $p_{A \setminus B} = q_{A \setminus B}$ and $p_{A \setminus B'} = q_{A \setminus B'}$. Since $B \rhd A$ and $B' \rhd A$, we have

$$
\begin{aligned}
& p \succsim_x q \\
\iff & (p_B, p_{A \setminus B}) \succsim_x (q_B, q_{A \setminus B}) \\
\iff & (p_{B \cap B'}, p_{A \setminus (B \cup B')}, p_{B \setminus B'}, p_{B' \setminus B}) \succsim_x (q_{B \cap B'}, q_{A \setminus (B \cup B')}, q_{B \setminus B'}, q_{B' \setminus B}) \\
\iff & (p_{B \cap B'}, r_{A \setminus (B \cup B')}, p_{B \setminus B'}, r_{B' \setminus B}) \succsim_x (q_{B \cap B'}, r_{A \setminus (B \cup B')}, q_{B \setminus B'}, r_{B' \setminus B}) \\
\iff & (p_{B \cap B'}, r_{A \setminus (B \cup B')}, r_{B \setminus B'}, r_{B' \setminus B}) \succsim_x (q_{B \cap B'}, r_{A \setminus (B \cup B')}, r_{B \setminus B'}, r_{B' \setminus B}) \\
\iff & (p_{B \cap B'}, r) \succsim_x (q_{B \cap B'}, r).
\end{aligned}
$$

Then we verify $B \setminus B' \rhd A$. The case for $B' \setminus B$ is symmetric and hence omitted. Take any $x \in X_{A^c}$, $r \in \Delta(X_{A \setminus (B \setminus B')})$, and $p, q \in \Delta(X_A)$ such that $p_{A \setminus (B \setminus B')} = q_{A \setminus (B \setminus B')}$. Since $B \rhd A, B' \rhd A$, and $B \cap B' \rhd A$, we can use $\succsim_x$ to represent the conditional preference on $\Delta(X_B), \Delta(X_{B'})$ and $\Delta(X_{B \cap B'})$ when there is no risk of confusion. Using the previous argument again, $p \succsim_x q$ if and only if

$$(p_{B\cap B'}, r_{A\setminus(B\cup B')}, p_{B\setminus B'}, \underline{x}_{B'\setminus B}) \succsim_x (q_{B\cap B'}, r_{A\setminus(B\cup B')}, q_{B\setminus B'}, \underline{x}_{B'\setminus B}).$$

We want to replace $p_{B\cap B'} = q_{B\cap B'}$ with $r_{B\cap B'}$ without changing the preference. This can be done immediately if $p_{B\cap B'} \sim_x r_{B\cap B'}$, because $B \cap B' \rhd A$. Without loss of generality, assume that $p_{B\cap B'} \succ_x r_{B\cap B'}$. First, by Lemma 3, there exist $y, y' \in X_{B\cap B'}$ such that $y \geqslant y'$ and $y \sim_x p_{B\cap B'} \succ_x r_{B\cap B'} \sim_x y'$. Since $B \cap B' \rhd A$, we have $p \succsim_x q$ if and only if

$$(y, r_{A\setminus(B\cup B')}, p_{B\setminus B'}, \underline{x}_{B'\setminus B}) \succsim_x (y, r_{A\setminus(B\cup B')}, q_{B\setminus B'}, \underline{x}_{B'\setminus B}).$$

Second, if $(y', \overline{x}_{B'\setminus B}) \succsim_x (y, \underline{x}_{B'\setminus B})$, then Axiom 5 (continuity) and Axiom 2 (outcome monotonicity) imply that there exists $z \in X_{B'\setminus B}$ such that $(y', z) \sim_x (y, \underline{x}_{B'\setminus B})$. Since $B \rhd A$, $B' \rhd A$, and $B \cap B' \rhd A$, we know that

$$\begin{aligned} & p \succsim_x q \\ \iff & (y', r_{A\setminus(B\cup B')}, p_{B\setminus B'}, z) \succsim_x (y', r_{A\setminus(B\cup B')}, q_{B\setminus B'}, z) \\ \iff & (r_{B\cap B'}, r_{A\setminus(B\cup B')}, p_{B\setminus B'}, z) \succsim_x (r_{B\cap B'}, r_{A\setminus(B\cup B')}, q_{B\setminus B'}, z) \\ \iff & (r_{B\cap B'}, r_{A\setminus(B\cup B')}, p_{B\setminus B'}, r_{B'\setminus B}) \succsim_x (r_{B\cap B'}, r_{A\setminus(B\cup B')}, q_{B\setminus B'}, r_{B'\setminus B}) \\ \iff & (p_{B\setminus B'}, r) \succsim_x (q_{B\setminus B'}, r). \end{aligned}$$

Third, suppose instead $(y, \underline{x}_{B'\setminus B}) \succ_x (y', \overline{x}_{B'\setminus B})$. Then there exist $y'', y''' \in X_{B\cap B'} \setminus \{\overline{x}_{B\cap B'}, \underline{x}_{B\cap B'}\}$ such that $y'' > y'''$, $(y'', \overline{x}_{B'\setminus B}) \succ_x (y, \underline{x}_{B'\setminus B})$, and $(y''', \underline{x}_{B'\setminus B}) \prec_x (y', \overline{x}_{B'\setminus B})$. For each $\hat{y} \in X_{B\cap B'} \setminus \{\overline{x}_{B\cap B'}, \underline{x}_{B\cap B'}\}$, let $\Gamma(\hat{y}) = \{\hat{y}' \in X_{B\cap B'} : (\hat{y}, \overline{x}_{B'\setminus B}) \succ_x (\hat{y}', \overline{x}_{B'\setminus B}) \succ_x (\hat{y}, \underline{x}_{B'\setminus B})\}$. By Axiom 5 (continuity), $\{\Gamma(\hat{y})\}_{y\in X_{B\cap B'}}$ is an open cover of $\{z \in X_{B\cap B'} : y''' \leq z \leq y''\}$, which is a compact set. Hence, it admits a finite subcover. This ensures the existence of a finite sequence $(y^k)_{k=0}^n$ in $\{z \in X_{B\cap B'} : y''' \leq z \leq y''\} \cup \{y\}$ such that (i) $y^0 = y$, (ii) $(y^{k+1}, \overline{x}_{B'\setminus B}) \sim_x (y^k, \underline{x}_{B'\setminus B})$ for each $k = 0, \ldots, n-1$, and (iii) $(y', \overline{x}_{B'\setminus B}) \succsim_x (y^n, \underline{x}_{B'\setminus B})$. By applying (ii) and the implications of $B \rhd A$ and $B' \rhd A$ repeatedly, we obtain

$$\begin{aligned}
& p \succsim_x q \\
\iff & (y, r_{A\setminus(B\cup B')}, p_{B\setminus B'}, \underline{x}_{B'\setminus B}) \succsim_x (y, r_{A\setminus(B\cup B')}, q_{B\setminus B'}, \underline{x}_{B'\setminus B}) \\
\iff & (y^1, r_{A\setminus(B\cup B')}, p_{B\setminus B'}, \overline{x}_{B'\setminus B}) \succsim_x (y^1, r_{A\setminus(B\cup B')}, q_{B\setminus B'}, \overline{x}_{B'\setminus B}) \\
\iff & (y^1, r_{A\setminus(B\cup B')}, p_{B\setminus B'}, \underline{x}_{B'\setminus B}) \succsim_x (y^1, r_{A\setminus(B\cup B')}, q_{B\setminus B'}, \underline{x}_{B'\setminus B}) \\
\iff & (y^2, r_{A\setminus(B\cup B')}, p_{B\setminus B'}, \overline{x}_{B'\setminus B}) \succsim_x (y^2, r_{A\setminus(B\cup B')}, q_{B\setminus B'}, \overline{x}_{B'\setminus B}) \\
\iff & (y^2, r_{A\setminus(B\cup B')}, p_{B\setminus B'}, \underline{x}_{B'\setminus B}) \succsim_x (y^2, r_{A\setminus(B\cup B')}, q_{B\setminus B'}, \underline{x}_{B'\setminus B}) \\
\iff & \cdots \\
\iff & (y^n, r_{A\setminus(B\cup B')}, p_{B\setminus B'}, \underline{x}_{B'\setminus B}) \succsim_x (y^n, r_{A\setminus(B\cup B')}, q_{B\setminus B'}, \underline{x}_{B'\setminus B}).
\end{aligned}$$

Since we have $(y', \overline{x}_{B'\setminus B}) \succsim_x (y^n, \underline{x}_{B'\setminus B})$ by (iii), we can apply the argument in the previous case to establish that $p \succsim_x q \iff (p_{B\setminus B'}, r) \succsim_x (q_{B\setminus B'}, r)$. This completes the proof for $B \setminus B' \rhd A$ and hence the first statement is true.

For the second statement, suppose $A = \bigcup_{B \rhd A} B$. We can follow the construction of the partition $\{B_k\}_{k=1}^n$ of A in the proof of Lemma 3. Then, the first statement we establish above implies the second. □

Proof of Lemma 6. To prove Lemma 6, we need some intermediate results. Without loss of generality, assume $i \notin A$.

Lemma 8. *Denote $B = \{i\} \cup A$. If $i \rightharpoonup A$, then for all $\alpha \in (0,1)$, $x \in X_{B^c}$, and $p, q, r, s \in \Delta(X_B)$ such that $p_i \perp r_i$ and $q_i \perp s_i$, the following properties hold:* (i) $p \succ_x r \implies p \succ_x p\alpha r \succ_x r$; (ii) $p \sim_x r \implies p \sim_x p\alpha r \sim_x r$; (iii) $p \sim_x q, r \sim_x s \implies p\alpha r \sim_x q\alpha s$; *and* (iv) $p \succ_x q, r \succ_x s \implies p\alpha r \succ_x q\alpha s$.

Proof of Lemma 8. For (i), we consider four cases. First, if $p = \overline{x}_B$ and $r = \underline{x}_B$, then the result is implied by Lemma 3. Second, if $p = \overline{x}_B$ and $r \succ_x \underline{x}_B$, then by Axiom 5 (continuity) and Axiom 2 (outcome monotonicity), we can find $\varepsilon \in \mathbb{R}^B_+$ such that $\varepsilon_i > 0$, $\varepsilon_j = 0$ for all $j \in A$, and $\overline{x}_B \succ_x \overline{x}_B - \varepsilon \succ_x r$. By the definition of $i \rightharpoonup A$ and Lemma 2, we have $p\alpha r \prec_x \delta_{\overline{x}_B}\, \alpha\, \delta_{\overline{x}_B - \varepsilon} \prec_x \overline{x}_B = p$. As $r \succ_x \underline{x}_B$, we can find $y, y' \in X_B$ such that $p \succ_x y \succ_x r \sim_x y'$, $y_i, y'_i \notin \operatorname{supp}(p_i) \cup \operatorname{supp}(r_i)$, $y_i \neq y'_i$, and $y_j = y'_j$ for all $j \in A$. Again by the definition of $i \rightharpoonup A$ and Lemma 2, $p\alpha r \succ_x \delta_y \alpha \delta_{y'} \succ_x y' \sim_x r$. Third, if $p \prec_x \overline{x}_B$ and $r = \underline{x}_B$, then the proof is symmetric to the second case. Finally, if $\overline{x}_B \succ_x p \succ_x r \succ_x \underline{x}_B$, then the proof is a simple combination of those of the above two cases.

For (ii), if $p \sim_x r$ and $p_i \perp r_i$, then $\overline{x}_B \succ_x p \sim_x r \succ_x \underline{x}_B$. By Axiom 5 (continuity) and Axiom 2 (outcome monotonicity), we can find $y, y' \in X_B$ such that $y \succ_x p \sim_x r \succ_x y'$, $y_i \neq y'_i$, and $y_i, y'_i \notin \text{supp}(p_i) \cup \text{supp}(r_i)$. For any $\beta \in (0,1)$, by applying (i) twice we get $p\beta\delta_y \succ_x p \sim_x r \succ_x p\beta\delta_{y'}$. Then apply (i) twice for $p\beta\delta_y, q$ and $p\beta\delta_{y'}$ and we derive $(p\beta\delta_y)\alpha r \succ_x r \succ_x (p\beta\delta_{y'})\alpha r$. Let β go to 1, and by Axiom 5 (continuity), $p\alpha r \succsim_x r \succsim_x p\alpha r$, which implies $p\alpha r \sim_x r \sim_x p$.

For (iii), if $p, r \in \{\overline{x}_B, \underline{x}_B\}$, then $p = q$ and $r = s$ and the result is trivial. Without loss of generality, assume that $\overline{x}_B \succ_x p \sim_x q \succ_x \underline{x}_B$. Using the argument in the proof for (ii), we can find y, y' such that $y \succ_x p \sim_x q \succ_x y'$, $y_i \neq y'_i$, $y_i, y'_i \notin \text{supp}(p_i) \cup \text{supp}(r_i)$, and $p\beta\delta_y \succ_x p \sim_x q \succ_x p\beta\delta_{y'}$ for every $\beta \in (0,1)$. Applying the definition of $i \rightharpoonup A$ twice, we obtain that $(p\beta\delta_y)\alpha r \succ_x q\alpha s \succ_x (p\beta\delta_{y'})\alpha r$. Let β go to 1, and by Axiom 5 (continuity), we have $p\alpha r \succsim_x q\alpha s \succsim_x p\alpha r$, which implies that $p\alpha r \sim_x q\alpha s$.

For (iv), it suffices to consider the case that $p \succ_x q \succ_x r \succ_x s$, as the other cases are implied by (i). There exists $y \in X_B$ such that $r \sim_x y$ and $y_i \notin \text{supp}(q_i)$. Applying the definition of $i \rightharpoonup A$ twice, we obtain that $p\alpha r \succ_x q\alpha\delta_y \succ q\alpha s$. $\square$

The following lemma describes a situation in which independence holds even if the condition of disjoint supports fails.

Lemma 9. *Denote $B = \{i\} \cup A$. If $i \rightharpoonup A$, then for all $\alpha \in (0,1)$, $x \in X_{B^c}$, and $p, q, r, s \in \Delta(X_B)$ such that $p_i \perp r_i$ and $\text{supp}(q) \cup \text{supp}(s) \subseteq \{\overline{x}_B, \underline{x}_B\}$, then $p \sim_x q, r \sim_x s \implies p\alpha r \sim_x q\alpha s$.*

Proof of Lemma 9. First, if $p, r \in \{\overline{x}_B, \underline{x}_B\}$, then $p = q$, $r = s$ and the result is trivial. Without loss of generality, assume that $\overline{x}_B \succ_x p \sim_x q \succ_x \underline{x}_B$. Then there exists $y \in X_B$ such that $y_i \notin \{\overline{x}_i, \underline{x}_i\}$ and $p \sim_x y$. Since $p \sim_x y$, $r \sim_x s$, $p_i \perp r_i$, and $y_i \perp s_i$, by part (iii) of Lemma 8, we have $p\alpha r \sim_x \delta_y \alpha s$. Hence, it suffices to show that $\delta_y \alpha s \sim_x q\alpha s$ for every $y \sim q$ with $y_i \notin \{\overline{x}_i, \underline{x}_i\}$. As $\overline{x}_B \succ_x p \sim_x q \succ_x \underline{x}_B$, by Axiom 5 (continuity), we can find $\varepsilon \in \mathbb{R}^B_+$ and $\gamma \in (0,1)$ such that $\varepsilon_i > 0$, $\varepsilon_j = 0$ for all $j \in A$, $\overline{x}_B - \varepsilon \succ_x y \succ_x \underline{x}_B + \varepsilon$, and $y \sim_x q \sim_x \delta_{\overline{x}_B - \varepsilon} \, \gamma \, \delta_{\underline{x}_B + \varepsilon}$. Denote

$\hat{q} = \delta_{\overline{x}_B - \varepsilon}\, \gamma\, \delta_{\underline{x}_B + \varepsilon}$ and $q^\beta = q\beta\hat{q}$ for each $\beta \in (0,1)$. Part (ii) of Lemma 8 implies that $q^\beta \sim_x q \sim_x y$.

We claim that $\delta_y \alpha s \sim_x q^\beta \alpha s$ for all $\beta, \alpha \in (0,1)$. To see this, first note that $q(\overline{x}_B) > 0$ and $q(\underline{x}_B) > 0$ as $\overline{x}_B \succ_x q \succ_x \underline{x}_B$. Then

$$
\begin{aligned}
q^\beta &= [\beta q(\overline{x}_B)\delta_{\overline{x}_B} + (1-\beta)\gamma\delta_{\overline{x}_B-\varepsilon}] + [\beta q(\underline{x}_B)\delta_{\underline{x}_B} + (1-\beta)(1-\gamma)\delta_{\underline{x}_B+\varepsilon}] \\
&\sim_x [\beta q(\overline{x}_B) + (1-\beta)\gamma]\delta_{\overline{x}_B-\varepsilon'} + [\beta q(\underline{x}_B) + (1-\beta)(1-\gamma)]\delta_{\underline{x}_B+\varepsilon''},
\end{aligned}
$$

in which the indifference follows from part (iii) of Lemma 8, and $\varepsilon', \varepsilon'' \in \mathbb{R}^B_+$ satisfy $\varepsilon'_i, \varepsilon''_i > 0$, $\varepsilon'_j = \varepsilon''_j = 0$ for all $j \in A$, and

$$
\begin{aligned}
\delta_{\overline{x}_B-\varepsilon'} &\sim_x \frac{\beta q(\overline{x}_B)}{\beta q(\overline{x}_B) + (1-\beta)\gamma}\delta_{\overline{x}_B} + \frac{(1-\beta)\gamma}{\beta q(\overline{x}_B) + (1-\beta)\gamma}\delta_{\overline{x}_B-\varepsilon}, \\
\delta_{\underline{x}_B+\varepsilon''} &\sim_x \frac{\beta q(\underline{x}_B)}{\beta q(\underline{x}_B) + (1-\beta)(1-\gamma)}\delta_{\underline{x}_B} + \frac{(1-\beta)(1-\gamma)}{\beta q(\underline{x}_B) + (1-\beta)(1-\gamma)}\delta_{\underline{x}_B+\varepsilon}.
\end{aligned}
$$

The existence of $\varepsilon', \varepsilon''$ is guaranteed by Lemma 2. Denote $\hat{q}^\beta := [\beta q(\overline{x}_B) + (1-\beta)\gamma]\delta_{\overline{x}_B-\varepsilon'} + [\beta q(\underline{x}_B) + (1-\beta)(1-\gamma)]\delta_{\underline{x}_B+\varepsilon''}$. Then $\hat{q}^\beta \sim_x q \sim_x y$. Note that

$$
\begin{aligned}
q^\beta \alpha s =& \Big[\alpha(\beta q(\overline{x}_B)\delta_{\overline{x}_B} + (1-\beta)\gamma\delta_{\overline{x}_B-\varepsilon}) + (1-\alpha)s(\overline{x}_B)\delta_{\overline{x}_B}\Big] \\
&+ \Big[\alpha(\beta q(\underline{x}_B)\delta_{\underline{x}_B} + (1-\beta)(1-\gamma)\delta_{\underline{x}_B+\varepsilon}) + (1-\alpha)s(\underline{x}_B)\delta_{\underline{x}_B}\Big].
\end{aligned}
$$

Again by applying Lemma 2 to the two terms above, respectively, and applying part (iii) of Lemma 8, we derive

$$
\begin{aligned}
q^\beta \alpha s \sim_x& \Big[\alpha(\beta q(\overline{x}_B) + (1-\beta)\gamma)\delta_{\overline{x}_B-\varepsilon'} + (1-\alpha)s(\overline{x}_B)\delta_{\overline{x}_B}\Big] \\
&+ \Big[\alpha(\beta q(\underline{x}_B) + (1-\beta)(1-\gamma))\delta_{\underline{x}_B+\varepsilon''} + (1-\alpha)s(\underline{x}_B)\delta_{\underline{x}_B}\Big] \\
=& \hat{q}^\beta \alpha s.
\end{aligned}
$$

Note that $\hat{q}^\beta_i \perp s_i$, $y_i \perp s_i$, and $\hat{q}^\beta \sim_x y$. Part (ii) of Lemma 8 implies that $\delta_y \alpha s \sim_x \hat{q}^\beta \alpha s \sim_x q^\beta \alpha s$, which holds for all $\alpha, \beta \in (0,1)$. Let β approach 1 and by Axiom 5 (continuity), we conclude that $q\alpha s \sim \delta_y \alpha s$. This completes the proof. □

For any $p \succ_x q$, the next result provides sufficient conditions for $p\alpha q$ to be better as $\alpha \in (0,1)$ increases.

Lemma 10. *Denote $B = \{i\} \cup A$. If $i \rightharpoonup A$, then for all $\alpha, \beta \in (0,1)$, $x \in X_{B^c}$, and $p, q \in \Delta(X_B)$ such that $\alpha > \beta$, $p \succ_x q$, and $p_i \perp q_i$, we have* (i) $\delta_{\overline{x}_B}\alpha\delta_{\underline{x}_B} \succ_x \delta_{\overline{x}_B}\beta\delta_{\underline{x}_B}$ *and* (ii) $p\alpha q \succ_x p\beta q$.

Proof of Lemma 10. For (i), first, note that $\delta_{\overline{x}_B}\beta\delta_{\underline{x}_B} = (\delta_{\overline{x}_B}\alpha\delta_{\underline{x}_B})\frac{\beta}{\alpha}\delta_{\underline{x}_B}$. By Lemma 3, there exists $y \in X_B$ such that $y_i \neq \overline{x}_i$, $y_i \neq \underline{x}_i$, and $y \sim_x \delta_{\overline{x}_B}\alpha\delta_{\underline{x}_B}$. By Lemma 9, we have $\delta_{\overline{x}_B}\beta\delta_{\underline{x}_B} = (\delta_{\overline{x}_B}\alpha\delta_{\underline{x}_B})\frac{\beta}{\alpha}\delta_{\underline{x}_B} \sim_x \delta_y\frac{\beta}{\alpha}\delta_{\underline{x}_B}$. Since $y \succ_x \underline{x}_B$ and $y_i \neq \underline{x}_i$, part (i) of Lemma 8 implies that $\delta_y\frac{\beta}{\alpha}\delta_{\underline{x}_B} \prec_x y \sim_x \delta_{\overline{x}_B}\alpha\delta_{\underline{x}_B}$. Hence, $\delta_{\overline{x}_B}\alpha\delta_{\underline{x}_B} \succ_x \delta_{\overline{x}_B}\beta\delta_{\underline{x}_B}$.

For (ii), by part (i), we can find unique $\gamma^1, \gamma^2 \in [0,1]$ such that $\gamma^1 > \gamma^2$ and $p \sim_x \delta_{\overline{x}_B}\gamma^1\delta_{\underline{x}_B} \succ_x q \sim_x \delta_{\overline{x}_B}\gamma^2\delta_{\underline{x}_B}$. Then Lemma 9 implies

$$p\alpha q \sim_x \delta_{\overline{x}_B}(\alpha\gamma^1 + (1-\alpha)\gamma^2)\delta_{\underline{x}_B},$$
$$p\beta q \sim_x \delta_{\overline{x}_B}(\beta\gamma^1 + (1-\beta)\gamma^2)\delta_{\underline{x}_B}.$$

Since $\alpha > \beta$ and $\gamma^1 > \gamma^2$, we know that $\alpha\gamma^1 + (1-\alpha)\gamma^2 > \beta\gamma^1 + (1-\beta)\gamma^2$ and hence $p\alpha q \succ_x p\beta q$ by part (i). □

Now we are ready to prove Lemma 6. For any $p \in \Delta(X_B)$, by Lemma 10, there exists a unique $\alpha(p) \in [0,1]$ such that $p \sim_x \delta_{\overline{x}_B}\alpha(p)\delta_{\underline{x}_B}$. Define $U_x : \Delta(X_B) \to \mathbb{R}$ such that $U_x(p) = \alpha(p)$ for every $p \in \Delta(X_B)$. Then $U_x(\delta_{\overline{x}_B}) = 1$ and $U_x(\delta_{\underline{x}_B}) = 0$. Lemma 10 ensures that $p \succsim_x q$ if and only if $U_x(p) \geqslant U_x(q)$ for all $p, q \in \Delta(X_B)$. Now we check condition (ii). Fix any $\alpha \in (0,1)$ and $p, q \in \Delta(X_B)$ with $p_i \perp q_i$. By definition of U_x, we know that $p \sim_x \delta_{\overline{x}_B}U_x(p)\delta_{\underline{x}_B}$ and $q \sim_x \delta_{\overline{x}_B}U_x(q)\delta_{\underline{x}_B}$. Since $p_i \perp q_i$, Lemma 9 implies $p\alpha q \sim_x \delta_{\overline{x}_B}(\alpha U_x(p) + (1-\alpha)U_x(q))\delta_{\underline{x}_B}$. Again, the definition of U_x implies $p\alpha q \sim_x \delta_{\overline{x}_B}U_x(p\alpha q)\delta_{\underline{x}_B}$. By Lemma 10, we conclude that $U_x(p\alpha q) = \alpha U_x(p) + (1-\alpha)U_x(q)$. Hence, $U_x(p) = \sum_{y_i} U_x(\delta_{y_i}, p_{B\setminus\{i\}|y_i})p_i(y_i)$. To verify (iii), define $w_x : X_B \to \mathbb{R}$ in which $w_x(y) = U_x(\delta_y)$ for all $y \in X_B$. By Axiom 2 (outcome monotonicity), w_x is strictly increasing. To see that w_x is continuous, suppose by contradiction that there exists a sequence (y^n) in X_B such that $y^n \to y \in X_B$ and $w_x(y^n) \not\to w_x(y)$. Without loss of generality and passing to a subsequence if necessary, suppose $w_x(y^n) \to a < b = w_x(y)$ and $w_x(y^n) < (a+b)/2$ for all n. By part (ii), we can find $r \in \Delta(X_B)$ with $U_x(r) = (a+b)/2$. That is, $y^n \prec_x r \prec_x y$ for all n. Axiom 5 (continuity) implies $y \precsim_x r \prec_x y$, a contradiction. Finally, by our construction, once $U_x(\delta_{\overline{x}_B})$ and $U_x(\delta_{\underline{x}_B})$ are determined, the utility function U_x is pinned down. Hence, U_x is unique up to a positive affine transformation. □

IV.3: Proof of Theorem 2

Suppose that $\succsim$ has an HEU representation with a tight $\mathcal{H}$. We start with two observations. The first is that the conditional preference of $\succsim$ on any nonempty subset of dimensions also has an HEU representation.

Lemma 11. *For any nonempty $I' \subseteq I$ and $z \in X_{I'^c}$, the conditional preference $\succsim_z$ has an HEU representation.*

Proof of Lemma 11. Fix any $I' \subseteq I$ and $z \in X_{I'^c}$. We ignore the dependence of functions on z when there is no risk of confusion. Define $\mathcal{H}' = \{A \subseteq I' : A \neq \emptyset, A = I' \cap B \text{ for some } B \in \mathcal{H}\}$. Clearly $\mathcal{H}'$ is a hierarchy on I'. We can similarly define mappings H', η', τ', and Φ'. For any $A \in \mathcal{H}'$ and $z' \in X_{\eta'(A)}$, denote by B_A the smallest component in $\mathcal{H}$ that includes A. Then $\tau'(A) \subseteq \tau(B_A)$ and $\eta'(A) \subseteq \eta(B_A)$. Define $\hat{U}^A_{z'}(p)$ for each $p \in \Delta(X_A)$ by

$$\hat{U}^A_{z'}(p) = U^{B_A}_{(z', z_{\eta(B_A)\cap I'^c})}(p, z_{B_A \setminus A}).$$

Then we recursively define $\hat{u}^A : X_{\eta'(A)} \times X_{\tau'(A)} \times \mathbb{R}^{\Phi'(A)} \to \mathbb{R}$, which satisfies

$$\hat{U}^A_{z'}(p) = \mathbb{E}^p_{\tau'(A)}\, \hat{u}^A(\, z',\ y,\ (\hat{U}^B_{(z',y)}(p_{B|y}))_{B\in\Phi'(A)}\,).$$

We can easily verify that $(\mathcal{H}', (\hat{u}^A)_{A\in\mathcal{H}'})$ is an HEU representation of $\succsim_z$. □

The second observation describes what happens if some hierarchy in $\mathbb{H}$ contains a bracket separable component.

Lemma 12. *Suppose $A \in \mathcal{H}$ for some $\mathcal{H} \in \mathbb{H}$. If A is bracket separable, then there exists $\mathcal{H}' \in \mathbb{H}$ such that $A \in \mathcal{H}'$ and $\tau^{\mathcal{H}'}(A) = \emptyset$.*

Proof of Lemma 12. Suppose $A \in \mathcal{H}$ and $(\mathcal{H}, (u^B)_{B\in\mathcal{H}})$ is an HEU representation of $\succsim$. We assume that $\tau^{\mathcal{H}}(A) \neq \emptyset$, since otherwise we can simply let $\mathcal{H}' = \mathcal{H}$. For any $z \in X_{A^c}$, Lemma 11 implies that $\succsim_z$ has an HEU representation with hierarchy $\mathcal{H}_A = \{B \subseteq A : B \neq \emptyset, B = A \cap A' \text{ for some } A' \in$

$\mathcal{H}\}$. Since A is bracket separable, applying our constructive proof of the sufficiency of Theorem 1 to $\succsim_z$ generates a different HEU representation $(\mathcal{H}'_A, (u'^B_z)_{B\in\mathcal{H}'_A})$ of $\succsim_z$ in which $\tau^{\mathcal{H}'_A}(A) = \emptyset$. Notably, $\mathcal{H}'_A$ does not depend on z and u'^A_z is only unique up to a monotone transformation—the expectation operator in equation (1) for component A is degenerate. Define $\mathcal{H}' = \mathcal{H} \cup \mathcal{H}'_A \setminus \mathcal{H}_A$. It is easy to verify that $\mathcal{H}'$ is a hierarchy—and importantly, is not tight. Let $\hat{u}^B = u^B$ for $B \in \mathcal{H}' \setminus \mathcal{H}'_A$ and extend the domain of $(u'^B_z)_{z\in X_{A^c}}$ to derive $\hat{u}^B$ for $B \in \mathcal{H}'_A$ (and adjust the values of u'^A_z if necessary). Then $(\mathcal{H}', (\hat{u}^B)_{B\in\mathcal{H}'})$ is also an HEU representation of $\succsim$ such that $A \in \mathcal{H}'$ and $\tau^{\mathcal{H}'}(A) = \tau^{\mathcal{H}'_A}(A) = \emptyset$. □

We are now ready to prove the uniqueness of the canonical hierarchy if it exists. Suppose that $\mathcal{H}^1, \mathcal{H}^2 \in \mathbb{H}$ are canonical for $\succsim$. For each $i = 1, 2$, define $\mathcal{H}^i_0 = \{I\}$ and $\mathcal{H}^i_k = \{A \in \mathcal{H}^i : A \in \Phi(B)$ for some $B \in \mathcal{H}^i_{k-1}\}$ for all $k \geqslant 1$. There exists a unique $K_i \geqslant 0$ such that $\mathcal{H}^i_{K_i} \neq \emptyset$ and $\mathcal{H}^i = \bigcup_{k=0}^{K_i} \mathcal{H}^i_k$. Indeed, $\{\mathcal{H}^i_k\}_{k=0}^{K_i}$ is a partition of $\mathcal{H}^i$ for both $i = 1, 2$.

We will show $\mathcal{H}^1_k = \mathcal{H}^2_k$ by induction on $k \geqslant 0$. Clearly, $\mathcal{H}^1_0 = \mathcal{H}^2_0 = \{I\}$. Suppose that there exists $t \geqslant 0$ such that $\mathcal{H}^1_k = \mathcal{H}^2_k$ for all $k \leqslant t$. Fix any $A \in \mathcal{H}^1_t = \mathcal{H}^2_t$. If there exists $\mathcal{H} \in \mathbb{H}$ such that $A \in \mathcal{H}$ and $\tau^{\mathcal{H}}(A) = \emptyset$, then $\tau^{\mathcal{H}^1}(A) = \tau^{\mathcal{H}^2}(A) = \emptyset$ by the first statement in Definition 4. Since $\mathcal{H}^1, \mathcal{H}^2$ are tight, we must have $A = I$ and $t = 0$. As a result, $\mathcal{H}^i_{t+1} = \Phi^{\mathcal{H}^i}(I)$, which is a bracket partition of I for both $i = 1, 2$. Any $B \in \mathcal{H}^i_{t+1}$ must not be bracket separable, since otherwise, by Lemma 12, there exists $\mathcal{H}' \in \mathbb{H}$ such that $B \in \mathcal{H}'$ and $\tau^{\mathcal{H}'}(B) = \emptyset$, which implies $\tau^{\mathcal{H}^i}(A) = \emptyset$ and contradicts $\mathcal{H}^i$ being tight. By Lemma 5, $\mathcal{H}^i_{t+1}$ must be the finest partition of I for both $i = 1, 2$ and hence $\mathcal{H}^1_{t+1} = \mathcal{H}^2_{t+1}$.

If instead $\tau^{\mathcal{H}}(A) \neq \emptyset$ for every $\mathcal{H} \in \mathbb{H}$ such that $A \in \mathcal{H}$, then $\tau^{\mathcal{H}^1}(A) \subseteq \tau^{\mathcal{H}^2}(A)$ and $\tau^{\mathcal{H}^2}(A) \subseteq \tau^{\mathcal{H}^1}(A)$ by the second statement in Definition 4. That is, $\tau^{\mathcal{H}^1}(A) = \tau^{\mathcal{H}^2}(A)$. Denote this by A'. By the same argument as in the previous case, $\Phi^{\mathcal{H}^1} \cup \Phi^{\mathcal{H}^2}$ contains no bracket separable components. Then the following three cases are possible: (i) $A \setminus A' = \emptyset$, in which case $\Phi^{\mathcal{H}^1}(A) = \Phi^{\mathcal{H}^2}(A) = \emptyset$; (ii) $A \setminus A'$ is not bracket separable, in which case $\Phi^{\mathcal{H}^1}(A) = \Phi^{\mathcal{H}^2}(A) = \{A \setminus A'\}$; and (iii) or $A \setminus A'$ is bracket separable, in which case $\Phi^{\mathcal{H}^1}(A) = \Phi^{\mathcal{H}^2}(A)$ is the finest bracket partition of $A \setminus A'$.

Hence, $\Phi^{\mathcal{H}^1} = \Phi^{\mathcal{H}^2}$ holds for all $A \in \mathcal{H}^1_t = \mathcal{H}^2_t$, which implies $\mathcal{H}^1_{t+1} =$

$\mathcal{H}^2_{t+1}$. By induction, $\mathcal{H}^1_t = \mathcal{H}^2_t$ for all $t \geqslant 0$ and hence $\mathcal{H}^1 = \mathcal{H}^2$.

Then we establish the existence of a canonical hierarchy. Denote by $(\mathcal{H}, (u^A)_{A\in\mathcal{H}})$ the HEU representation constructed in the proof of the sufficiency of Theorem 1. Recall that $\mathcal{H} = \bigcup_{t=0}^n \mathcal{H}_t$ in which $\mathcal{H}_t$ is the set of components constructed in each stage $0 \leqslant t \leqslant n$. Notice that this procedure may not generate a unique hierarchy, since we apply an arbitrary order to dimensions in $M(A)$ when $|M(A)| \geqslant 2$ and hence it is not necessarily the canonical hierarchy. However, the next lemma provides insights regarding how to remove components in $\mathcal{H}$ to generate a canonical hierarchy.

Lemma 13. *Suppose $\mathcal{H} \in \mathbb{H}$ and $A \in \mathcal{H}$ such that $\tau(A) \neq \emptyset$. If $A \setminus \tau(A) \in \mathcal{H}$ and $\tau(A \setminus \tau(A)) = \{i\}$ for some $i \in M(A)$, then $\mathcal{H} \setminus \{A \setminus \tau(A)\} \in \mathbb{H}$.*

Proof of Lemma 13. Suppose $A \in \mathcal{H}$ and $(\mathcal{H}, (u^B)_{B\in\mathcal{H}})$ is an HEU representation of $\succsim$, and they satisfy the conditions stated in the lemma. Denote $\mathcal{H}' = \mathcal{H} \setminus \{A \setminus \tau^{\mathcal{H}}(A)\}$. Then $\tau^{\mathcal{H}'}(A) = \tau^{\mathcal{H}}(A) \cup \{i\}$, $\Phi^{\mathcal{H}'}(A) = \Phi^{\mathcal{H}}(A \setminus \tau^{\mathcal{H}}(A))$, and $\eta^{\mathcal{H}'}(A) = \eta^{\mathcal{H}}(A)$. The three functions of $\mathcal{H}$ and $\mathcal{H}'$ agree on $B \in \mathcal{H}' \setminus \{A\}$. Let $\hat{u}^B = u^B$ for all $B \in \mathcal{H}' \setminus \{A\}$ and define $\hat{u}^A : X_{\eta^{\mathcal{H}'}(A)} \times X_{\tau^{\mathcal{H}'}(A)} \times \mathbb{R}^{\Phi^{\mathcal{H}'}(A)} \to \mathbb{R}$ by

$$\hat{u}^A(z, x, a) = u^A\Big(z,\ x_{\tau^{\mathcal{H}}(A)},\ u^{A\setminus\tau^{\mathcal{H}}(A)}\Big(\,(z, x_{\tau^{\mathcal{H}}(A)}),\ x_i,\ a\Big)\Big).$$

Define U^B_z and $\hat{U}^B_z$ accordingly using the recursive equation (1). It is easy to see that $\hat{U}^B_z(\delta_x) = U^B_z(\delta_x)$ for all $B \in \mathcal{H}', z \in X_{\eta^{\mathcal{H}}(A)}$ and $x \in X_B$. This guarantees that $\hat{U}^B_z(\delta_x)$ is continuous and strictly increasing in $x \in X_B$. Moreover, $\hat{U}^B_z(p) = U^B_z(p)$ for all $B \in \mathcal{H}'$ with $B \subsetneq A$, $p \in \Delta(X_B)$ and $z \in X_{\eta^{\mathcal{H}}(B)}$. To show that $(\mathcal{H}', (\hat{u}^B)_{B\in\mathcal{H}'})$ is an HEU representation of $\succsim$, it suffices to show that $\hat{U}^A_z(p) = U^A_z(p)$ for all $p \in \Delta(X_A)$ and $z \in X_{\eta^{\mathcal{H}}(A)}$, since by recursion, it implies that $\hat{U}^B_z(p) = U^B_z(p)$ for all $B \in \mathcal{H}'$ with $A \subsetneq B$, $p \in \Delta(X_B)$ and $z \in X_{\eta^{\mathcal{H}}(B)}$.

Denote $A' = A \setminus \tau^{\mathcal{H}}(A)$. Fix $z' \in X_{A^c}$ and denote $z = z'_{\eta^{\mathcal{H}}(A)} \in X_{\eta^{\mathcal{H}}(A)}$. The utilities of $p \in \Delta(X_A)$ in the two representations are given by

$$\begin{aligned}
U^A_z(p) &= \mathbb{E}^p_{\tau^{\mathcal{H}}(A)}\, u^A\Big(z, y, \mathbb{E}^p_{i|y}\, u^{A'}((z,y), y',\ (U^B_{(z,y,y')}(p_{B|y,y'}))_{B\in\Phi^{\mathcal{H}'}(A)})\Big),\\
\hat{U}^A_z(p) &= \mathbb{E}^p_{\tau^{\mathcal{H}}(A)\cup\{i\}}\, \hat{u}^A\Big(\ z,\ (y,y'),\ (U^B_{(z,y,y')}(p_{B|y,y'}))_{B\in\Phi^{\mathcal{H}'}(A)}\Big)\\
&= \mathbb{E}^p_{\tau^{\mathcal{H}}(A)\cup\{i\}}\, u^A\Big(z, y,\ u^{A'}((z,y), y',\ (U^B_{(z,y,y')}(p_{B|y,y'}))_{B\in\Phi^{\mathcal{H}'}(A)})\Big).
\end{aligned}$$

Note that in the above expressions, $y \in X_{\tau^{\mathcal{H}}(A)}$ and $y' \in X_i$.

Since $i \in M(A)$, Lemma 6 ensures the existence of a function $V_z : \Delta(X_A) \to \mathbb{R}$ such that (i) $p \succsim_z q$ if and only if $V_z(p) \geqslant V_z(q)$ for all $p, q \in \Delta(X_A)$ and (ii) $V_z(p\alpha q) = \alpha V_z(p) + (1-\alpha)V_z(q)$ for all $\alpha \in (0,1)$ and $p, q \in \Delta(X_A)$ with $p_i \perp q_i$.

Fix any $y \in X_{\tau^{\mathcal{H}}(A)}$ and $p \in \Delta(X_{A'})$. Denote $\operatorname{supp}(p_i) = \{y'^1, \ldots, y'^n\}$ such that $y'^1 < y'^2 < \cdots < y'^n$. We can find $y^k \in X_{\tau^{\mathcal{H}}(A)}$ and $y''^k \in X_i$ for every $k \geqslant 1$ such that $(y, y'^k, p_{A'\setminus\{i\}|y'^k}) \sim_z (y^k, y''^k, p_{A'\setminus\{i\}|y'^k})$, elements in $\{y^1, \ldots, y^n\}$ are mutually distinct, and elements in $\{y''^1, \ldots, y''^n\}$ are mutually distinct. Denote $q = \sum_{k=1}^n p_i(y'^k) \cdot (\delta_{y^k}, \delta_{y''^k}, p_{A'\setminus\{i\}|y'^k})$. Properties (i) and (ii) ensure that $V_z(\delta_y, p) = V_z(q)$, which implies $(\delta_y, p) \sim_z q$. Since $\succsim_z$ is also represented by U_z^A, we have

$$\begin{aligned}
&U_z^A(y,p)\\
&= u^A\Big(z, y, \mathbb{E}_i^p\, u^{A'}((z,y), y', (U^B_{(z,y,y')}(p_{B|y'}))_{B\in\Phi^{\mathcal{H}'}(A)})\Big)\\
&= U_z^A(q)\\
&= \sum_{k=1}^n p_i(y'^k)\, u^A\Big(z, y^k, u^{A'}((z,y^k), y''^k, (U^B_{(z,y^k,y''^k)}(p_{B|y'^k}))_{B\in\Phi^{\mathcal{H}'}(A)})\Big)\\
&= \mathbb{E}_i^p\, u^A\Big(z, y, u^{A'}((z,y), y', (U^B_{(z,y,y')}(p_{B|y'}))_{B\in\Phi^{\mathcal{H}'}(A)})\Big).
\end{aligned}$$

The above result holds for all (y,p). Combining the above two sets of equations, for any $r \in \Delta(X_A)$, we have

$$\begin{aligned}
U_z^A(r) &= \mathbb{E}^r_{\tau^{\mathcal{H}}(A)}\, u^A\Big(z, y, \mathbb{E}^r_{i|y}\, u^{A'}((z,y), y', (U^B_{(z,y,y')}(r_{B|y,y'}))_{B\in\Phi^{\mathcal{H}'}(A)})\Big)\\
&= \mathbb{E}^r_{\tau^{\mathcal{H}}(A)}\mathbb{E}^r_{i|y}\, u^A\Big(z, y, u^{A'}((z,y), y', (U^B_{(z,y,y')}(r_{B|y,y'}))_{B\in\Phi^{\mathcal{H}'}(A)})\Big)\\
&= \mathbb{E}^r_{\tau^{\mathcal{H}}(A)\cup\{i\}}\, u^A\Big(z, y, u^{A'}((z,y), y', (U^B_{(z,y,y')}(r_{B|y,y'}))_{B\in\Phi^{\mathcal{H}'}(A)})\Big)\\
&= \hat{U}_z^A(r).
\end{aligned}$$

Hence, $\hat{U}_z^A(r) = U_z^A(r)$ for all $r \in \Delta(X_A)$ and $z \in X_{\eta^{\mathcal{H}}(A)}$. This implies that $(\mathcal{H}', (\hat{u}^B)_{B\in\mathcal{H}'})$ is an HEU representation of $\succsim$ and completes the proof. □

Consider the following process of removing elements from $\mathcal{H}$. For each $t = 0, \ldots, n-1$ and $A \in \mathcal{H}_t$ such that $\Phi(A) \subseteq \mathcal{H}_{t+1}$, we remove any $B \in \Phi(A)$ such that $A \setminus M(A) \subsetneq B$. By construction of $\Phi(A) \in \mathcal{H}_{t+1}$, we can guarantee that each elimination step satisfies the conditions in Lemma

13, and hence the resulting hierarchy $\mathcal{H}'$ belongs to $\mathbb{H}$. Moreover, if such a B exists, then $\tau^{\mathcal{H}'}(A) = M(A)$. Repeat the above process for all $t = 0, \dots, n-1$ and $A \in \mathcal{H}_t$ such that $\Phi(A) \in \mathcal{H}_{t+1}$, and denote the resulting hierarchy by $\mathcal{H}^*$. By repeatedly applying Lemma 13, we know that $\mathcal{H}^* \in \mathbb{H}$ and is tight.

It remains to show that $\mathcal{H}^*$ is a canonical hierarchy for $\succsim$. Note that $\mathcal{H}^*$ is tight. Fix any $A \in \mathcal{H}^*$. For the first statement in Definition 4, if there exists $\mathcal{H} \in \mathbb{H}$ with $A \in \mathcal{H}$ and $\tau^{\mathcal{H}}(A) = \emptyset$, then A is bracket separable. By the construction of $\mathcal{H}^*$, we conclude that $A = I$ and $\tau^{\mathcal{H}^*}(A) = \emptyset$. For the second statement in Definition 4, if $\tau^{\mathcal{H}}(A) \neq \emptyset$ for all $\mathcal{H} \in \mathbb{H}$ such that $A \in \mathcal{H}$, then $\tau^{\mathcal{H}^*}(A) = M(A)$. By the definition of an HEU representation, $i \in \tau^{\mathcal{H}}(A)$ implies $i \rightharpoonup A$ and hence $i \in M(A)$. Thus, $\tau^{\mathcal{H}}(A) \subseteq \tau^{\mathcal{H}^*}(A)$. This completes the proof.

Remark: Indeed, $\mathcal{H}^*$ can be constructed using the following algorithm, which slightly modifies the one in the sufficiently proof of Theorem 1:

Stage 0. We start with $\mathcal{H}_0^* = \{I\}$.

Stage 1. Consider the following cases:

(1) If I is bracket separable, then denote by $\{A_k\}_{k=1}^m$ the finest bracket partition of I as defined in Lemma 5. Note that A_k is not bracket separable for all k. Let $\mathcal{H}_1^* = \{A_k\}_{k=1}^n$ and move to Stage 2.

(2) If I is not bracket separable—that is, $I \neq \bigcup_{B \triangleright I} B$ (Lemma 4)—then the contrapositive of Axiom 4 (separability under bracketing) implies $M(I) = \{i \in I : i \rightharpoonup I\} \neq \emptyset$.

   (i) If $I \setminus M(I) = \emptyset$, then the procedure terminates.

   (ii) If $I \setminus M(I) \neq \emptyset$ is not bracket separable, then let $\mathcal{H}_1^* = \{I \setminus M(I)\}$ and move to Stage 2.

   (iii) If $I \setminus M(I) \neq \emptyset$ is bracket separable, then denote by $\{B_k\}_{k=1}^m$ the finest bracket partition of $I \setminus M(I)$ as defined in Lemma 5. Let $\mathcal{H}_1^* = \{B_1, \dots, B_m\}$ and move to Stage 2.

Stage $t \geqslant 2$. Consider any $A \in \mathcal{H}_{t-1}^*$ such that $|A| \geqslant 2$. By construction, A is not bracket separable. Again by Axiom 4 (separability under bracketing), $M(A) \neq \emptyset$.

(i) If $A \setminus M(A) = \emptyset$, then we make no change.

(ii) If $A \setminus M(A) \neq \emptyset$ is not bracket separable, then let $\{A \setminus M(A)\} \subseteq H_t^*$.

(iii) If $A \setminus M(A) \neq \emptyset$ is bracket separable, then denote by $\{B_k\}_{k=1}^m$ the finest bracket partition of $A \setminus M(A)$ as defined in Lemma 5. Let $\{B_1, \ldots, B_m\} \subseteq H_t^*$.

Repeat the process for all $A \in \mathcal{H}_{t-1}$ such that $|A| \geqslant 2$ and we get all elements in $\mathcal{H}_k$. Move on to Stage $t+1$. This procedure terminates in finitely many stages n. Define $\mathcal{H}^* = \bigcup_{t=0}^n \mathcal{H}_t^*$.

IV.4: Proofs in Section 5

Proof of Theorem 3. First, if $\succsim$ has a FATE representation, then clearly $i \rightharpoonup I$ for all $i \in I$. If $i \rightharpoonup I$ for all $i \in I$, then $M(I) = I$. We can modify the construction procedure in the proof of the sufficiency of Theorem 1 by following only Stage 0 and case (2.i) of Stage 1. The resulting hierarchy can be chosen to be $\mathcal{H} = \{I, I \setminus \{N\}, \ldots, \{1, 2\}, \{1\}\}$ and we can construct Bernoulli indices $(u^A)_{A \in \mathcal{H}}$ such that $(\mathcal{H}, (u^A)_{A \in \mathcal{H}})$ is an HEU representation of $\succsim$ by following the rest of that proof. Then we can apply Lemma 13 repeatedly to show that $\mathcal{H}' = \{I\} \in \mathbb{H}$ and hence $\succsim$ has a FATE representation.

Second, if $\succsim$ has a FETA representation, then by definition, $\{\{i\} : i \in I\}$ is a bracket partition of I, which implies that $\{i\} \rhd I$ for all $i \in I$. Inversely, if $\{i\} \rhd I$ for all $i \in I$, then $\{\{i\} : i \in I\}$ is the finest bracket partition of I, which implies that the canonical hierarchy for $\succsim$ is $\mathcal{H}^* = \{I, \{1\}, \ldots, \{N\}\}$. Hence, $\succsim$ has a FETA representation.

Finally, if $\succsim$ has a recursive representation, then by definition, for any two different A, B in the canonical hierarchy $\mathcal{H}^*$ (which is also the unique hierarchy), we have $A \subsetneq B$ or $B \subsetneq A$. Moreover, since $|\mathcal{H}^*| = N = |I|$, we can find a permutation (i.e., a bijective function) $\pi : I \to I$ such that $\mathcal{H}^* = \{A_k\}_{k=1}^N$ in which $A_k = \{\pi(N), \pi(N-1), \ldots, \pi(k)\}$ for each $k = 1, \ldots, N$. Since $\mathcal{H}^*$ is the canonical hierarchy, $M(A_k) = \{\pi(k)\}$ for all $k = 1, \ldots, N$. This implies that $\pi(k) \rightharpoondown \pi(k+1)$ for all $k = 1, \ldots, N-1$. Now suppose that there exists a bijective function $\pi : I \to I$ such that $\pi(k) \rightharpoondown \pi(k+1)$, $k = 1, \ldots, N-1$. Then in any hierarchy $\mathcal{H} \in \mathbb{H}$, we must have $H(\pi(k+1)) \subsetneq H(\pi(k))$ for all $k = 1, \ldots, N-1$. This implies that $\mathcal{H} = \{H(\pi(k))\}_{k=1}^{N}$, in which $H(\pi(k)) = \{\pi(N), \pi(N-1), \ldots, \pi(k)\}$ for all $k = 1, \ldots, N$. Indeed, this is the unique hierarchy that can support an HEU representation of $\succsim$. Hence, $\succsim$ has a recursive representation. □

Proof of Theorem 4. If $\succsim$ has a generalized bracketing representation with partition $\{A_k\}_{k=1}^{n}$, then $i \rightharpoonup A_k$ for every $i \in A_k$ and $k = 1, \ldots, n$. For any nonempty $A \subseteq I$, if $\bigcup_{B \rhd A} B \neq A$—that is, A is not bracket separable—then $A \subseteq A_k$ for some $k = 1, \ldots, n$, which implies that $i \rightharpoonup A$ for every $i \in A$.

Conversely, assume that for every nonempty $A \subseteq I$, if $\bigcup_{B \rhd A} B \neq A$ (that is, if A is not bracket separable), then $i \rightharpoonup A$ for every $i \in A$ (that is, $M(A) = A$). Denote by $\mathcal{H}^*$ the canonical hierarchy for $\succsim$. By definition, for any $A \in \mathcal{H}^* \setminus \{I\}$, we know that A is not bracket separable, which implies $M(A) = A$. Hence, if I is not bracket separable, then $\mathcal{H}^* = \{I\}$. If instead I is bracket separable, then $I = \mathcal{H}^* = \{I, \{A_k\}_{k=1}^{n}\}$ in which $\{A_k\}_{k=1}^{n}$ is the finest bracket partition of I. In both cases, $\succsim$ has a generalized bracketing representation.

If $\succsim$ has a generalized recursive representation with hierarchy $\mathcal{H}$ such that for all $A, B \in \mathcal{H}$, either $A \subseteq B$ or $B \subseteq A$, then we can write $\mathcal{H} = \{A_k\}_{k=1}^{n}$ in which $A_n = I$ and $A_{k-1} \subsetneq A_k$ for every $k = 2, \ldots, n$. By the HEU representation, $\tau(A_k) = A_k \setminus A_{k-1}$ for every $k = 2, \ldots, n$ and $\tau(A_1) = A_1$. For any $k = 1, \ldots, n$ and $i \in \tau(A_k)$, we have $i \rightharpoonup A_k$. For any nonempty $A \subseteq I$, denote by k^* the largest k such that $A \cap \tau(A_k) \neq \emptyset$. Choose any $i \in A \cap \tau(A_{k^*})$. Then $A \subseteq A_{k^*}$ and $i \rightharpoonup A_{k^*}$, which implies $i \rightharpoonup A$.

Conversely, suppose for every nonempty $A \subseteq I$ there exists $i \in A$ such that $i \rightharpoonup A$—that is, $M(A) \neq \emptyset$. Consider the following induction. Denote $A_1 = I$. For each $1 \leqslant t \leqslant N-1$, choose any $i_t \in M(A_t)$ and denote $A_{t+1} = A_t \setminus \{i_t\}$. Let $\mathcal{H} = \{A_t\}_{t=1}^{N}$. Then $\mathcal{H}$ is a tight hierarchy, $\tau(A_t) = \{i_t\}$ for each $t = 1, \ldots, N-1$ and $\tau(A_N) = A_N$. We can follow the construction of Bernoulli indices in the proof of the sufficiency of Theorem 1 to generate an HEU representation of $\succsim$ with hierarchy $\mathcal{H}$. Hence, $\succsim$ has a generalized recursive representation. □

IV.5: Proofs in Section 6.1

We first prove Proposition 2.

Proof of Proposition 2. For sufficiency, note that if $\succsim$ is represented by $U(p) = \mathbb{E}^{f[p]}u(x)$, then $\succsim$ satisfies dominance. Suppose instead u exhibits CARA—that is, $c(q,u) = \frac{1}{a}\log \mathbb{E}^q\, e^{ax}$ for $a \neq 0$ or $c(q,u) = \mathbb{E}^q x$. In the latter case, u is also represented by $U(p) = \mathbb{E}^{f[p]}u(x)$ and hence dominance holds. In the former case, for any $p \in \Delta(Z^N)$ such that $p = (p_1, \ldots, p_N)$, we have

$$U(p) = \frac{1}{a}\sum_{i=1}^{n} \log \mathbb{E}^{f[p_{A_i}]}\, e^{ax} = \frac{1}{a}\,\log \prod_{i=1}^{n}\, \mathbb{E}^{f[p_{A_i}]}\, e^{ax} = \frac{1}{a}\,\log\, \mathbb{E}^{f[p]}\, e^{ax}.$$

The last equality holds because p_{A_i} and p_{A_j} are statistically independent for all $i \neq j$. Hence, $\succsim$ satisfies dominance without correlation.

For necessity, suppose that $\succsim$ satisfies dominance without correlation and $\succsim$ is not represented by $U(p) = \mathbb{E}^{f[p]}u(x)$. Then $\{A_i\}_{i=1}^n$ must be a nontrivial partition of I. Choose any $i \in A_1$ and $j \in A_2$ and fix $z \in X_{\{i,j\}^c}$ such that $z_l = 0$ for all $l \in \{i,j\}^c$, we consider the conditional preference $\succsim_z$ on $\Delta(X_{i,j})$ and utility function can be written as $\hat{U}(p) = c(p_i, u) + c(p_j, u)$ for all $p \in \Delta(X_{\{i,j\}})$. By Proposition 1 of Rabin and Weizsäcker (2009), if u does not exhibit CARA, then $\succsim_z$ (and hence $\succsim$) must violate dominance without correlation, a contradiction. This completes the proof. □

Proof of Proposition 1. The sufficiency part is trivial. For necessity, suppose that $\succsim$ satisfies dominance and $\succsim$ is not represented by $U(p) = \mathbb{E}^{f[p]}u(x)$. Then Proposition 2 implies that u must exhibit CARA. We can construct $\succsim_z$ on $\Delta(X_{i,j})$ as in the proof of Proposition 2, which is represented by $\hat{U}(p) = c(p_i, u) + c(p_j, u)$. Normalize $u(0) = 0$. By dominance, for any $x \in Z$, we have $\delta_{(x,0)}\frac{1}{2}\delta_{(0,x)} \sim \delta_{(x,0)}$, which implies $u(x/2) = u(x)/2$. Hence, $c(q,u) = \mathbb{E}^q x$ and $\succsim$ is represented by $U(p) = \mathbb{E}^{f[p]}u(x)$, which leads to a contradiction. This completes the proof. □

Proof of Proposition 3. Recall that u exhibits CARA if and only if $A(x)$ is a constant for all $x \in Z$.

For the first statement, the sufficiency part is trivial. Indeed, by Proposition 2, if $\succsim$ is represented by $U(p) = \mathbb{E}^{f[p]}u(x)$ or u exhibits CARA, then $\succsim$ satisfies dominance without correlation and the decision maker is indifferent regardless of whether her income comes from one source or multiple sources. For necessity, suppose that $\succsim$ is not represented by $U(p) = \mathbb{E}^{f[p]}u(x)$. We can construct $\succsim_z$ on $\Delta(X_{i,j})$ as in the proof of Proposition 2, which is represented by $\hat{U}(p) = c(p_i, u) + c(p_j, u)$. Set $p_j = \delta_x$ for some $x \in Z$. Since both gains and losses are allowed, avoidance of multidimensional risk implies that $(p_i, \delta_x) \sim (f[(p_i, \delta_x)], \delta_0)$—that is, $c(p_i, u) + x = c(f[(p_i, \delta_x), u])$. Hence, u must exhibit CARA.

For the second statement, suppose $\succsim$ is not represented by $U(p) = \mathbb{E}^{f[p]}u(x)$. It suffices to show that $\succsim$ satisfies avoidance of multidimensional risk for gains if and only if $A(x)$ is decreasing in $x \in Z_+$. To show the "if" part, suppose $A(x)$ is decreasing in $x \in Z_+$. Then $c(q, u) + c(q', u) \leqslant c(f[(q, q')], u)$ for all $q, q' \in \Delta(Z_+)$. Hence, for any $p \in \Delta(Z_+^N)$ such that $p = (p_1, \ldots, p_N)$, we have

$$U(p) = \sum_{i=1}^{n} c(f[p_{A_i}], u) \leqslant c\Big(f\Big[(f[p_{A_1}], \ldots, f[p_{A_n}])\Big], u\Big) = c(f[p], u),$$

which implies $f[p] \succsim p$. To show the "only if" part, we can follow the same proof idea of the first statement by observing that avoidance of multidimensional risk for gains implies that $c(q, u) + x \leqslant c(f[(q, \delta_x)], u)$ for $x \in Z_+$ and $q \in \Delta(Z_+)$. Hence, $A(x)$ is decreasing in $x \in Z_+$.

The proof of the third statement is symmetric to that of the second one and is omitted. $\square$

IV.6: Proofs in Section 6.2

Proof of Proposition 4. Suppose $\mathcal{H} = \{\{1, 2\}\}$ or $\mathcal{H} = \{\{1, 2\}, \{1\}\}$. Since ϕ is affine, the utility of time lottery (z, p) is $v(z)\mathbb{E}^p[e^{-rt}]$, which implies that the decision maker cannot be risk-averse over time lotteries.

If $\mathcal{H} = \{\{1, 2\}, \{1\}, \{2\}\}$, then using the arguments before the statement of Proposition 4, we know that $\succsim$ is risk-averse over time lotteries if and only if u^2 is convex.

If $\mathcal{H} = \{\{1, 2\}, \{2\}\}$, the utility of (z, p) is $v(z)e^{-rc(p, u^2(z, \cdot))}$. Compared with the case with $\mathcal{H} = \{\{1,2\},\{1\},\{2\}\}$, the only change is that $u^2(z,\cdot)$ may depend on z. Hence, $\succsim$ is risk-averse over time lotteries if and only if $u^2(z,\cdot)$ is convex for all z. □

Proof of Proposition 5. Note that our notion of stochastic impatience is stronger than that of DeJarnette et al. (2020) because we exclude the trivial case in which the decision maker is always indifferent between the two options.

If $\mathcal{H} = \{\{1,2\}\}$ or $\mathcal{H} = \{\{1,2\},\{1\}\}$, then by Propositions 2 and 4 of DeJarnette et al. (2020), $\succsim$ satisfies stochastic impatience if and only if it is risk-seeking over time lotteries and not risk-neutral over time lotteries, which implies that $\succsim$ is not risk-averse over time lotteries.

If $\mathcal{H} = \{\{1,2\},\{1\},\{2\}\}$, then $\succsim$ violates stochastic impatience, because the decision maker is indifferent among any pairing between prizes and payment dates.

If $\mathcal{H} = \{\{1,2\},\{2\}\}$, then by Proposition 4, $\succsim$ is risk-averse over time lotteries if and only if $u^2(z,\cdot)$ is convex for all $z \in Z$. Since the notion of stochastic impatience only involves lotteries without conditional risk in the time dimension given each realization of money, $\succsim$ satisfies stochastic impatience if and only if the expected utility function with Bernoulli index (3) satisfies stochastic impatience, which, by Propositions 2 and 4 of DeJarnette et al. (2020), is equivalent to ϕ being a nontrivial convex transformation of ln. □

References for Online Appendix

CAMARA, M. K. (2021): “Computationally Tractable Choice,” Working Paper.

DEJARNETTE, P., D. DILLENBERGER, D. GOTTLIEB, AND P. ORTOLEVA (2020): “Time Lotteries and Stochastic Impatience,” *Econometrica*, 88, 619–656.

DILLENBERGER, D., D. GOTTLIEB, AND P. ORTOLEVA (2020): “Stochastic Impatience and the separation of Time and Risk Preferences,” Working Paper.

ELLIS, A. AND D. J. FREEMAN (2021): “Revealing Choice Bracketing,” Working Paper.

Epstein, L. G. and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 937–969.

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

Kihlstrom, R. E. and L. J. Mirman (1981): "Constant, Increasing and Decreasing Risk Aversion with Many Commodities," *Review of Economic Studies*, 48, 271–280.

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

Mu, X., L. Pomatto, P. Strack, and O. Tamuz (2023): "Background Risk and Small-Stakes Risk Aversion," Working Paper.

Selden, L. (1978): "A New Representation of Preferences Over "Certain X Uncertain" Consumption Pairs: The "Ordinal Certainty Equivalent" Hypothesis," *Econometrica*, 1045–1060.

Selden, L. and I. Stux (1978): "Consumption Trees, OCE Utility and the Consumption/Savings Decision," Working Paper.

Vorjohann, P. (2023): "Reference-Dependent Choice Bracketing," Working Paper.