

PROCEDURAL MIXTURE SETS

HENDRIK ROMMESWINKEL*

OCTOBER 2026

Abstract

This paper provides two variations of the mixture space theorem of Herstein and Milnor (1953). The first replaces the reduction of compound mixtures property with associativity, giving rise to *procedural mixture sets* in which mixing an element with itself need not return an equivalent element. The second retains standard mixture sets but weakens independence to *disjoint independence*, restricting it to mixtures whose components have disjoint support. In both cases, the standard axioms of weak order, continuity, and (disjoint) independence yield a mixture entropy representation that contains the Shannon or Tsallis entropy of the mixture weights as a component. We apply the first result to event splitting effects in decisions under risk and the second to deliberate randomization.

KEYWORDS: Decision theory, procedural value, mixture sets, entropy, reduction of compound mixtures, associativity

*Graduate School of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-0004, and Waseda Institute for Advanced Study, 1-104 Totsukamachi, Shinjuku City, Tokyo 169-8050; hendrik.r@r.hit-u.ac.jp. Jaffray Lecture at Risk, Uncertainty, and Decision 2024. Many thanks to participants at the Society for the Advancement of Economic Theory conference 2019 and Risk, Uncertainty, and Decision Conference 2024, seminar participants at Erasmus Universiteit Rotterdam and Waseda University and Chun-ting Chen, Soo Hong Chew, Patrick DeJarnette, Takashi Ui, coeditor Mira Frick, and three anonymous referees for helpful comments.

1 INTRODUCTION

The mixture set theorem of Herstein and Milnor (1953) is a foundational result in decision theory. Mixture sets are sets endowed with a commutative mixture operator for which zero-weight mixtures do not matter (a property termed connectedness) and compound mixtures containing identical elements can be reduced by coalescing the identical elements. The theorem states that a preference relation over a mixture set satisfies weak order, continuity, and independence if and only if it admits a utility representation that is linear in mixture weights. The theorem is frequently invoked for preferences over lotteries, Anscombe-Aumann acts, and menus to generate expected-utility-like representations. The present paper provides two variations of the mixture set theorem. In both, the entropy measures of Shannon (1948) and Tsallis (1988) appear as a component of the representation.

The first result (Theorem 1) replaces the reduction of compound mixtures property with associativity of the mixture operation. Reduction of compound mixtures is implied by associativity together with *coalescing*, the requirement that $\mu a \oplus (1 - \mu)a = a$. We call sets satisfying connectedness, commutativity, and associativity—but not necessarily coalescing—*procedural mixture sets*. The failure of coalescing means that the mixing operation can have intrinsic value: mixing an element with itself need not return an equivalent element.

Examples of procedural mixture sets are lotteries presented to decision makers in split form and Markov chains. These have in common that their elements combine naturally under a mixture operation that violates coalescing: Lotteries in split form distinguish between a lottery that yields an outcome for sure and a lottery that displays this outcome twice, with probabilities that sum to one. Running two Markov chains, each with probability $1/2$, yields a new Markov chain with a state space that holds a copy of each original chain's states; a mixture of a chain with itself has therefore twice as many states.

On procedural mixture sets, the standard von Neumann-Morgenstern axioms characterize a *mixture entropy preference* rather than expected utility: there exist parameters $q \in \mathbb{R}$ and $r > 0$ such that

$$U(\mu a \oplus (1 - \mu)b) = \mu^r U(a) + (1 - \mu)^r U(b) + q \cdot H_r(\mu),$$

where H_r is the Shannon entropy when $r = 1$ and the Tsallis entropy when $r \neq 1$.

When $q = 0$ and $r = 1$, the standard expected utility representation is recovered.

We apply the representation result to lotteries in split form. In the spirit of Birnbaum (2004a, 2004b) and Birnbaum and Bahra (2007), we formalize *split form lotteries* as finite multisets of outcome-probability pairs, called *branches*. Branches with identical outcomes are kept separate rather than merged. Mixing two such lotteries concatenates their branches with rescaled probabilities. The literature on event splitting effects (Birnbaum, 2004a, 2004b; Birnbaum & Bahra, 2007; Humphrey, 1995, 1999; Marley & Luce, 2005; Quispe-Torreblanca et al., 2025; Starmer & Sugden, 1993; Weber, 2007) documents that subjects systematically value lotteries differently depending on whether branches with identical outcomes are merged or listed separately: the procedural mixture set framework models this directly without requiring any modification to the preference axioms. Theorem 1 shows that decision makers who satisfy the vNM axioms but do not coalesce identical outcomes do not maximize expected utility but instead a mixture entropy preference.

The second result (Theorem 2) retains the standard mixture set structure, adds associativity, and weakens the independence axiom. The result is a utility representation that, when restricted to mixtures of elements with disjoint supports, takes the form of a mixture entropy. Restricting independence to only hold for such disjoint mixtures is plausible in many contexts in which entropies occur, such as measures of inequality (Theil, 1967), diversity (Nehring & Puppe, 2009), or segregation (Frankel & Volij, 2011) of distributions. We can mix two very unequal distributions and obtain a much more equal one, a failure of full independence. How equal the resulting distribution is depends on where the two distributions concentrate their probability: *ceteris paribus* the mixture is more even when the states with high probability under one distribution have low probability under the other than when the same states have high probability. Under disjoint supports no state receives probability from both mixture components, and this dependence is absent. This motivates restricting the independence axiom to disjoint mixtures.

We apply the second result to deliberate randomization. A decision maker exhibits deliberate randomization when she strictly prefers a mixture of distinct alternatives to either alternative alone. Machina's Mom (Machina, 1989) famously prefers to flip a coin when allocating an indivisible good between her children. Mom is dynamically consistent but violates consequentialism when she rejects

giving a second chance to a child who lost the coin flip. Our disjoint independence condition effectively imposes consequentialism on Mom's decision whenever the randomization involves only children who have not yet had a chance to win the good. By our second result, this condition, together with some technical assumptions, guarantees that Mom has a mixture entropy preference over allocations with disjoint supports. The entropy component expresses Mom's preference for randomizing among her children. The result contributes to the axiomatic literature on deliberate randomization by providing foundations for the deliberate randomization model of Mattsson and Weibull (2002), in which the decision maker maximizes an expectation augmented by a divergence. Unlike most of the literature, our approach directly models preference over the randomization instead of relying on stochastic choice functions.

The remainder of the paper is organized as follows. Section 2 develops procedural mixture sets, states the two representation theorems, and applies them to split form lotteries and to deliberate randomization. Section 3 relates our results to the literature on deliberate randomization, on the reduction of compound mixtures property, and on entropy functions. Section 4 concludes. Proofs of the main theorems and further comparative statics results are collected in the appendices.

2 REPRESENTATION THEOREMS

For ease of comparison, we first recapitulate the mixture sets of Herstein and Milnor (1953).

Definition 1 (Mixture Set). A *mixture set* $\langle \mathcal{M}, \oplus \rangle$ is a set $\mathcal{M}$ endowed with a mixture operator $\oplus$ that maps any $(a, b, \mu) \in \mathcal{M} \times \mathcal{M} \times [0, 1]$ to $\mu a \oplus (1 - \mu)b \in \mathcal{M}$ and that satisfies:

$$\begin{aligned} 1a \oplus (1 - 1)b &= a, \\ \mu a \oplus (1 - \mu)b &= (1 - \mu)b \oplus \mu a, \\ \lambda[\mu a \oplus (1 - \mu)b] \oplus (1 - \lambda)b &= (\lambda\mu)a \oplus (1 - \lambda\mu)b \end{aligned}$$

where each property holds for all $a, b \in \mathcal{M}$ and all $\mu, \lambda \in [0, 1]$.

We may call the three properties, respectively, *connectedness*, *commutativity*, and

reduction of compound mixtures. λ and μ are called the *mixture weights*. An element a is called an *outcome* if there do not exist distinct b and c such that $a = \mu b \oplus (1 - \mu)c$ for some $\mu \in (0, 1)$.

We define the following axioms for a preference relation on the mixture set. These are slightly stronger than those of Herstein and Milnor (1953) but are still necessary conditions of their representation.

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

We further say that a weak order is *nontrivial* if for some a, b , $a \succsim b$ but not $b \succsim a$.

Axiom 2 (Continuity). For any $a, b, c, d \in \mathcal{M}$ and any $\mu \in (0, 1]$, the sets

$$\{\nu \mid \mu[\nu a \oplus (1 - \nu)d] \oplus (1 - \mu)b \succsim c\} \quad \text{and} \quad \{\nu \mid c \succsim \mu[\nu a \oplus (1 - \nu)d] \oplus (1 - \mu)b\}$$

are closed.

At $\mu = 1$ the axiom is the standard mixture continuity employed by Herstein and Milnor (1953). The axiom additionally requires mixture continuity also for nested mixtures.

Axiom 3 (Independence). If $a, a', b \in \mathcal{M}$, $\mu \in (0, 1)$ then $a \succsim a' \Leftrightarrow \mu a \oplus (1 - \mu)b \succsim \mu a' \oplus (1 - \mu)b$.

Herstein and Milnor (1953) proved the following result:

Theorem (Herstein and Milnor, 1953). *Let $\succsim$ be a relation on a mixture set $\mathcal{M}$. The following statements are equivalent:*

1. $\succsim$ satisfies Weak Order, Continuity, and Independence.
2. $\succsim$ can be represented by a function U such that $U(\alpha a \oplus (1 - \alpha)b) = \alpha U(a) + (1 - \alpha)U(b)$ for all $\alpha \in [0, 1]$ and $a, b \in \mathcal{M}$.

The reduction of compound mixtures property is implied by two distinct properties, associativity,

$$\begin{aligned} & \lambda [\mu a \oplus (1 - \mu)b] \oplus (1 - \lambda)c \\ &= \mu \lambda a \oplus (1 - \mu \lambda) \left[\frac{\lambda(1 - \mu)}{1 - \mu \lambda} b \oplus \frac{1 - \lambda}{1 - \mu \lambda} c \right] \end{aligned}$$

for all $a, b, c \in \mathcal{M}$ and all $\mu, \lambda \in [0, 1]$ with $\mu\lambda < 1$, and coalescing:

$$\mu a \oplus (1 - \mu)a = a$$

for all $a \in \mathcal{M}$ and all $\mu \in [0, 1]$. Associativity states that the order of mixing does not matter. It holds in the standard examples of mixture sets: convex subsets of a vector space, lotteries over a finite set of prizes, and menus of lotteries. Coalescing is a structural property of $\oplus$: mixing an alternative with itself returns that alternative; thus, the mixture operation itself cannot affect preference.

By associativity and coalescing, respectively,

$$\begin{aligned} & \lambda[\mu a \oplus (1 - \mu)b] \oplus (1 - \lambda)b \\ &= (\lambda\mu)a \oplus (1 - \lambda\mu) \left[\frac{\lambda(1 - \mu)}{1 - \lambda\mu} b \oplus \frac{1 - \lambda}{1 - \lambda\mu} b \right] \\ &= (\lambda\mu)a \oplus (1 - \lambda\mu)b, \end{aligned}$$

and thus associativity and coalescing imply reduction of compound mixtures. Mongin (2001, p. 67) showed that the reverse implication fails: associativity is not implied by connectedness, commutativity, and reduction of compound mixtures together. Only coalescing is implied by reduction of compound mixtures together with connectedness (Jensen, 1967).

Example. The classical example of a mixture set is the set of probability distributions on a finite set of prizes, $\Delta(\mathcal{X}) = \{p : \mathcal{X} \rightarrow [0, 1] \mid \sum_{x \in \mathcal{X}} p(x) = 1\}$: if a and b are probability distributions, $c = \alpha a \oplus (1 - \alpha)b$ is the probability distribution in which for all $x \in \mathcal{X}$: $c(x) = \alpha a(x) + (1 - \alpha)b(x)$. For this mixture set, a relation that satisfies Axioms 1–3 can be represented by expected utility: $U(a) = \sum_{x \in \mathcal{X}} a(x)u(x)$ for some Bernoulli utility function $u : \mathcal{X} \rightarrow \mathbb{R}$. $\square$

2.1 PROCEDURAL MIXTURES

In our definition of procedural mixture sets we replace reduction of compound mixtures by associativity:

Definition 2 (Procedural Mixture Set). A procedural mixture set $\langle \mathcal{S}, \oplus \rangle$ is a set $\mathcal{S}$ endowed with a mixture operator $\oplus$ that maps any $(a, b, \mu) \in \mathcal{S} \times \mathcal{S} \times [0, 1]$ to

$\mu a \oplus (1 - \mu)b \in \mathcal{S}$ such that for all $a, b, c \in \mathcal{S}$ and all $\mu, \lambda \in [0, 1]$ with $\mu\lambda < 1$,

$$\begin{aligned} 1a \oplus (1 - 1)b &= a, \\ \mu a \oplus (1 - \mu)b &= (1 - \mu)b \oplus \mu a, \\ \lambda [\mu a \oplus (1 - \mu)b] \oplus (1 - \lambda)c \\ &= \mu\lambda a \oplus (1 - \mu\lambda) \left[\frac{\lambda(1 - \mu)}{1 - \mu\lambda} b \oplus \frac{1 - \lambda}{1 - \mu\lambda} c \right]. \end{aligned}$$

Connectedness and commutativity remain unchanged. Coalescing is not assumed.

Example. The literature on *event splitting effects* documents that subjects systematically value lotteries differently depending on whether identical outcomes are coalesced or not. In these experiments, lotteries are presented to subjects as lists of probabilities combined with prizes. A common consequence can be written as one list element or split across several, and the experiments test whether subjects choose differently between the two presentations.

Problem 1D	1–55	56–100	Problem 2D	1–45	46–90	91–100
a	0	11	$\check{a}$	0	11	0
b	7	0	$\check{b}$	7	0	7

Table 1: Two decision problems of Starmer and Sugden (1993). A number between 1 and 100 is drawn uniformly, determining the monetary payoff in pounds.

Concretely, Starmer and Sugden (1993) presented decision makers with the two decision problems displayed in Table 1. Each decision maker chose either between a and b in problem 1D or between $\check{a}$ and $\check{b}$ in problem 2D. Across two groups of 90 subjects, one for each decision problem, 76 chose a in 1D, but only 50 chose $\check{a}$ in 2D, a statistically significant difference. This indicates the presence of event splitting effects that can be modeled by the preferences

$$\begin{aligned} a &\equiv \left\{ \begin{pmatrix} .55 \\ 0 \end{pmatrix}, \begin{pmatrix} .45 \\ 11 \end{pmatrix} \right\} \succ \left\{ \begin{pmatrix} .55 \\ 7 \end{pmatrix}, \begin{pmatrix} .45 \\ 0 \end{pmatrix} \right\} \equiv b, \\ \check{a} &\equiv \left\{ \begin{pmatrix} .45 \\ 0 \end{pmatrix}, \begin{pmatrix} .45 \\ 11 \end{pmatrix}, \begin{pmatrix} .1 \\ 0 \end{pmatrix} \right\} \prec \left\{ \begin{pmatrix} .45 \\ 7 \end{pmatrix}, \begin{pmatrix} .45 \\ 0 \end{pmatrix}, \begin{pmatrix} .1 \\ 7 \end{pmatrix} \right\} \equiv \check{b}, \end{aligned}$$

where the top entries are the probabilities and the bottom entries are the outcomes. The only difference between the two problems is the presentation of the lotteries: in 2D the outcomes 0 and 7 have been split across two probability ranges. If the decision maker is sensitive to such differences in presentation, it makes sense to employ a formal structure that can account for them.

Formally, we define a *split form lottery* over $\mathcal{X}$ as a finite multiset $\left\{ \begin{pmatrix} p_1 \\ x_1 \end{pmatrix}, \dots, \begin{pmatrix} p_n \\ x_n \end{pmatrix} \right\}$ such that $\sum_{i=1}^n p_i = 1$, $p_i > 0$, and $x_i \in \mathcal{X}$ for all $i = 1, \dots, n$; duplicate branches $\begin{pmatrix} p_i \\ x_i \end{pmatrix} = \begin{pmatrix} p_j \\ x_j \end{pmatrix}$ are retained, and the order of elements does not matter. Denote by $\mathcal{S}$ the set of split form lotteries over a set of consequences $\mathcal{X}$.

Let c and d be split form lotteries,

$$c = \left\{ \begin{pmatrix} p_1^c \\ x_1^c \end{pmatrix}, \dots, \begin{pmatrix} p_{n_c}^c \\ x_{n_c}^c \end{pmatrix} \right\}, \quad d = \left\{ \begin{pmatrix} p_1^d \\ x_1^d \end{pmatrix}, \dots, \begin{pmatrix} p_{n_d}^d \\ x_{n_d}^d \end{pmatrix} \right\},$$

and let $\mu \in (0, 1)$. Define

$$\mu c \oplus (1 - \mu)d = \left\{ \begin{pmatrix} \mu p_1^c \\ x_1^c \end{pmatrix}, \dots, \begin{pmatrix} \mu p_{n_c}^c \\ x_{n_c}^c \end{pmatrix}, \begin{pmatrix} (1 - \mu)p_1^d \\ x_1^d \end{pmatrix}, \dots, \begin{pmatrix} (1 - \mu)p_{n_d}^d \\ x_{n_d}^d \end{pmatrix} \right\},$$

and $1c \oplus 0d = c$, $0c \oplus 1d = d$. Mixing two split form lotteries creates a new split form lottery whose branches are the union of the branches of the two, with probabilities scaled by μ and $1 - \mu$ respectively. It is straightforward to see that the set of split form lotteries satisfies commutativity, connectedness, and associativity but not coalescing. It is thus a procedural mixture set, but not a mixture set.

Write $[x] \equiv \left\{ \begin{pmatrix} 1 \\ x \end{pmatrix} \right\}$ for the split form lottery that yields x with certainty. The options of problem 1D are the procedural mixtures $a = .55[0] \oplus .45[11]$ and $b = .55[7] \oplus .45[0]$. The options of problem 2D split the branches with probability .55 into one of probability .45 and one of probability .1,

$$\check{a} = .55 \left[\frac{9}{11}[0] \oplus \frac{2}{11}[0] \right] \oplus .45[11], \quad \check{b} = .55 \left[\frac{9}{11}[7] \oplus \frac{2}{11}[7] \right] \oplus .45[0].$$

Coalescing would identify $\frac{9}{11}[x] \oplus \frac{2}{11}[x]$ with $[x]$ and hence each option of 2D with its counterpart in 1D. Without coalescing, mixing an element with itself can have value, and that value can differ from element to element. The observed choices are those of a decision maker for whom mixing $[7]$ with itself is worth more than

mixing $[0]$ with itself: the decision maker attaches a higher value to splitting the outcome 7 than to splitting the outcome 0. $\square$

We are interested in the following representation of binary relations on procedural mixture sets.

Definition 3 (Mixture Entropy). A binary relation $\succsim$ on $\mathcal{S}$ is a *mixture entropy preference* if there exists a function $U : \mathcal{S} \rightarrow \mathbb{R}$ and parameters $q \in \mathbb{R}, r \in \mathbb{R}_{>0}$ such that U represents $\succsim$ and for all $a, b \in \mathcal{S}$ and $\mu \in [0, 1]$,

$$U(\mu a \oplus (1 - \mu)b) = \mu^r U(a) + (1 - \mu)^r U(b) + q \cdot H_r(\mu)$$

$$H_r(\mu) = \begin{cases} -\mu \ln \mu - (1 - \mu) \ln(1 - \mu), & r = 1 \\ \frac{1 - \mu^r - (1 - \mu)^r}{r - 1}, & r \neq 1 \end{cases} \quad (1)$$

where $0 \ln 0 \equiv 0$ here and thereafter. We call U, r, q a *mixture entropy representation* of $\succsim$.

Importantly, the mixture entropy does not satisfy $U(\mu a \oplus (1 - \mu)a) = U(a)$ unless $q = 0$ and $r = 1$, or $U(a) = q/(r - 1)$ and $r \neq 1$. This will be formalized in Proposition 1 which defines the intrinsic value of mixing that, depending on the parameters, can be negative, zero, or positive.

Theorem 1. Let $\succsim$ be a binary relation on a procedural mixture set $\langle \mathcal{S}, \oplus \rangle$. Then the following two statements are equivalent.

1. $\succsim$ satisfies axioms 1–3.
2. $\succsim$ is a mixture entropy preference.

If U^1, r^1, q^1 and U^2, r^2, q^2 are mixture entropy representations of the same nontrivial weak order, then $r^1 = r^2$ and $U^1 = \phi U^2 + \psi$ for some $\phi \in \mathbb{R}_{>0}$ and $\psi \in \mathbb{R}$. The parameter q satisfies $q^1 = \phi q^2$ when $r^1 = 1$ and $q^1 = \phi q^2 + \psi(r^1 - 1)$ when $r^1 \neq 1$.

When $q = 0$ and $r = 1$, the representation reduces to $U(\mu a \oplus (1 - \mu)b) = \mu U(a) + (1 - \mu)U(b)$, i.e., standard expected utility. In this special case $U(\mu a \oplus (1 - \mu)a) = U(a)$, so the agent is indifferent between $\mu a \oplus (1 - \mu)a$ and a : coalescing holds at the level of preferences and the equivalence classes of $\sim$ form a mixture set.¹

¹More precisely, $\langle \mathcal{S}, \oplus, \sim \rangle$ is a generalized mixture set in the sense of Fishburn (1982).

Example. Returning to the example of event splitting effects, the literature has found that coalescing is not preference-neutral. Birnbaum (2008) and Birnbaum and Navarrete (1998) developed models that account not only for event splitting effects but also for other deviations from expected utility, such as rank dependence. A natural question therefore arises: if we leave the preference axioms of expected utility unchanged, how can behavior differ due to the absence of coalescing? In other words: what is the “pure” effect of the failure to coalesce identical outcomes when individuals are otherwise perfectly rational?

Theorem 1 identifies coalescing of identical outcomes as the only difference between an expected utility maximizer and a maximizer of a mixture entropy representation: if a decision maker “edits” split form lotteries by coalescing identical outcomes before making a choice, then they maximize expected utility. If they do not coalesce outcomes but satisfy all of the expected utility axioms, they maximize a mixture entropy representation. $\square$

Sketch. A rough outline of the proof follows. The details are provided in Appendix A. We first show that $\succsim$ has a quasi-separable representation (Gilboa et al., 2016; Qin & Rommeswinkel, 2024) of the form $U(\mu a \oplus (1 - \mu)b) = f(a, \mu) + g(b, 1 - \mu)$. Commutativity guarantees that $f = g$ without loss of generality. Independence implies that f and g must be increasing functions of the utilities of a and b , respectively. With some abuse of notation, associativity of the procedural mixture set then implies that

$$\begin{aligned} & f(f(U(a), \mu) + f(U(b), 1 - \mu), \lambda) + f(U(c), 1 - \lambda) \\ &= f(U(a), \mu\lambda) + f\left(f\left(U(b), \frac{\lambda(1 - \mu)}{1 - \mu\lambda}\right) + f\left(U(c), \frac{1 - \lambda}{1 - \mu\lambda}\right), 1 - \mu\lambda\right) \end{aligned} \quad (2)$$

Holding c , μ , and λ fixed, the first term of the left-hand side and the entire right-hand side are additively separable representations over the codomains of $U(a)$ and $U(b)$. Two continuous additive representations of the same ordering are related by a common positive affine transformation, so both their sums and their component functions are affine transformations of each other. It follows that f is affine in its first argument, $f(u, \mu) = A(\mu) \cdot u + B(\mu)$. Moreover, $A(\mu) \cdot A(\lambda) = A(\mu\lambda)$, the multiplicative Cauchy equation, with solution $A(\mu) = \mu^r$, and $r > 0$ by

Independence. Substituting our solution for f into (2) yields:

$$\begin{aligned} & \lambda^r [B(\mu) + B(1 - \mu)] + B(\lambda) + B(1 - \lambda) \\ &= B(\mu\lambda) + (1 - \mu\lambda)^r \left[B\left(\frac{\lambda(1-\mu)}{1-\mu\lambda}\right) + B\left(\frac{1-\lambda}{1-\mu\lambda}\right) \right] + B(1 - \mu\lambda) \end{aligned}$$

which after substituting $H(\mu) \equiv B(\mu) + B(1 - \mu)$, $\lambda = 1 - x$, and $\mu\lambda = y$, is the fundamental equation of information,

$$(1 - x)^r H\left(\frac{y}{1-x}\right) + H(x) = (1 - y)^r H\left(\frac{x}{1-y}\right) + H(y),$$

whose solutions (Ebanks et al., 1987) yield the desired representation. $\square$

Remark 1. The multi-mixture representations of Theorem 1 follow from compounding in a straightforward way, e.g.:

$$\begin{aligned} & U\left(\alpha_1 a_1 \oplus (1 - \alpha_1) \left(\frac{\alpha_2}{1 - \alpha_1} a_2 \oplus \frac{1 - \alpha_1 - \alpha_2}{1 - \alpha_1} \left(\cdots \oplus \frac{\alpha_n}{1 - \alpha_1 - \cdots - \alpha_{n-2}} a_n\right)\right)\right) \\ &= \sum_{i=1}^n (\alpha_i)^r U(a_i) + q \cdot H_r(\alpha_1, \dots, \alpha_n) \end{aligned}$$

where $\alpha_i > 0$ and $\sum_{i=1}^n \alpha_i = 1$, and

$$H_r(\alpha_1, \dots, \alpha_n) \equiv \begin{cases} -\sum_i \alpha_i \ln \alpha_i, & r = 1 \\ \frac{1 - \sum_i (\alpha_i)^r}{r-1} & r \neq 1. \end{cases}$$

The case $r = 1$ gives the Shannon (1948) entropy, which in many contexts is the standard measure of entropy. A decision maker whose preference it represents thereby accepts the following property.

Axiom 4 (Mixture Cancellation). For all $a, a', b, b' \in \mathcal{S}$, all $\mu \in (0, 1]$ and all $\lambda \in (0, 1)$,

$$\begin{aligned} & \mu(\lambda a \oplus (1 - \lambda)a) \oplus (1 - \mu)b \sim \mu(\lambda a' \oplus (1 - \lambda)a') \oplus (1 - \mu)b' \\ & \Leftrightarrow \mu a \oplus (1 - \mu)b \sim \mu a' \oplus (1 - \mu)b'. \end{aligned}$$

The axiom expresses that same-element mixtures are not irrelevant but cancel out: replacing a by a' on both branches of an internal mixture has the same

comparative effect as replacing a by a' directly.

It is straightforward to apply mixture cancellation to Theorem 1 to obtain the following corollary.

Corollary 1. *Let $\succsim$ be a binary relation on a procedural mixture set $\langle \mathcal{S}, \oplus, \sim \rangle$. Then the following two statements are equivalent.*

1. $\succsim$ satisfies axioms 1–4.
2. $\succsim$ is a mixture entropy preference with $r = 1$.

We now provide comparative statics. Given that coalescing is not satisfied, the most interesting comparative static concerns the value of mixing an element with itself. The parameter q sets a threshold for $U(a)$ that determines whether mixing an element with itself, $U(\mu a \oplus (1 - \mu)a)$, increases U or not. To make this precise, we introduce a positive and a negative value of mixing.

Definition 4 (Value of Mixing). $\succsim$ exhibits a *negative (positive) value of mixing* at $a \in \mathcal{S}$ if $a \succ (<) \mu a \oplus (1 - \mu)a$ for all $\mu \in (0, 1)$.

The following result is then straightforward:

Proposition 1 (Monotone Mixing). *If $\succsim$ has a mixture entropy representation U, r, q , then the following statements are equivalent:*

1. $\succsim$ exhibits a *negative (positive) value of mixing* at a ,
2. $U(a)(r - 1) > (<) q$.

It follows that repeatedly mixing an element with itself at a fixed weight $\mu \in (0, 1)$, so that $a_{n+1} = \mu a_n \oplus (1 - \mu)a_n$, yields a sequence with $U(a_{n+1}) = (\mu^r + (1 - \mu)^r)U(a_n) + qH_r(\mu)$. If $U(a_0)(r - 1) = q$, the sequence of utilities is constant. Otherwise, if $r > 1$, then U converges to $q/(r - 1)$, and if $r \leq 1$, then U diverges to ∞ when $U(a_0)(r - 1) < q$ and to $-\infty$ when $U(a_0)(r - 1) > q$.

Example. As earlier claimed, the model can capture the behavior $a \succ b$ and $\check{b} \succ \check{a}$ in decision problems 1D and 2D. By Proposition 1 it suffices that the preferences exhibit a positive value of mixing at $[7]$, a negative value of mixing at the worse outcome $[0]$, and a small enough utility difference between a and b . More generally,

we can show that whenever $r < 1$ there exists a monotone utility function over monetary payoffs that generates this behavior:

By (1), mixing $[x]$ with itself at weight μ changes its value by

$$U(\mu[x] \oplus (1 - \mu)[x]) - U([x]) = (\mu^r + (1 - \mu)^r - 1) U([x]) + qH_r(\mu).$$

For $r < 1$ we have $\mu^r + (1 - \mu)^r > \mu + (1 - \mu) = 1$, so the first term is increasing in $U([x])$. The second term does not depend on x . Splitting the outcome 7 is thus worth more than splitting the outcome 0 if and only if $r < 1$. Furthermore, if utilities over consequences satisfy

$$\left(\frac{11}{9}\right)^r < \frac{U([11]) - U([0])}{U([7]) - U([0])} < 1 + \left(\frac{2}{9}\right)^r,$$

then $a \succ b$ and $\check{b} \succ \check{a}$. □

Further comparative statics results are presented in Appendix C. The parameter r is well known in information theory to control the importance of small versus large probabilities. In the context of procedural mixtures, this means that the higher r , the more inclined a decision maker is to prefer a mixture $\alpha d \oplus (1 - \alpha)d$ to a mixture $\beta(\gamma d \oplus (1 - \gamma)d) \oplus (1 - \beta)(\gamma d \oplus (1 - \gamma)d)$ when the value of mixing at d is positive.

In a procedural mixture such as $U(\mu a \oplus (1 - \mu)b)$, the parameter q controls the importance of the value derived from the mixture weight, $H_r(\mu)$, relative to the value of the mixed elements, $U(a)$ and $U(b)$. This means q controls the direction and degree to which the decision maker favors more even distributions. However, the magnitude of q is only behaviorally relevant if there are two outcomes (i.e., elements that are not generated from mixtures themselves) $x \succ y$, otherwise q is preference neutral.

2.2 DISJOINT MIXTURES

Theorem 1 obtained a mixture entropy representation by maintaining the standard mixture set axioms and applying them to preferences on a non-standard mixture set. We now derive an analogous representation for standard mixture sets, relaxing the independence axiom instead.

If decision makers value randomization for its own sake, they violate expected utility: strictly preferring a lottery over two alternatives to either alternative alone is inconsistent with the independence axiom. Such preferences are studied in the literature on deliberate randomization (Border & Segal, 1994; Cerreia-Vioglio et al., 2019; Fudenberg et al., 2015; Grant, 1995; Hashidate, 2018; Machina, 1989; Mattsson & Weibull, 2002; Miao & Zhong, 2018; Trautmann, 2009). Restricting independence to mixtures with *disjoint support* is not only consistent with deliberate randomization but may be normatively compelling, as the following example shows:

Example. Machina’s Mom (Machina, 1989) must allocate a single indivisible good between her children ABIGAIL and BENJIE. Wanting to give both children a fair chance, Mom prefers a coin flip to giving the good outright to one child, let’s say ABIGAIL:

$$\frac{1}{2}\text{ABIGAIL} \oplus \frac{1}{2}\text{BENJIE} \succ \text{ABIGAIL}.$$

Independence (Axiom 3) with $b = \text{BENJIE}$ and $\mu = \frac{1}{2}$ then requires Mom to prefer substituting ABIGAIL with another coin flip within $\frac{1}{2}\text{ABIGAIL} \oplus \frac{1}{2}\text{BENJIE}$. Mom instead has the reverse preference:

$$\frac{1}{2}\text{ABIGAIL} \oplus \frac{1}{2}\text{BENJIE} \succ \frac{1}{2}(\frac{1}{2}\text{ABIGAIL} \oplus \frac{1}{2}\text{BENJIE}) \oplus \frac{1}{2}\text{BENJIE}.$$

The reasoning is that BENJIE already has a fair chance on the first coin flip: the substitution would be equivalent to giving him a second chance in case he loses.

Suppose now Mom has a third child, CHARLIE. Just like for BENJIE, Mom prefers to flip a coin between ABIGAIL and CHARLIE over a direct allocation to ABIGAIL:

$$\frac{1}{2}\text{ABIGAIL} \oplus \frac{1}{2}\text{CHARLIE} \succ \text{ABIGAIL}.$$

Unlike the previous substitution, it would be plausible for Mom to accept a substitution that replaces ABIGAIL with a mixture between her and CHARLIE after

BENJIE has just lost the first coin flip:

$$\frac{1}{2}(\frac{1}{2}\text{ABIGAIL} \oplus \frac{1}{2}\text{CHARLIE}) \oplus \frac{1}{2}\text{BENJIE} \succ \frac{1}{2}\text{ABIGAIL} \oplus \frac{1}{2}\text{BENJIE}.$$

Given Mom's fairness motivation, restricting independence to mixtures that do not share outcomes seems normatively convincing because disjointness of the outcomes restricts substitutions to those in which the affected children have not yet had a fair chance. $\square$

In this subsection we show that, together with the other standard mixture-space axioms, restricting independence to disjoint mixtures yields the same mixture entropy representation as Theorem 1.

To model a notion of distinctness in standard mixture sets, we introduce additional notation. Let $\mathcal{Z}$ be a set. For a given mixture set $\mathcal{M}$, a *support* is a nonempty-valued correspondence $\text{supp} : \mathcal{M} \rightrightarrows \mathcal{Z}$ such that for all $a, b \in \mathcal{M}$ and all $\alpha \in (0, 1)$: $\text{supp}(\alpha a \oplus (1 - \alpha)b) = \text{supp}(a) \cup \text{supp}(b)$. For simple lotteries $p \in \Delta(\mathcal{X})$, taking $\mathcal{Z} = \mathcal{X}$ and $\text{supp}(p) = \{x \in \mathcal{X} \mid p(x) > 0\}$ recovers the usual support definition for lotteries.² A subset Z of $\mathcal{Z}$ is *essential* with respect to a preference relation $\succsim$ on $\mathcal{M}$ if there exist $a, b \in \mathcal{M}$ such that $\text{supp}(a), \text{supp}(b) \subseteq Z$ and $a \succ b$.

Definition 5 (Decomposability). A mixture set $\mathcal{M}$ with support $\text{supp} : \mathcal{M} \rightrightarrows \mathcal{Z}$ is *decomposable* if for every $m \in \mathcal{M}$ and every nonempty proper subset $P \subsetneq \text{supp}(m)$, there exist $m', m'' \in \mathcal{M}$ and $\alpha \in (0, 1)$ such that $\text{supp}(m') = P$, $\text{supp}(m'') = \text{supp}(m) - P$, and $m = \alpha m' \oplus (1 - \alpha)m''$.

Decomposability requires that any element can be expressed as a binary mixture of two elements whose supports form any desired bipartition of the original support. The elements of $\mathcal{Z}$ are thus the indivisible components of the support. Decomposability only imposes structure on the support, not the mixture set, e.g., every mixture set is decomposable if $\mathcal{Z}$ is a singleton set and multiple distinct elements can have an identical singleton support.

²Our definition of a support of a mixture set is broader than this example suggests: for example, the support of a consequence could be a set of attributes that describe this consequence and these attributes could overlap with those of other consequences.

Axiom 5 (Disjoint Independence). For all $a, a', b \in \mathcal{M}$, $\mu \in (0, 1)$, if $(\text{supp}(a) \cup \text{supp}(a')) \cap \text{supp}(b) = \emptyset$, then

$$\begin{aligned} & a \succsim a' \\ \Leftrightarrow & \mu a \oplus (1 - \mu)b \succsim \mu a' \oplus (1 - \mu)b. \end{aligned}$$

If there are sufficiently many essential subsets of $\mathcal{Z}$, a mixture set satisfying disjoint independence admits a representation of the same functional form as Theorem 1:

Theorem 2 (Disjoint Independence Representation Theorem). *Let $\mathcal{M}$ with support supp be a mixture set satisfying decomposability and associativity, and $\succsim^*$ a binary relation on $\mathcal{M}$ such that there exist at least three mutually disjoint essential subsets of $\mathcal{Z}$. Then the following statements are equivalent:*

1. *The relation $\succsim^*$ satisfies Axioms 1, 2, and 5.*
2. *There exist parameters $q \in \mathbb{R}$, $r \in \mathbb{R}_{>0}$, and a function $U : \mathcal{M} \rightarrow \mathbb{R}$, continuous in mixture weights, representing $\succsim^*$ such that for all $a, b \in \mathcal{M}$ with $\text{supp}(a) \cap \text{supp}(b) = \emptyset$:*

$$\begin{aligned} U(\mu a \oplus (1 - \mu)b) &= \mu^r U(a) + (1 - \mu)^r U(b) + q \cdot H_r(\mu) \\ H_r(\mu) &= \begin{cases} -\mu \ln \mu - (1 - \mu) \ln(1 - \mu), & r = 1 \\ \frac{1 - \mu^r - (1 - \mu)^r}{r - 1}, & r \neq 1 \end{cases} \end{aligned}$$

Besides requiring the mixture entropy representation on disjoint mixtures, Theorem 2 states that U is continuous in mixture weights for both disjoint and non-disjoint mixtures.

The most important difference between Theorem 1 and Theorem 2 is the requirement of a sufficient number of essential subsets. An essential subset must contain two elements between which the decision maker is not indifferent; for lotteries over finitely many outcomes and the usual support, three disjoint essential subsets thus require at least six distinct outcomes.

Remark 2. For simple lotteries on a finite prize space $\mathcal{X}$ with the standard support definition and outcomes that can be arranged into three pairs with a nontrivial

preference over their mixtures, Theorem 2 yields

$$U(p) = \sum_{x \in \mathcal{X}} p(x)^r u(x) + q \cdot H_r(p), \quad (3)$$

the prize-by-prize representation familiar from expected utility extended by an entropy term in the mixture weights.

Example. Our essentiality requirement is satisfied if Mom can allocate the good to six different children and strictly prefers coin flips over direct allocations between any two of them. She need not have a strict preference between the direct allocations. The utility representation (3) in the remark applies in this case.

However, to apply Theorem 2, Mom need not have six children. Three children ABIGAIL, BENJIE, CHARLIE combined with two qualities HIGH, and LOW, of the indivisible good suffice, provided the support records only which child receives the good: $\mathcal{Z} = \{z_A, z_B, z_C\}$ with $\text{supp}(\text{ABIGAIL, HIGH}) = \text{supp}(\text{ABIGAIL, LOW}) = \{z_A\}$, and likewise for the other two children. Each of $\{z_A\}$, $\{z_B\}$, $\{z_C\}$ is essential as soon as Mom strictly prefers one quality to the other for that child, e.g., $\text{ABIGAIL, HIGH} \succ \text{ABIGAIL, LOW}$. The mixture entropy representation for disjoint support outcomes is then implied by the axioms with the entropy of the probability of receiving either quality entering the representation. For mixtures of the two qualities for a child, e.g., $U(\mu \text{BENJIE, HIGH} \oplus (1 - \mu) \text{BENJIE, LOW})$, utility can be any arbitrary continuous function of the mixture weight μ . $\square$

We now explain the connection between the two results and why the additional essentiality condition is needed. With only two disjoint outcomes, disjoint independence places no constraint on preferences; all we obtain is a representation $U(\alpha a \oplus (1 - \alpha)b) = f(\alpha, U(a), U(b))$, which need not be quasi-separable given α . With three outcomes a, b, c , disjoint independence and associativity imply that any continuous representation U must satisfy, for all $\alpha, \beta \in [0, 1)$ with $\alpha + \beta \leq 1$

(with some abuse of notation ignoring specifications of supports):

$$\begin{aligned}
& U\left(\alpha a \oplus (1 - \alpha)\left(\frac{\beta}{1 - \alpha} b \oplus \frac{1 - \alpha - \beta}{1 - \alpha} c\right)\right) \\
&= f\left(\alpha, U(a), f\left(\frac{\beta}{1 - \alpha}, U(b), U(c)\right)\right) \\
&= f\left(\beta, U(b), f\left(\frac{\alpha}{1 - \beta}, U(a), U(c)\right)\right) \\
&= U\left(\beta b \oplus (1 - \beta)\left(\frac{\alpha}{1 - \beta} a \oplus \frac{1 - \alpha - \beta}{1 - \beta} c\right)\right).
\end{aligned}$$

To obtain from this the fundamental equation of information that drives the proof of Theorem 1, we require f to be additively separable in the mixture elements. To prove this, the range of f needs to be a continuum. In the proof of Theorem 1 we had infinitely many compounding stages of mixtures available, which allowed us to replace a , b , and c by mixtures covering a rich set of indifference curves; here we are restricted by disjointness.³ Thus, to prove that f is indeed quasi-separable we need three mutually disjoint essential subsets of mixtures in the support. The proof then proceeds largely as in the procedural mixture set case, with additional bookkeeping for the disjoint supports.

Example. Machina (1989) originally describes the example of Mom in a decision-tree setting rather than via preferences over lotteries. The two formulations are connected via the following equivalence: under dynamic consistency, reduction of compound mixtures, and consequentialism, choices over a decision tree reduce to a preference over the lottery induced by the optimal strategy, and this preference satisfies full independence (Hammond, 1988; Volij, 1994). Machina argued for retaining dynamic consistency in non-expected utility preferences. By reducing preferences to a mixture space, we implicitly assumed dynamic consistency and reduction of compound mixtures. Disjoint independence requires Mom to satisfy consequentialism only at decision nodes whose options have in their support only children who would have had no chance of winning the good had that node not been reached.

The resulting representation is intuitive: Mom may have some preferences over which child should receive the good, expressed by the utility of the outcomes. But Mom also cares about randomization: the entropy is increasing in how evenly the

³We may think of procedural mixture sets as effectively duplicating the support with every mixture operation.

probabilities are distributed among the children. □

3 LITERATURE

There are three branches of literature related to the present paper: the literature on deliberate randomization, the literature on mixture sets, and the literature on axiomatic characterizations of entropies.

3.1 DELIBERATE RANDOMIZATION

Border and Segal (1994) shows that as long as Machina's Mom satisfies dynamic consistency, her behavior is approximately expected utility after conditioning on sufficiently small events. Grant (1995) characterizes when Machina's Mom can remain probabilistically sophisticated, working with preferences over Savage acts and dropping monotonicity rather than independence. Trautmann (2009) formalizes Machina's Mom's behavior by augmenting expected utility on monetary payoffs with a process-fairness term defined on the social allocation generated by the randomization. Miao and Zhong (2018) document experimentally that randomization in dictator allocations is widespread and that ex-ante and ex-post fairness preferences coexist in the same population.

Axiomatizations of deliberate randomization commonly use choice-theoretic primitives. Fudenberg et al. (2015) take a stochastic choice rule over outcomes as primitive and recover an additive perturbed utility in which expected utility is supplemented by a nonlinear perturbation function that induces randomization. Cerreia-Vioglio et al. (2019) work with stochastic choice over lotteries and characterize the case in which the chosen distribution maximizes a convex preference over lotteries on each menu. Hashidate (2018) instead takes preferences over menus as primitive and obtains a two-stage representation in which a probability-weighting function captures attitudes toward randomization at the menu-choice stage. Mattsson and Weibull (2002) derive a divergence cost function, but their axioms are imposed on the disutility of effort of choosing a distribution different from a "default choice vector". In the case of $r = 1$, our representation coincides with their objective once the prize utilities are redefined to absorb the default choice vector. Thus, our result can be seen as a characterization of their objective

function with a subjective default choice vector.

3.2 MIXTURE SETS

There is a large literature of decision-theoretic papers that employ the mixture sets introduced by Herstein and Milnor (1953). The mixture set theorem is usually employed as a starting point to obtain a representation linear in mixture weights and additional conditions are then imposed depending on the application. In contrast, Mongin (2001) examines under which conditions mixture sets can be treated as convex subsets of a vector space. Jensen (1967) defined coalescing, his property M5, $\mu a \oplus (1 - \mu)a = a$ and derived it from reduction of compound mixtures and connectedness. Fishburn and Roberts (1978) extended the theorem by replacing $=$ by the indifference relation $\sim$ in the mixture set conditions. Fishburn (1982) formally introduced the thus generalized mixture sets. Payró (2022) extends the mixture space theorem to preferences over menus of lotteries satisfying mixture betweenness, characterizing utility as a solution of $V(x) = \Phi(x, V(x))$ where Φ is mixture linear in the first argument.

3.3 ENTROPY FUNCTIONS

Economic theory frequently uses information measures.⁴ The present paper characterizes the entropy information measures by Shannon (1948) and Tsallis (1988) in a simple algebraic framework. Krantz et al. (1971, ch. 3.12) defined entropy structures algebraically and showed that a relation represented by H_r satisfies the assumptions of an entropy structure, but did not characterize H_r or H_1 . For a general survey of the characterization of information measures, see Csiszár (2008).

Closely related to the present paper are Luce et al. (2008a, 2008b), who introduced the mixture entropy representation to characterize the utility of gambling. As in our setting, their gambles do not coalesce identical outcomes. The key differences are that 1) utility over money is assumed to be linear, 2) certainty

⁴The entropy is used to model inequality (Shorrocks, 1980; Theil, 1967), segregation (Frankel & Volij, 2011), the utility of gambling (Luce et al., 2008a, 2008b), diversity (Nehring & Puppe, 2009), consumer demand (Theil, 1965), freedom of choice (Suppes, 1996), market concentration (Hennessy & Lapan, 2007), and information costs (Caplin et al., 2017; Sims, 2003).

equivalents are assumed to always exist, and 3) their axioms rely on the joint receipt of multiple gambles.

Theorem 2 also yields a comparatively simple axiomatization of the Tsallis (1988) and Rényi (1961) entropies. Rényi’s original characterization works with positive measures rather than probability distributions, and imposes scale invariance. Since a special case of our representation is a monotone transformation of the Rényi entropy, Theorem 2 provides axioms for it that apply to probability distributions.

Regarding the disjoint independence result, a precursor result appears in Chen et al. (2021) in the context of consumption expenditure shares. There a similar result is proven for at least four disjoint subsets instead of three, without relying on quasi-separable preference representations.

4 CONCLUSION

We presented two variations of the mixture set theorem — replacing reduction of compound mixtures by associativity (Theorem 1), or weakening independence to disjoint independence (Theorem 2) — both of which yield the Shannon and Tsallis entropies as components of the utility representation.

The procedural aspect of Theorem 1 is most clearly visible in settings where mixing two structures duplicates their components. An application outside the scope of this paper is the class of Markov chains mentioned in the introduction. These form a procedural mixture set under the natural mixture operation that runs either chain with probability equal to its weight; states must be duplicated for this to be well defined, so mixing a chain with itself violates coalescing. Our mixture entropy representation for $r > 1$ provides a complexity measure for the chain that captures how dispersed the probability distribution over states is as the chain progresses. The parameter r controls whether the long run probabilities or the short run probabilities matter more.

For many applications it may be helpful to generalize procedural mixture sets analogously to Fishburn (1982) by replacing equality in Definition 2 with an equivalence relation $\approx$ on $\mathcal{S}$, imposing $a \approx b \Rightarrow a \sim b$ on preferences. This admits further examples, such as weighted binary trees under the equivalence that rearranges the branching structure while holding the product of weights at each

terminal node fixed.

ACKNOWLEDGMENTS

This work was financially supported by the Center for Research in Econometric Theory and Applications (Grant No. 109-L900-203) from the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan, and by the Ministry of Science and Technology (MOST), Taiwan, under Grant No. 107-2410-H-002-031 and 109-2634-F-002-045.

A PROOF OF THEOREM 1

Proof. Necessity is straightforward. We prove sufficiency.

Let $\mathcal{Q} = \mathcal{S} / \sim$ be the quotient set of $\mathcal{S}$ with respect to the equivalence relation $\sim$. Note that whenever $a, b \in \mathcal{S}$ and $a \sim b$, then any utility representation U must satisfy $U(a) = U(b)$. Note further that $\mathcal{Q}$ is a procedural mixture set when endowed with $\succsim^*$ such that $q \succsim^* r$ if and only if $a \succsim b$ for some $a \in q$ and $b \in r$. Thus, finding a utility on $\mathcal{Q}$ is equivalent to finding a utility on $\mathcal{S}$. We therefore assume for the remainder of the proof that $\mathcal{S} = \mathcal{Q}$.

Let the order topology on $\mathcal{S}$ be the topology generated by the subbase of upper and lower contour sets of the asymmetric part of $\succsim$.

Lemma 1. *A (procedural) mixture set $\mathcal{T}$ endowed with a mixture continuous weak order is topologically connected in the order topology.*

Proof. If $\mathcal{S}$ is not connected, it is the union of two nonempty disjoint open sets $\mathcal{S}'$ and $\mathcal{S}''$. Take any $s' \in \mathcal{S}'$ and $s'' \in \mathcal{S}''$. The set $\mathcal{S}''' = \{a \mid \exists \mu : a = \mu s' \oplus (1 - \mu)s''\}$ lies in $\mathcal{S}$ since $\mathcal{S}$ is a mixture set, and is disconnected in the subspace topology by the disjoint nonempty open sets $\mathcal{S}' \cap \mathcal{S}'''$ and $\mathcal{S}'' \cap \mathcal{S}'''$. Since the upper and lower contour sets of $\succsim$ form a subbase of $\mathcal{S}$, their intersections with $\mathcal{S}'''$ form a subbase of the subspace topology on $\mathcal{S}'''$. By Axiom 2, the bijection $f : \mu \mapsto \mu s' \oplus (1 - \mu)s''$ is continuous. But then $f^{-1}(\mathcal{S}' \cap \mathcal{S}''')$ and $f^{-1}(\mathcal{S}'' \cap \mathcal{S}''')$ are open, disjoint, and disconnect $[0, 1]$, a contradiction. $\square$

Lemma 2. $\succsim$ is coseparable, i.e.,

$$\begin{aligned}\mu a \oplus (1 - \mu)b &\sim \bar{\mu}\bar{a} \oplus (1 - \bar{\mu})\bar{b} \\ \mu a' \oplus (1 - \mu)b &\sim \bar{\mu}\bar{a}' \oplus (1 - \bar{\mu})\bar{b} \\ \mu a \oplus (1 - \mu)b' &\sim \bar{\mu}\bar{a} \oplus (1 - \bar{\mu})\bar{b}'\end{aligned}$$

jointly imply

$$\mu a' \oplus (1 - \mu)b' \sim \bar{\mu}\bar{a}' \oplus (1 - \bar{\mu})\bar{b}'$$

Proof. Using commutativity and associativity it is straightforward to show that

$$\begin{aligned}&1/2[\mu a \oplus (1 - \mu)b] \oplus 1/2[\mu a' \oplus (1 - \mu)b'] \\ &= 1/2[\mu a' \oplus (1 - \mu)b] \oplus 1/2[\mu a \oplus (1 - \mu)b']\end{aligned}$$

for any μ, a, b, a', b' . Axiom 3 together with the assumptions stated above then guarantees the desired result. $\square$

Lemma 3. $\succsim$ can be represented by continuous U, F such that for all $a, b \in \mathcal{S}$ and $\mu \in [0, 1]$:

$$U(\mu a \oplus (1 - \mu)b) = F(a, \mu) + F(b, 1 - \mu)$$

and

$$F(a, 0) = 0 \quad \text{and} \quad F(a, 1) = U(a). \quad (4)$$

Proof. If $a \sim b$ for all $a, b \in \mathcal{S}$, the representation is trivial. Otherwise we use the main theorem of Qin and Rommeswinkel (2024), which provides a representation theorem for weak orders on (open subsets of) $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ with the representation $f(x, z) + g(y, z)$. Here we choose $\mathcal{X} = \mathcal{S}$, $\mathcal{Y} = \mathcal{S}$, and $\mathcal{Z} = (0, 1)$ and endow the space with the product topology of the order topologies and the subspace topology of the reals. Thus, we will first obtain the representation on $\mu \in (0, 1)$ and then extend it to $[0, 1]$ using Axiom 2. To apply the main theorem of Qin and Rommeswinkel (2024), we require the following conditions: essentiality, conditional independence of the $\mathcal{X}$ and $\mathcal{Y}$ dimension given $\mathcal{Z}$, coseparability of $\mathcal{X}$ and $\mathcal{Y}$ given $\mathcal{Z}$, continuity

in the product topology, and topological connectedness of $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$.

Since we have a product space, the well-behavedness assumptions of Qin and Rommeswinkel (2024) are not needed and we also only need essentiality instead of strict essentiality. Essentiality requires that for at least some μ and some a there exist some b, b' such that $\mu a \oplus (1 - \mu)b \not\succ \mu a \oplus (1 - \mu)b'$ and for some a, b there exist some μ, μ' such that $\mu a \oplus (1 - \mu)b \not\succ \mu' a \oplus (1 - \mu')b$. The former is guaranteed by Axiom 3 and the exclusion of the case $a \sim b$ for all $a, b \in \mathcal{S}$. The latter is guaranteed by Axiom 2 and the exclusion of the case $a \sim b$ for all $a, b \in \mathcal{S}$. Next, we need conditional independence of the $\mathcal{X}$ and $\mathcal{Y}$ dimensions given the $\mathcal{Z}$ dimension. This holds by Axiom 3. Further, coseparability of the $\mathcal{X}$ and $\mathcal{Y}$ dimensions given the $\mathcal{Z}$ dimension has been shown above.

For continuity in the product topology on $\mathcal{S} \times \mathcal{S} \times (0, 1)$ we show that the upper sets $\{(a, b, \mu) \mid \mu a \oplus (1 - \mu)b \succ c\}$ are open; the lower sets follow from a symmetric argument. We construct for every mixture $\mu x \oplus (1 - \mu)y \succ c$ an open neighborhood $\mathcal{N}_x \times \mathcal{N}_y \times I$ of (x, y, μ) every element of which is again strictly preferred to c . We take $\mathcal{N}_x = \{x'' \mid x'' \succ x'\}$ for an $x' \prec x$ that still satisfies $\mu x' \oplus (1 - \mu)y \succ c$. x' is constructed as a mixture, $x' = \nu x \oplus (1 - \nu)z$ for some $z \prec x$.⁵ Connectedness of the order topology (Lemma 1) together with Axiom 2 guarantees the existence of a ν for which x' ranks below x while $\mu x' \oplus (1 - \mu)y$ still ranks above c . An analogous construction holds for $\mathcal{N}_y$ by commutativity. Finally for I we take an open set of weights containing μ on which $\lambda x' \oplus (1 - \lambda)y' \succ c$ which holds by Axiom 2 at $\mu = 1$. Axiom 3 closes the argument: replacing x' by any better x'' and y' by any better y'' preserves the strict preference, so every point of the neighborhood is preferred to c .

Topological connectedness of the product topology follows from the connectedness of its components $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$. The interval $(0, 1)$ is connected and each component $\mathcal{S}$ is connected in the order topology.

From Qin and Rommeswinkel (2024) then follows the existence of functions F and E such that $\succsim$ can be represented by

$$U(\mu a \oplus (1 - \mu)b) = F(a, \mu) + E(b, \mu)$$

Commutativity of the mixture set guarantees that we can redefine E and F such

⁵If x is preference minimal there is no such z , and we can simply take $\mathcal{N}_x = \mathcal{S}$.

that $E(b, \mu) = F(b, 1 - \mu)$. To see this, note that from $U(\mu a \oplus (1 - \mu)b) = U((1 - \mu)b \oplus \mu a)$ follows that $F(b, 1 - \mu) + E(a, 1 - \mu) = F(a, \mu) + E(b, \mu)$. Since this holds for all a , $E(b, \mu) = F(b, 1 - \mu) + E(a^*, 1 - \mu) - F(a^*, \mu)$ for some arbitrarily chosen a^* . Thus, $E(b, \mu) = F(b, 1 - \mu) + f(\mu)$ where $f(\mu) \equiv E(a^*, 1 - \mu) - F(a^*, \mu)$. Substituting in the original representation, $U(\mu a \oplus (1 - \mu)b) = F(a, \mu) + F(b, 1 - \mu) + f(\mu)$. Redefining $\hat{F}(a, \mu) = F(a, \mu) + \begin{cases} f(\mu)/2 & \mu < 1/2 \\ f(1 - \mu)/2 & \mu \geq 1/2 \end{cases}$, it follows that $U(\mu a \oplus (1 - \mu)b) = \hat{F}(a, \mu) + \hat{F}(b, 1 - \mu)$.

$\hat{F}(a, \mu)$ is only unique up to addition of a term $\psi(\mu)$ with $\psi(\mu) + \psi(1 - \mu) = 0$. Choose some $a_0 \in \mathcal{S}$ and redefine $\hat{F}$ by addition of

$$\psi(\mu) = \hat{F}(a_0, 1 - \mu) - \mu U(\mu a_0 \oplus (1 - \mu)a_0),$$

which gives $\lim_{\mu \rightarrow 0} \hat{F}(a_0, \mu) = 0$. Since

$$\hat{F}(a, \mu) = U(\mu a \oplus (1 - \mu)b) - U(\mu a_0 \oplus (1 - \mu)b) + \hat{F}(a_0, \mu),$$

we have $\lim_{\mu \rightarrow 0} \hat{F}(a, \mu) = 0$ for every a , so the representation extends to $\mu \in \{0, 1\}$ as desired. $\square$

Lemma 4. $F(a, \mu) = A(\mu)U(a) + B(\mu)$ for all μ and all $a \in \mathcal{S}$.

Proof. For fixed μ , $F(a, \mu)$ is a monotone transformation of U :

$$\begin{aligned} F(a, \mu) &\geq F(b, \mu) \\ \Leftrightarrow F(a, \mu) + F(c, 1 - \mu) &\geq F(b, \mu) + F(c, 1 - \mu) \\ \Leftrightarrow \mu a \oplus (1 - \mu)c &\succsim \mu b \oplus (1 - \mu)c \\ \Leftrightarrow a &\succsim b \\ \Leftrightarrow U(a) &\geq U(b) \end{aligned}$$

Therefore, we can write $F(a, \mu) = G(U(a), \mu)$ with G strictly increasing in its first argument.

From Associativity and the representation in Lemma 3:

$$\begin{aligned}
& U(\mu a \oplus (1 - \mu)[\lambda b \oplus (1 - \lambda)c]) \\
&= G(U(a), \mu) + G(G(U(b), \lambda) + G(U(c), 1 - \lambda), 1 - \mu) \\
&= G(U(b), (1 - \mu)\lambda) \\
&\quad + G\left(G\left(U(a), \frac{\mu}{1 - (1 - \mu)\lambda}\right) + G\left(U(c), \frac{(1 - \mu)(1 - \lambda)}{1 - (1 - \mu)\lambda}\right), 1 - (1 - \mu)\lambda\right) \\
&= U((1 - \mu)\lambda b \oplus (1 - (1 - \mu)\lambda) \left[\frac{\mu}{1 - (1 - \mu)\lambda} a \oplus \frac{(1 - \mu)(1 - \lambda)}{1 - (1 - \mu)\lambda} c \right]). \tag{5}
\end{aligned}$$

Both sides are continuous additive representations over $\mathcal{S} \times \mathcal{S}$ in the elements a and b , so by the uniqueness of additive representations $G(\cdot, 1 - \mu)$ in (5) is affine in its first argument. Since b, c and μ, λ are arbitrary, this holds for all utility levels. Therefore $G(U(a), \mu) = A(\mu)U(a) + B(\mu)$ for all $a \in \mathcal{S}$ and $\mu \in [0, 1]$. (4) gives $A(0) = B(0) = 0$, $A(1) = 1$ and $B(1) = 0$. $\square$

Lemma 5. Suppose $U(\mu a \oplus (1 - \mu)b) = A(\mu)U(a) + A(1 - \mu)U(b) + H(\mu)$ for a continuous, strictly monotone A with $A(1) = 1$, and associativity holds. Then $A(\mu) = \mu^r$, $r \in \mathbb{R}_{>0}$.

Proof. We define $H(\mu) = H(1 - \mu) = B(\mu) + B(1 - \mu)$. Using associativity, we can derive that

$$\begin{aligned}
& A(\lambda) [A(\mu)U(a) + A(1 - \mu)U(b) + H(\mu)] + A(1 - \lambda)U(c) + H(\lambda) \\
&= A(\lambda\mu)U(a) + H(\lambda\mu) \\
&\quad + A(1 - \lambda\mu) \left[A\left(\frac{\lambda(1 - \mu)}{1 - \lambda\mu}\right) U(b) + A\left(\frac{1 - \lambda}{1 - \lambda\mu}\right) U(c) + H\left(\frac{\lambda(1 - \mu)}{1 - \lambda\mu}\right) \right]
\end{aligned}$$

Consider a substitution a' for a under which the above condition must still hold. If $\Delta U = U(a) - U(a')$, then it follows that

$$A(\lambda)A(\mu)\Delta U = A(\lambda\mu)\Delta U$$

and therefore A is multiplicative. Using the multiplicative Cauchy equation it is straightforward to derive that $A(\mu) = \mu^r$, $r \in \mathbb{R}$. By Axiom 3, $r > 0$. $\square$

We now finish the proof. We obtain

$$\lambda^r H(\mu) + H(\lambda) = (1 - \lambda\mu)^r \left[H\left(\frac{\lambda(1 - \mu)}{1 - \lambda\mu}\right) \right] + H(\lambda\mu)$$

and substitute: $\lambda = 1 - x$ and $\lambda\mu = y$. Using $H(x) = H(1 - x)$ we obtain:

$$(1 - x)^r H\left(\frac{y}{1 - x}\right) + H(x) = (1 - y)^r H\left(\frac{x}{1 - y}\right) + H(y)$$

with two types of solutions (Ebanks et al., 1987):

$$\begin{aligned} A(\mu) &= \mu; & H(\mu) &= -(\mu \ln \mu + (1 - \mu) \ln(1 - \mu))\tilde{q} + s \\ A(\mu) &= \mu^r; & H(\mu) &= -(\mu^r + (1 - \mu)^r - 1)\tilde{q} + s \end{aligned}$$

where $\tilde{q}, s \in \mathbb{R}$. Substituting a solution of this form into the functional equation, the terms not involving s cancel and leave $s[(1 - x)^r - (1 - y)^r] = 0$, which for $x \neq y$ forces $s = 0$. Writing $q = \tilde{q}$ when $r = 1$ and $q = (r - 1)\tilde{q}$ when $r \neq 1$, we have therefore obtained the desired representation:

$$\begin{aligned} U(\mu a \oplus (1 - \mu)b) &= \mu^r U(a) + (1 - \mu)^r U(b) + q \cdot H_r(\mu) \\ \text{with } H_r(\mu) &= \begin{cases} -\mu \ln \mu - (1 - \mu) \ln(1 - \mu), & r = 1 \\ \frac{1 - \mu^r - (1 - \mu)^r}{r - 1}, & r \neq 1 \end{cases} \end{aligned}$$

Regarding uniqueness, note that if preferences are nontrivial, then we immediately have an additively separable preference $U(\mu x \oplus (1 - \mu)y)$ over a continuum of x and y and thus U is unique up to affine transformations. The uniqueness properties of r and q follow immediately. $\square$

B PROOF OF PROPOSITION 1

The utility condition for a negative (positive) value of mixing is:

$$U(a) > (<) (\mu^r + (1 - \mu)^r) U(a) + q H_r(\mu)$$

for all $\mu \in (0, 1)$, which is equivalent to

$$U(a)(r-1)H_r(\mu) > (<)qH_r(\mu)$$

Since $H_r(\mu)$ is positive for all $\mu \in (0, 1)$, the desired equivalence is obtained after division by $H_r(\mu)$.

C COMPARATIVE STATICS

Definition 6 (Comparative Value of Mixing). $\succsim_1$ yields a *higher value of mixing* than $\succsim_2$ if for all $\alpha, \beta, \gamma < 1/2$ and some $d \in \mathcal{S}$ at which $\succsim_1$ and $\succsim_2$ exhibit a positive value of mixing,

$$\alpha d \oplus (1 - \alpha)d \succsim_1 \beta(\gamma d \oplus (1 - \gamma)d) \oplus (1 - \beta)(\gamma d \oplus (1 - \gamma)d),$$

then

$$\alpha d \oplus (1 - \alpha)d \succsim_2 \beta(\gamma d \oplus (1 - \gamma)d) \oplus (1 - \beta)(\gamma d \oplus (1 - \gamma)d).$$

In words, if under $\succsim_1$ a binary mixture has a greater value than a mixture across four elements, then this must also be the case under $\succsim_2$.

We now turn to the interpretation of the parameter r . The parameter r controls the degree of the effect of mixing on the value. That is, it controls how much the value increases (or decreases) by an additional mixing stage.

Proposition 2. Let $\succsim_1$ and $\succsim_2$ be mixture entropy preferences with representations U_1 and U_2 and parameters r_1, q_1 and r_2, q_2 , respectively. Then the following statements are equivalent.

1. $\succsim_1$ yields a higher value of mixing than $\succsim_2$.
2. $r_1 \leq r_2$.

Proof. Write $a = \alpha d \oplus (1 - \alpha)d$, $c = \gamma d \oplus (1 - \gamma)d$, and $b = \beta c \oplus (1 - \beta)c$. For

$\mu \in [0, 1]$ and $r > 0$ let

$$T_r(\mu) = \mu^r + (1 - \mu)^r, \quad R_r(\mu) = \begin{cases} \frac{\ln T_r(\mu)}{1 - r}, & r \neq 1, \\ H_1(\mu), & r = 1, \end{cases}$$

the latter being the entropy of Rényi (1961). For $r \neq 1$, (1) gives $H_r = (1 - T_r)/(r - 1)$. Applying (1) to a , to c , and then to b , and collecting terms,

$$U(a) - U(b) = (q - (r - 1)U(d)) \frac{T_r(\beta) T_r(\gamma) - T_r(\alpha)}{r - 1},$$

where at $r = 1$ the second factor equals $H_1(\alpha) - H_1(\beta) - H_1(\gamma)$, its limit. The second factor has the sign of $R_r(\alpha) - R_r(\beta) - R_r(\gamma)$: for $r \neq 1$, $(1 - r)(R_r(\alpha) - R_r(\beta) - R_r(\gamma)) = \ln T_r(\alpha) - \ln(T_r(\beta)T_r(\gamma))$ has the sign of $T_r(\alpha) - T_r(\beta)T_r(\gamma)$. By Proposition 1, the first factor is strictly positive at every d at which $\succsim$ exhibits a positive value of mixing. At such d , therefore,

$$a \succsim b \iff R_r(\alpha) \geq R_r(\beta) + R_r(\gamma). \quad (6)$$

The proposition thus compares the Rényi entropies of two-point distributions across orders. The comparison rests on the following lemma.

Lemma 6. *Let $0 < r < s$. Then R_r and R_s are continuous and strictly increasing on $[0, 1/2]$, with $R_r(0) = R_s(0) = 0$ and $R_r(1/2) = R_s(1/2) = \ln 2$, and*

$$\Theta = R_s \circ R_r^{-1} : [0, \ln 2] \rightarrow [0, \ln 2]$$

is strictly convex.

Proof. Continuity and strict monotonicity of R_r and R_s on $[0, 1/2]$, and with them the endpoint values, are immediate, so Θ is a continuous increasing bijection of $[0, \ln 2]$. We prove convexity.

Throughout, $u > 0$ denotes the order of the entropy and $\varphi(t) = \ln(1 + e^t)$. For $\alpha \in (0, 1/2]$ let $\ell = \ln((1 - \alpha)/\alpha) \in [0, \infty)$ be the log odds, so that $\alpha = 1/(1 + e^\ell)$

and $T_u(\alpha) = (1 + e^{u\ell}) / (1 + e^\ell)^u$. Taking logarithms and differentiating,

$$R_u(\alpha) = \frac{\varphi(u\ell) - u\varphi(\ell)}{1 - u}, \quad -\frac{dR_u}{d\ell} = \frac{\alpha u}{1 + e^{-u\ell}} \cdot \frac{e^{(1-u)\ell} - 1}{1 - u}, \quad (7)$$

where for $u = 1$ both expressions are the limits as $u \rightarrow 1$; in particular the last quotient then equals ℓ . Both factors on the right are strictly positive for $\ell > 0$.

We first show that convexity of Θ depends on the comparative slopes of R_s and R_r : by the chain rule,

$$\Theta'(R_r(\alpha)) = \frac{dR_s/d\ell}{dR_r/d\ell}.$$

As $R_r(\alpha)$ increases over $(0, \ln 2)$, ℓ decreases over $(0, \infty)$, so Θ is strictly convex if the ratio on the right is strictly decreasing in ℓ . Taking logarithms, it suffices that for every fixed $\ell > 0$,

$$\frac{\partial}{\partial \ell} \ln\left(-\frac{dR_u}{d\ell}\right) \text{ is strictly decreasing in } u. \quad (8)$$

It remains to verify (8). Differentiating the logarithm of the right-hand side of (7) in ℓ ,

$$\frac{\partial}{\partial \ell} \ln\left(-\frac{dR_u}{d\ell}\right) = \frac{d \ln \alpha}{d\ell} + \frac{1}{\ell} \left[g((u-1)\ell) - g(u\ell) + \frac{u\ell}{\sinh(u\ell)} \right],$$

where $g(w) = w/(e^w - 1)$ with $g(0) = 1$, using $x/(e^x + 1) = x/\sinh x - g(x)$. The first term does not depend on u . In the bracket substitute $x = u\ell$, which is strictly increasing in u , so that it reads $[g(x - \ell) - g(x)] + x/\sinh x$. We now show that both terms inside the bracket are strictly decreasing in $x > 0$ and thus u . The function g is strictly convex on $\mathbb{R}$: $g(w) = (w/2) \coth(w/2) - w/2$, and $z \coth z$ has second derivative $2(z \coth z - 1)/\sinh^2 z$, which is strictly positive because $|\tanh z| < |z|$ for $z \neq 0$ and tends to $2/3$ as $z \rightarrow 0$. Hence $g'(x - \ell) < g'(x)$ for fixed $\ell > 0$, so $g(x - \ell) - g(x)$ is strictly decreasing in x . Finally, the derivative of $x/\sinh x$ is $\cosh x (\tanh x - x)/\sinh^2 x < 0$ for $x > 0$. So the bracket is strictly decreasing in u , which proves (8). $\square$

We now prove the proposition. Fix d at which both preferences exhibit a

positive value of mixing. By (6), statement 1 says that for all $\alpha, \beta, \gamma \in [0, 1/2]$,

$$R_{r_1}(\alpha) \geq R_{r_1}(\beta) + R_{r_1}(\gamma) \implies R_{r_2}(\alpha) \geq R_{r_2}(\beta) + R_{r_2}(\gamma). \quad (9)$$

If $r_1 = r_2$ this is trivially true, so let $r_1 \neq r_2$ and let $\Theta = R_{r_2} \circ R_{r_1}^{-1}$, which by Lemma 6, applied with $\{r, s\} = \{r_1, r_2\}$, is a strictly increasing bijection of $[0, \ln 2]$ with $\Theta(0) = 0$. Substituting $x = R_{r_1}(\beta)$, $y = R_{r_1}(\gamma)$, and $z = R_{r_1}(\alpha)$, which range over $[0, \ln 2]$, (9) reads: $z \geq x + y$ implies $\Theta(z) \geq \Theta(x) + \Theta(y)$. Because Θ is increasing, this holds for all such x, y, z if and only if

$$\Theta(x + y) \geq \Theta(x) + \Theta(y) \quad \text{for all } x, y \geq 0 \text{ with } x + y < \ln 2. \quad (10)$$

Indeed, given (10) and $z \geq x + y$, $\Theta(z) \geq \Theta(x + y) \geq \Theta(x) + \Theta(y)$; conversely, take $z = x + y$.

If $x = 0$ or $y = 0$, (10) holds with equality, so fix $x, y > 0$ with $x + y < \ln 2$. Writing x as the convex combination of $x + y$ and 0 with weights $x/(x + y)$ and $y/(x + y)$, y as the one with weights $y/(x + y)$ and $x/(x + y)$, and using $\Theta(0) = 0$, strict convexity of Θ gives $\Theta(x) < x\Theta(x + y)/(x + y)$ and $\Theta(y) < y\Theta(x + y)/(x + y)$, and adding the two yields $\Theta(x) + \Theta(y) < \Theta(x + y)$; strict concavity of Θ gives the same inequalities reversed. If $r_1 < r_2$, then Θ is strictly convex by Lemma 6, so (10) holds and statement 1 follows. If $r_1 > r_2$, then Θ is the inverse of $R_{r_1} \circ R_{r_2}^{-1}$, which is strictly increasing and strictly convex by the lemma, so Θ is strictly concave and (10) fails for every $x, y > 0$; hence statement 1 fails. $\square$

In the main text we mentioned that the magnitude of the parameter q only matters for preference if there are at least two outcomes with a strict preference. We make this precise in the following remark:

Remark 3. If $\mathcal{S}$ is generated from finite procedural mixtures of members of a set $\mathcal{X}$ and for all $x, y \in \mathcal{X}$, $U^1(x) = U^1(y)$, and $U^2(x) = U^2(y)$ then U^1 and U^2 represent the same relation if and only if the signs of $U^1(x)(r^1 - 1) - q^1$ and $U^2(x)(r^2 - 1) - q^2$ are identical and $r^1 = r^2$.

It follows from this remark that the magnitude of the parameter q only matters in comparison to a cardinal value of outcomes. If there do not exist outcomes $x \succ y$, then by the uniqueness properties of the representation we can find an

affine transformation of U such that the valuation of any existing outcomes is equal to zero. Any subsequent multiplication of q by a positive factor results in an increasing linear transformation of U (which does not change the represented relation).

We now provide additional results that require the existence of nontrivial preference on a set of *consequences*, i.e., elements of $\mathcal{S}$ that cannot be written in the form $\mu a \oplus (1 - \mu)b$ with $\mu \in (0, 1)$. Let $\mathcal{X}$ be a set of consequences. The following definitions are the standard definitions of certainty equivalents and comparative risk aversion for mixture sets adjusted to the procedural mixture setting.

Definition 7 (Certainty Equivalent). The certainty equivalent $c = ce(\alpha x \oplus (1 - \alpha)y) \in \mathcal{X}$ of a procedural mixture of outcomes x and y is an outcome that satisfies $\alpha x \oplus (1 - \alpha)y \sim \alpha c \oplus (1 - \alpha)c$.

Definition 8 (Comparative Risk Aversion). $\succsim_1$ is at least as risk averse as $\succsim_2$ if for all $\alpha \in (0, 1)$ and all $x, y, z \in \mathcal{X}$, we have that

$$\begin{aligned} & \alpha x \oplus (1 - \alpha)y \succsim_1 \alpha z \oplus (1 - \alpha)z \\ \Rightarrow & \alpha x \oplus (1 - \alpha)y \succsim_2 \alpha z \oplus (1 - \alpha)z \end{aligned}$$

With these definitions, the standard comparative risk aversion result for decisions under risk extends to the procedural case:

Proposition 3. Let $\succsim_1$ and $\succsim_2$ be mixture entropy preferences with representations U_1 and U_2 and parameters r_1, q_1 and r_2, q_2 , respectively. Let $U_1(\mathcal{X})$ and $U_2(\mathcal{X})$ be convex sets. Suppose $\succsim_1$ and $\succsim_2$ coincide on $\mathcal{X}$ and there exist some $x, y \in \mathcal{X}$ such that $x \succ_1 y$. Then the following statements are equivalent.

1. $\succsim_1$ is at least as risk averse as $\succsim_2$.
2. The restriction of U_1 to $\mathcal{X}$ is a concave monotone transformation of U_2 and $r_1 = r_2$.

Proof. $\Leftarrow$ is immediate; we prove $\Rightarrow$. Since the preferences coincide on outcomes, utilities over outcomes must be continuous monotone transformations of one another, i.e., $U_1 = T \circ U_2$ when restricted to $\mathcal{X}$.

Since the values of outcomes are a convex set, for every $x, y \in \mathcal{X}$ such that $x \succ_2 y$ we can find z such that $z = ce_1(1/2x \oplus 1/2y)$. Notice that by the definition

of a certainty equivalent we have $U_1(x)/2 + U_1(y)/2 = U_1(z)$. If T is not concave, then for some such x and y , $U_2(x) + U_2(y) = T^{-1}(U_1(x))/2 + T^{-1}(U_1(y))/2 < T^{-1}(U_1(z)) = U_2(z)$. But then $1/2x \oplus 1/2y \succsim_1 1/2z \oplus 1/2z$ but $1/2x \oplus 1/2y \prec_2 1/2z \oplus 1/2z$, contradicting that $\succsim_1$ is at least as risk averse as $\succsim_2$. We have thus established that T is concave.

Notice now that if $r_1 \neq r_2$, then $p_1(\alpha) \equiv \alpha^{r_1}/(\alpha^{r_1} + (1-\alpha)^{r_1}) \neq \alpha^{r_2}/(\alpha^{r_2} + (1-\alpha)^{r_2}) \equiv p_2(\alpha)$. Without loss of generality, assume that $p_1(\alpha) \geq p_2(\alpha)$ for $\alpha \geq 1/2$. Since T is monotone and continuous, it is differentiable almost everywhere. Without loss of generality, assume T is differentiable at $U_2(x)$ and $U_2(x) = U_1(x)$ and $\partial T(U_2(x))/\partial U = 1$. Then we can find outcomes $x' \succ x \succ x''$ such that $p_1(\alpha)T(U_2(x')) + (1-p_1(\alpha))T(U_2(x'')) > U_2(x) > p_2(\alpha)U_2(x') + (1-p_2(\alpha))U_2(x'')$ contradicting that $\succsim_1$ is at least as risk averse as $\succsim_2$. $\square$

Mixture entropy preferences that are equally risk averse can be compared by the weight they place on the value of consequences relative to the value of mixing.

Definition 9 (Comparative Consequentialism). $\succsim_2$ is at least as consequentialist as $\succsim_1$ if for all $\alpha \in (0, 1)$ and all $x, y \in \mathcal{X}$, we have that

$$\begin{aligned} \alpha x \oplus (1-\alpha)x &\succsim_2 y \\ \Rightarrow \alpha x \oplus (1-\alpha)x &\succsim_1 y \end{aligned} \tag{11}$$

Proposition 4. Let $\succsim_1$ and $\succsim_2$ be equally risk averse mixture entropy preferences with representations U_1 and U_2 and parameters r_1, q_1 and r_2, q_2 , respectively. Let $U_1(\mathcal{X})$ and $U_2(\mathcal{X})$ be convex sets. Suppose there exist some $x, y \in \mathcal{X}$ such that $x \succ_1 y$. Then the following statements are equivalent.

1. $\succsim_2$ is at least as consequentialist as $\succsim_1$.
2. There exist $s \in \mathbb{R}_{>0}$ and $t \in \mathbb{R}$ such that $U_2(x) = sU_1(x) + t$ for all $x \in \mathcal{X}$ and $sq_1 + t(r_1 - 1) \geq q_2$.

Proof. $\Leftarrow$ is immediate; we prove $\Rightarrow$. If $\succsim_1$ and $\succsim_2$ are equally risk averse, then $r_1 = r_2 = r$ and U_1 is an affine transformation of U_2 on outcomes. Assume for the moment that $U_1 = U_2$ on outcomes. Notice that since $H_r(\alpha) \geq 0$ for all r and all α , (11) holds if and only if $q_1 \geq q_2$. In general $U_2 = sU_1 + t$ on outcomes, and by

the uniqueness properties of U the representation of $\succsim_1$ written on the scale of U_2 carries the parameter $sq_1 + t(r - 1)$. Comparing that parameter with q_2 gives the desired result. $\square$

This provides a link between the parameters of our representation and the intensity of the value of mixing relative to the value of consequences. Notice that this comparison is only meaningful when $r_1 = r_2$ and the values of consequences are cardinally comparable.

D PROOF OF THEOREM 2

Proof. We follow the same proof structure as the proof of Theorem 1. Wherever the exact same lemma applies, we will avoid restating it.

Lemma 1 directly applies because it holds both for procedural mixture sets and mixture sets.

The proof of coseparability differs for disjoint mixtures:

Lemma 7. *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be mutually disjoint, essential components of the support. $\succsim^*$ is coseparable on disjoint components: for any $\mu, \bar{\mu} \in (0, 1)$, elements $a, a', \bar{a}, \bar{a}'$ with support in $\mathcal{A}$ and $b, b', \bar{b}, \bar{b}'$ with support in $\mathcal{B}$, the conditions*

$$\begin{aligned}\mu a \oplus (1 - \mu)b &\sim \bar{\mu} \bar{a} \oplus (1 - \bar{\mu})\bar{b} \\ \mu a' \oplus (1 - \mu)b &\sim \bar{\mu} \bar{a}' \oplus (1 - \bar{\mu})\bar{b} \\ \mu a \oplus (1 - \mu)b' &\sim \bar{\mu} \bar{a} \oplus (1 - \bar{\mu})\bar{b}'\end{aligned}\tag{12}$$

jointly imply $\mu a' \oplus (1 - \mu)b' \sim \bar{\mu} \bar{a}' \oplus (1 - \bar{\mu})\bar{b}'$.

Proof. Let $\mathcal{C}$ be a third essential component disjoint from $\mathcal{A}$ and $\mathcal{B}$, and fix c with $\text{supp}(c) \subseteq \mathcal{C}$. For sufficiently small η , by continuity of $\succsim$ and essentiality of $\mathcal{C}$ we can find c' with $\text{supp}(c') \subseteq \mathcal{C}$ such that

$$\eta[\mu a \oplus (1 - \mu)b'] \oplus (1 - \eta)c \sim \eta[\mu a \oplus (1 - \mu)b] \oplus (1 - \eta)c'.\tag{13}$$

By (12) and disjoint independence (same c' on both sides, disjoint from $\mathcal{A} \cup \mathcal{B}$):

$$\eta[\mu a' \oplus (1 - \mu)b] \oplus (1 - \eta)c' \sim \eta[\bar{\mu} \bar{a}' \oplus (1 - \bar{\mu})\bar{b}] \oplus (1 - \eta)c'.\tag{14}$$

Define

$$\zeta := \frac{\eta(1-\mu)}{1-\eta\mu}, \quad \theta := \frac{\eta(1-\bar{\mu})}{1-\eta\bar{\mu}}.$$

By associativity applied to both sides of (14):

$$(\mu\eta)a' \oplus (1-\mu\eta)[\zeta b \oplus (1-\zeta)c'] \sim (\bar{\mu}\eta)\bar{a}' \oplus (1-\bar{\mu}\eta)[\theta\bar{b} \oplus (1-\theta)c']. \quad (15)$$

By (13) and associativity:

$$(\mu\eta)a \oplus (1-\mu\eta)[\zeta b' \oplus (1-\zeta)c] \sim (\mu\eta)a \oplus (1-\mu\eta)[\zeta b \oplus (1-\zeta)c'].$$

By disjoint independence this latter equation is equivalent to:

$$(\mu\eta)a' \oplus (1-\mu\eta)[\zeta b' \oplus (1-\zeta)c] \sim (\mu\eta)a' \oplus (1-\mu\eta)[\zeta b \oplus (1-\zeta)c']. \quad (16)$$

Analogously:

$$(\bar{\mu}\eta)\bar{a}' \oplus (1-\bar{\mu}\eta)[\theta\bar{b}' \oplus (1-\theta)c] \sim (\bar{\mu}\eta)\bar{a}' \oplus (1-\bar{\mu}\eta)[\theta\bar{b} \oplus (1-\theta)c']. \quad (17)$$

Combining (16), (15), and (17):

$$(\mu\eta)a' \oplus (1-\mu\eta)[\zeta b' \oplus (1-\zeta)c] \sim (\bar{\mu}\eta)\bar{a}' \oplus (1-\bar{\mu}\eta)[\theta\bar{b}' \oplus (1-\theta)c].$$

Reversing the associativity operation on both sides:

$$\eta[\mu a' \oplus (1-\mu)b'] \oplus (1-\eta)c \sim \eta[\bar{\mu}\bar{a}' \oplus (1-\bar{\mu})\bar{b}'] \oplus (1-\eta)c.$$

By disjoint independence, $\mu a' \oplus (1-\mu)b' \sim \bar{\mu}\bar{a}' \oplus (1-\bar{\mu})\bar{b}'$ as desired. $\square$

Lemma 8. *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be mutually disjoint, essential components of the support. Then $\succsim^*$ can be represented by continuous U, F such that for all $\mu \in [0, 1]$, all a with $\text{supp}(a) \subseteq \mathcal{A}$ and all b with $\text{supp}(b) \subseteq \mathcal{B}$,*

$$U(\mu a \oplus (1-\mu)b) = F(a, \mu) + F(b, 1-\mu),$$

where F is continuous in the mixture weight on $[0, 1]$ and, for every x with support in $\mathcal{A}$

or in $\mathcal{B}$,

$$F(x, 0) = 0 \quad \text{and} \quad F(x, 1) = U(x). \quad (18)$$

Proof. For the representation of $\mu \in (0, 1)$ the proof is identical to that of Lemma 3 with one change: $\mathcal{X}$ is the subset of the mixture set restricted to a support within $\mathcal{A}$ and $\mathcal{Y}$ is the subset of the mixture set restricted to a support within $\mathcal{B}$. All mixture properties and associativity are inherited by these subsets.

$F(a, \mu)$ is determined only up to addition of a term $\psi(\mu)$ for any function ψ satisfying $\psi(\mu) + \psi(1 - \mu) = 0$. Choose some a_0 with $\text{supp}(a_0) \subseteq \mathcal{A}$ and redefine F by addition of:

$$\psi(\mu) = F(a_0, 1 - \mu) - \mu U(\mu a_0 \oplus (1 - \mu)a_0)$$

which guarantees that $\lim_{\mu \rightarrow 0} F(a_0, \mu) = 0$. Since

$$F(a, \mu) = U(\mu a \oplus (1 - \mu)b) - U(\mu a_0 \oplus (1 - \mu)b) + F(a_0, \mu),$$

$\lim_{\mu \rightarrow 0} F(a, \mu) = 0$ and $\lim_{\mu \rightarrow 1} F(a, \mu) = U(a)$ and we can thus extend the additive representation to $\mu \in \{0, 1\}$. $\square$

Lemma 9. *Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ be mutually disjoint, essential components of the support. There are functions A and B with $F(a, \mu) = A(\mu)U(a) + B(\mu)$ for all μ and all a with support in a single essential component.*

Proof. Let a, b, c have support in $\mathcal{A}, \mathcal{B}, \mathcal{C}$, three mutually disjoint essential components. Disjoint independence implies separability of preferences over a, b, c when holding fixed the mixture weights μ, λ . Together with essentiality of each component and connectedness, the additive representation theorem of Wakker (1989) yields continuous functions $\bar{f}(\cdot, \mu, \lambda)$, $\bar{g}(\cdot, \mu, \lambda)$, $\bar{h}(\cdot, \mu, \lambda)$ and $\bar{k}(\cdot, \mu, \lambda)$, the last strictly increasing in its first argument, such that

$$U(\lambda[\mu a \oplus (1 - \mu)b] \oplus (1 - \lambda)c) = \bar{k}(\bar{f}(a, \mu, \lambda) + \bar{g}(b, \mu, \lambda) + \bar{h}(c, \mu, \lambda), \mu, \lambda) \quad (19)$$

The rest of the argument is analogous to that of Lemma 4, with each ingredient replaced by its disjoint counterpart: disjoint independence takes the role of

Axiom 3, Lemma 8 takes the role of Lemma 3, and (18) takes the role of (4). $\square$

We now finish the proof. Define $H(\mu) = B(\mu) + B(1 - \mu)$, so that $H(\mu) = H(1 - \mu)$. Substituting the affine form of Lemma 9 into the quasi-separable representation, Lemma 5 gives $A(\mu) = \mu^r$. Using associativity as in the procedural mixture case then yields the functional equation

$$\lambda^r H(\mu) + H(\lambda) = (1 - \lambda\mu)^r H\left(\frac{\lambda(1 - \mu)}{1 - \lambda\mu}\right) + H(\lambda\mu).$$

Substituting $\lambda = 1 - x$ and $\lambda\mu = y$ and using $H(x) = H(1 - x)$:

$$(1 - x)^r H\left(\frac{y}{1 - x}\right) + H(x) = (1 - y)^r H\left(\frac{x}{1 - y}\right) + H(y)$$

with two types of solutions (Ebanks et al., 1987):

$$\begin{aligned} A(\mu) &= \mu; & H(\mu) &= -(\mu \ln \mu + (1 - \mu) \ln(1 - \mu))\tilde{q} + s \\ A(\mu) &= \mu^r; & H(\mu) &= -(\mu^r + (1 - \mu)^r - 1)\tilde{q} + s \end{aligned}$$

where $\tilde{q}, s \in \mathbb{R}$. The functional equation is solved on the interior of the simplex, where substituting a solution of this form leaves $s[(1 - x)^r - (1 - y)^r] = 0$ once the terms not involving s cancel, and hence forces $s = 0$; Axiom 2 and connectedness of the mixture set then extend $s = 0$ to the boundary. Writing $q = \tilde{q}$ when $r = 1$ and $q = (r - 1)\tilde{q}$ when $r \neq 1$, we have therefore obtained the desired representation for all a, b with $\text{supp}(a) \subseteq \mathcal{A}$ and $\text{supp}(b) \subseteq \mathcal{B}$:

$$\begin{aligned} U(\mu a \oplus (1 - \mu)b) &= \mu^r U(a) + (1 - \mu)^r U(b) + q \cdot H_r(\mu) \\ \text{with } H_r(\mu) &= \begin{cases} -\mu \ln \mu - (1 - \mu) \ln(1 - \mu), & r = 1 \\ \frac{1 - \mu^r - (1 - \mu)^r}{r - 1}, & r \neq 1 \end{cases} \end{aligned} \quad (20)$$

Substituting the solutions for $H(\mu) = B(\mu) + B(1 - \mu)$ and $F(a, \mu) = \mu^r U(a)$ into (19) yields the desired representation for disjoint supports $\mathcal{A}, \mathcal{B}, \mathcal{C}$.

Finally, we extend the representation to arbitrary disjoint supports. Let $\mathcal{Z}', \mathcal{Z}'' \subseteq \mathcal{Z}$ be disjoint and let $z', z'' \in \mathcal{M}$ be elements with $\text{supp}(z') = \mathcal{Z}'$ and $\text{supp}(z'') = \mathcal{Z}''$. We are trying to establish the functional form of $U(\mu z' \oplus (1 - \mu)z'')$ for $\mu \in (0, 1)$.

Since every superset of an essential set is essential, we may assume that $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ partition $\mathcal{Z}$. By decomposability, z' can be expressed as a mixture of elements a' , b' , c' with supports $\mathcal{A}' = \mathcal{Z}' \cap \mathcal{A}$, $\mathcal{B}' = \mathcal{Z}' \cap \mathcal{B}$, and $\mathcal{C}' = \mathcal{Z}' \cap \mathcal{C}$, and z'' as a mixture of elements a'' , b'' , c'' with supports $\mathcal{A}'' = \mathcal{Z}'' \cap \mathcal{A}$, $\mathcal{B}'' = \mathcal{Z}'' \cap \mathcal{B}$, and $\mathcal{C}'' = \mathcal{Z}'' \cap \mathcal{C}$.

Any two of these elements whose supports lie in different components are already covered by the representation obtained so far. It is sufficient to establish the desired functional form for mixtures containing a' and a'' . Commutativity yields then the same result for b', b'', c', c'' .

Let b and c be some arbitrary elements with support in $\mathcal{B}$ and $\mathcal{C}$ respectively and let $\mu, \lambda, \nu \in (0, 1)$. By (20) and an analogous result for supports $\mathcal{A}$ and $\mathcal{C}$:

$$\begin{aligned} U(x') &\equiv U(\mu a' \oplus (1 - \mu)b) = \mu^r U(a') + (1 - \mu)^r U(b) + qH_r(\mu) \\ U(x'') &\equiv U(\lambda a'' \oplus (1 - \lambda)c) = \lambda^r U(a'') + (1 - \lambda)^r U(c) + qH_r(\lambda). \end{aligned}$$

Although x' and x'' have disjoint supports $\mathcal{A}' \cup \mathcal{B}$ and $\mathcal{A}'' \cup \mathcal{C}$, we cannot use Lemma 8 to conclude a quasi-separable representation for this pair since Lemma 7 requires a third essential component disjoint from both, which may not be available. However, since the elements with support $\mathcal{A}' \cup \mathcal{B}$ and $\mathcal{A}'' \cup \mathcal{C}$ are essential and preferences separable given a fixed ν , it is straightforward to obtain a representation $U(\nu x' \oplus (1 - \nu)x'') = V(U(x'), U(x''), \nu)$. Using associativity, we can rearrange the left-hand side and obtain

$$U(\nu x' \oplus (1 - \nu)x'') = (\nu(1 - \mu))^r U(b) + ((1 - \nu)(1 - \lambda))^r U(c) + G(a', a'', \nu, \mu, \lambda),$$

where the remaining term G contains the utility of the mixture of a' and a'' , and the entropy of $\nu(1 - \mu)$, $(1 - \nu)(1 - \lambda)$, and $\nu\mu + (1 - \nu)\lambda$. It follows that $V(u', u'', \nu) = u'\phi(\nu) + u''\psi(\nu) + \chi(\nu)$ for functions ϕ, ψ, χ . It is then straightforward to determine $\phi(\nu) = \nu^r$, $\psi(\nu) = (1 - \nu)^r$, and $\chi(\nu) = qH_r(\nu)$ by comparison with the left-hand side.

Uniqueness follows as in the procedural mixture case: nontrivial preferences yield an additively separable representation over a continuum, so U is unique up to affine transformations, and the uniqueness properties of r and q follow immediately. $\square$

REFERENCES

- Birnbaum, M. H. (2004a). Causes of Allais common consequence paradoxes: An experimental dissection. *Journal of Mathematical Psychology*, 48, 87–106.
- Birnbaum, M. H. (2004b). Tests of rank-dependent utility and cumulative prospect theory in gambles represented by natural frequencies: Effects of format, event framing, and branch splitting. *Organizational Behavior and Human Decision Processes*, 95, 40–65.
- Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review*, 115(2), 463–501.
- Birnbaum, M. H., & Bahra, J. P. (2007). Gain-loss separability and coalescing in risky decision making. *Management Science*, 53(6), 1016–1028.
- Birnbaum, M. H., & Navarrete, J. B. (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. *Organizational Behavior and Human Decision Processes*, 73(2–3), 237–273.
- Border, K. C., & Segal, U. (1994). Dynamic consistency implies approximately expected utility preferences. *Journal of Economic Theory*, 63(2), 170–188.
- Caplin, A., Dean, M., & Leahy, J. (2017, August). *Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy* (w23652). National Bureau of Economic Research. Cambridge, MA. <https://doi.org/10.3386/w23652>
- Cerreia-Vioglio, S., Dillenberger, D., Ortoleva, P., & Riella, G. (2019). Deliberately stochastic. *American Economic Review*, 109(7), 2425–2445.
- Chen, T.-Y., Rommeswinkel, H., Wang, G.-Y., & Yang, T.-S. (2021). *Measuring Consumer Freedom*. 3807668. <https://doi.org/10.2139/ssrn.3807668>
- Csiszár, I. (2008). Axiomatic Characterizations of Information Measures. *Entropy*, 10, 261–273.
- Ebanks, B. R., Kannappan, P., & Ng, C. T. (1987). Generalized fundamental equation of information of multiplicative type. *Aequationes Mathematicae*, 32(1), 19–31. <https://doi.org/10.1007/BF02311295>
- Fishburn, P. C. (1982). *Foundations of expected utility*. Reidel.
- Fishburn, P. C., & Roberts, F. S. (1978). Mixture axioms in linear and multilinear utilities. *Theory and Decision*, 9, 161–171.
- Frankel, D., & Volij, O. (2011). Measuring school segregation. *Journal of Economic Theory*, 146(1), 1–38.

- Fudenberg, D., Iijima, R., & Strzalecki, T. (2015). Stochastic choice and revealed perturbed utility. *Econometrica*, 83(6), 2371–2409.
- Gilboa, I., Postlewaite, A., & Samuelson, L. (2016). Memorable consumption. *Journal of Economic Theory*, 165, 414–455. <https://doi.org/10.1016/j.jet.2016.05.004>
- Grant, S. (1995). Subjective probability without monotonicity: Or how Machina's mom may also be probabilistically sophisticated. *Econometrica*, 63(1), 159–189.
- Hammond, P. J. (1988). Consequentialist Foundations for Expected Utility. *Theory and Decision*, 25(1), 25–78.
- Hashidate, Y. (2018). *Preferences for randomization and anticipated utility* [CIRJE Discussion Paper F-1083, University of Tokyo].
- Hennessy, D. A., & Lapan, H. (2007). When different market concentration indices agree. *Economics Letters*, 95(2), 234–240. <https://doi.org/10.1016/j.econlet.2006.10.011>
- Herstein, I. N., & Milnor, J. (1953). An Axiomatic Approach to Measurable Utility. *Econometrica*, 21(2), 291–297. <https://doi.org/10.2307/1905540>
- Humphrey, S. J. (1995). Regret aversion or event-splitting effects? More evidence under risk and uncertainty. *Journal of Risk and Uncertainty*, 11, 263–274.
- Humphrey, S. J. (1999). Probability learning, event-splitting effects and the economic theory of choice. *Journal of Economic Psychology*.
- Jensen, N.-E. (1967). An introduction to Bernoullian utility theory. I. Utility functions. *Swedish Journal of Economics*, 69, 163–183.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement volume 1: Additive and polynomial representations* (Vol. 1). Dover Publications.
- Luce, R. D., Ng, C. T., Marley, A. A. J., & Aczél, J. (2008a). Utility of gambling I: Entropy modified linear weighted utility. *Economic Theory*, 36(1), 1–33. <https://doi.org/10.1007/s00199-007-0260-5>
- Luce, R. D., Ng, C. T., Marley, A. A. J., & Aczél, J. (2008b). Utility of gambling II: Risk, paradoxes, and data. *Economic Theory*, 36(2), 165–187. <https://doi.org/10.1007/s00199-007-0259-y>
- Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature*, 27(4), 1622–1668.

- Marley, A. A. J., & Luce, R. D. (2005). Independence properties vis-à-vis several utility representations. *Theory and Decision*, 58, 77–143.
- Mattsson, L.-G., & Weibull, J. W. (2002). Probabilistic choice and procedurally bounded rationality. *Games and Economic Behavior*, 41(1), 61–78.
- Miao, B., & Zhong, S. (2018). Probabilistic social preference: How Machina’s mom randomizes her choice. *Economic Theory*, 65(1), 1–24. <https://doi.org/10.1007/s00199-016-1015-y>
- Mongin, P. (2001). A note on mixture sets in decision theory. *Decisions in Economics and Finance*, 24(1), 59–69. <https://doi.org/10.1007/s102030170010>
- Nehring, K., & Puppe, C. (2009). Diversity. In *Handbook of Rational and Social Choice*. Oxford University Press.
- Payró, F. (2022, September). *Mixture-dependent preference for commitment* (BSE Working Paper No. 1365), Barcelona School of Economics.
- Qin, W.-z., & Rommeswinkel, H. (2024). Quasi-separable Preferences. *Theory and Decision*, 96, 555–595.
- Quispe-Torreblanca, E., Stewart, N., & Birnbaum, M. H. (2025). Surprisingly robust violations of stochastic dominance despite splitting training: A quasi-adversarial collaboration. *Judgment and Decision Making*, 20.
- Rényi, A. (1961). On measures of information and entropy. *Proceedings of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability 1960*, 547–561.
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(3), 379–423, 623–656.
- Shorrocks, A. F. (1980). The Class of Additively Decomposable Inequality Measures. *Econometrica*, 48(3), 613.
- Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3), 665–690. [https://doi.org/10.1016/S0304-3932\(03\)00029-1](https://doi.org/10.1016/S0304-3932(03)00029-1)
- Starmer, C., & Sugden, R. (1993). Testing for juxtaposition and event-splitting effects. *Journal of Risk and Uncertainty*, 6, 235–254.
- Suppes, P. (1996). The nature and measurement of freedom. *Social Choice and Welfare*, 13(2), 183–200.
- Theil, H. (1965). The Information Approach to Demand Analysis. *Econometrica*, 33(1), 67–87.
- Theil, H. (1967). *Economics and Information Theory*. North-Holland.

- Trautmann, S. T. (2009). A tractable model of process fairness under risk. *Journal of Economic Psychology*, 30(5), 803–813.
- Tsallis, C. (1988). Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, 52, 479–487.
- Volij, O. (1994). Dynamic consistency, consequentialism and reduction of Compound Lotteries. *Economics Letters*, 46(2), 121–129.
- Wakker, P. P. (1989). *Additive representations of preferences: A new foundation of decision analysis* (Vol. 4). Springer Science & Business Media.
- Weber, B. J. (2007). The effects of losses and event splitting on the Allais paradox. *Judgment and Decision Making*, 2(2), 115–125.