Discrete Random Expected Utility and Self-Selection

Igor Kopylov* Kevin Peng†‡

November 14, 2025

JOB MARKET PAPER
[click here for latest version]

SELECT SLIDES FOR ASSOCIATED EXPERIMENTAL PROJECT INCLUDED IN APPENDIX B. (Peng, 2025)

Abstract

We model stochastic choice rules via finitely many types θ that maximize distinct expected utility functions and use endogenous tie-breaking rules. First, we characterize *discrete random expected utility* (DREU) where the likelihoods $\mu(\theta)$ of such types are preserved across all menus A. This model is a discrete version for the *random expected utility* of Gul and Pesendorfer (2006), but our axioms, identification, and tie-breaking methods are distinct. More generally, we propose *discrete-map expected utility* (DMEU) where the likelihoods $\mu_A(\theta)$ are contingent on the menu A and hence, capture various kinds of *context dependence*. This extension violates monotonicity together with other assumptions of DREU. If monotonicity does hold, then our model can be interpreted in terms of *self-selection*, where types can increase their participation across distinct menus but only if their best choices are improved. All components of our representations are identified uniquely. Finally, we discuss applications to heterogeneous risk attitudes, beliefs, and Cobb-Douglas utility functions.

*Department of Economics, University of California, Irvine. E-mail: ikopylov@uci.edu.

†Department of Economics, University of California, Irvine. E-mail: kuangyup@uci.edu

‡We are grateful to John Duffy, Stergios Skaperdas, Jay Lu, Yusufcan Masatligolu, and Chris Chambers for their comments.

1 Introduction

Observed choices in consumption menus can be naturally *stochastic* when they are produced by distinct rational (i.e., utility maximizing) agents. Data is often presented as an aggregation of choices from diverse agents. Such data is used in a wide range of empirical work, such as preferred modes of transportation (McFadden, 2001) and choice of fishing sites (Train, 1998). This heterogeneity may also result from more behavioral motivations where a single individual can exhibit distinct random responses in identical choice problems (see Agranov and Ortoleva (2017)).

The classic random utility model (RUM) by Block and Marschak (1960) aims to represent stochastic choice data by a stable probability distribution μ across *types* that include all possible utility functions. In the general case where types are unrestricted, Falmagne (1978) characterizes RUM via non-negativity of Block and Marschak's polynomials. Gul and Pesendorfer (2006) (henceforth, GP) refine this model to *random expected utility* (REU) where all types should maximize expected utility functions rather than arbitrary ones. Other refinements of RUM have restricted types by single-crossing conditions (Apesteguia, Ballester, and Lu, 2017), quasi-linearity (McFadden, 1973; Williams, 1977; Daly and Zachery, 1979; Yang and Kopylov, 2023), and other assumptions. Behavioral patterns, such as ambiguity aversion (Lu, 2021), have also been discussed within the context of random utility. All of these refinements preserve the assumption that μ is invariant of the consumption menu.

In this paper, we propose a novel version of random utility that has

- a discrete (i.e., finite) type space Θ where all types $\theta \in \Theta$ combine expected utility maximization with endogenous tie-breaking rules;
- a distribution of types μ_A that can depend on the menu A and capture context dependence and self-selection;
- a unique identification for both Θ and μ_A.

We still consider a representation where the distribution $\mu(\theta)$ is invariant of A. This case is a discrete version of the GP's REU, but our axioms and identifications are more parsimonious than theirs and rely more on algebraic rather than analytical techniques. Moreover, the algebraic approach makes the consumption domain X more flexible than in GP, where X is a finite-dimensional simplex of lotteries. In our framework, X is convex, but need not be closed, or bounded, or finite dimensional. For example, X may consist of monetary gambles with payoffs on the real line, or X can be the non-negative orthant of consumption bundles.

More substantially, we depart from REU by allowing the distribution μ_A to vary with the menu A. In general, it is only required that the support $\Theta = \{\theta : \mu_A(\theta) > 0\}$ be finite and invariant of A. This Θ is called a discrete type space. It is used to represent observable likelihoods $\rho(x, A)$ of choosing any element x in any finite consumption menu A. Our most general representation takes the form

$$\rho(x, A) = \mu_A\{\theta \in \Theta : x \text{ is the best element for } \theta \text{ in } A\} \tag{1}$$

where $\mu_A(\theta)$ is the probability of type θ making a choice in the menu A.

Each type θ is assumed to maximize some expected utility function u_θ in each menu A. However, in order for representation (1) to be well-defined, each type $\theta \in \Theta$ must also incorporate some endogenous tie-breaking rule that selects a unique maximizer for u_θ in A. Formally, each θ is associated with a total, complete, transitive binary relation that satisfies von Neumann-Morgenstern's Independence axiom as well. Such total types have been used in finite-dimensional contexts by Fishburn (1982), Myerson (1986), Blume, Branderburger, and Dekel (1991) and others. We argue that there is a rich family of total types even if X is not finite dimensional. Thus we provide a novel way to accommodate tie-breaking in random expected utility. Without totality, Piermont (2022) relaxes GP's model to represent only those choices that are strict maximizers for corresponding types.

We show several characterization results for stochastic choice rules. First, Theorem 2 below characterizes representation where $\mu_A = \mu$ is menu-invariant and relates this result to GP's model. Our main result (Theorem 3 below) characterizes representation (1) and provides a convenient platform for various refinements that impose more structure on the menu-contingent distribution μ_A and/or types θ. We obtain several such refinements.

Theorem 4 summarizes the implications of the standard monotonicity principle that asserts

$$\rho(x, A) \geq \rho(x, A \cup B)$$

for all menus $A, B \in \mathcal{M}$ and elements $x \in A$. It turns out to be equivalent for the function μ_A to satisfy

$$\mu_A(\theta) \geq \mu_{A \cup B}(\theta)$$

for all types $\theta \in \Theta$ that have the same maximal element in A and in $A \cup B$. This finding can be interpreted in terms of *self-selection*: types can increase their participation in the menu $A \cup B$, but only if they find better choices in the larger menu. For example, it should be reasonable for gamblers to show up in higher numbers to bookmakers that offer them more favorable odds on some sporting event. What's more, the expansion of bet types with the rise of online gambling can incentivize new customers to enter the market (Hing et al., 2022). Self-selection can also be observed in the dining industry. Garnett et al. (2019) found that doubling the options of vegetarian meals increased sales of such meals by around 15%.

By contrast, self-selection cannot be detected in the general RUM examples because it can keep all Block-Marschak polynomials non-negative. In that example, the presence of self-selection makes the general RUM *misspecified* without rejecting this model outright. Our model does not have this identification issue and disentangles self-selection from the composition of the type space Θ.

More broadly, there are some behavioral patterns, such as reasoned-based heuristics in Shafir, Simonson, and Tversky (1993), where Monotonicity does not hold and hence, self-selection is no longer a plausible explanation. We use *context dependence* as a blanket term to refer to all such patterns (see a literature review by Rabin (1998)). An example of such patterns is extremeness aversion. Sharpe et al. (2008) observed that the removal of the largest drink size (44 oz) from the menu led consumers to shift their purchases from the 32-oz option—now the largest available size—to the 21-oz size.

Finally, we discuss applications where the type space Θ captures heterogeneous *risk attitudes* and *beliefs*. For example, agents make more risk neutral choices when lotteries are presented in a binary format (i.e., only two states of the world have positive probabilities) instead of lotteries that are more complex or compound lotteries (Harrison et al., 2013). Changes in beliefs arise in the form of partition dependence. Ahn and Ergin (2010) observed that agents, when choosing deductible plans, place a higher weight on surgeries when they are individually listed than when they are bundled under the umbrella term "surgery". In such situations, the consumption domain X can be taken to consist of lotteries or Anscombe-Aumann (1963) uncertain prospects.

2 Preliminaries

Let $X = \{x, y, z, \dots\}$ be a consumption domain. Assume that X is a non-singleton convex subset of a normed linear space L. For example, X can be a simplex

$$\Delta^n = \left\{ x \in \mathbb{R}^n_+ : \sum_{i=1}^{n} x_i = 1 \right\}$$

of all *lotteries* on a finite set of deterministic prizes $\{1, ..., n\}$. Then X is a convex subset of $L = \mathbb{R}^n$. In general, X need not be finite-dimensional, and its elements need not be interpreted as lotteries.

Let $\mathcal{M} = \{A, B, \dots\}$ be the set of all *menus*—non-empty finite subsets of X. Singletons $\{x\}$ are written as x.

A binary relation θ on X is called a *type* if it is complete, transitive, and satisfies Independence: for all $x, y, z \in X$ and $\gamma \in (0, 1]$,

$$x\theta y \iff [\gamma x + (1-\gamma)z]\theta[\gamma y + (1-\gamma)z].$$

For any type θ and menu $A \in \mathcal{M}$, say that $x \in A$ is *rational* for θ in A if $x\theta y$ for all $y \in A$. Let

$$R_\theta(A) = \{x \in A : x \text{ is rational for } \theta \text{ in } A\}.$$

This set is non-empty because θ is complete and transitive, and A is finite.

Let $\mathcal{T}$ be the set of all *total* types such that for all $x, y \in X$, $x\theta y\theta x$ implies $x = y$. For any $A \in \mathcal{M}$ and $x \in A$, let

$$\mathcal{T}_x(A) = \{\theta \in \mathcal{T} : x \text{ is rational for } \theta \text{ in } A\}.$$

As $R_\theta(A)$ is a singleton for each $\theta \in \mathcal{T}$, $\mathcal{T}$ is partitioned into a disjoint union

$$\mathcal{T} = \bigcup_{x \in A} \mathcal{T}_x(A).$$

For any $A, B \in \mathcal{M}$ and $\gamma \in [0, 1]$, define a mixture (see Figure 1 for an example)

$$\gamma A + (1-\gamma)B = \{\gamma x + (1-\gamma)y : x \in A, y \in B\}.$$

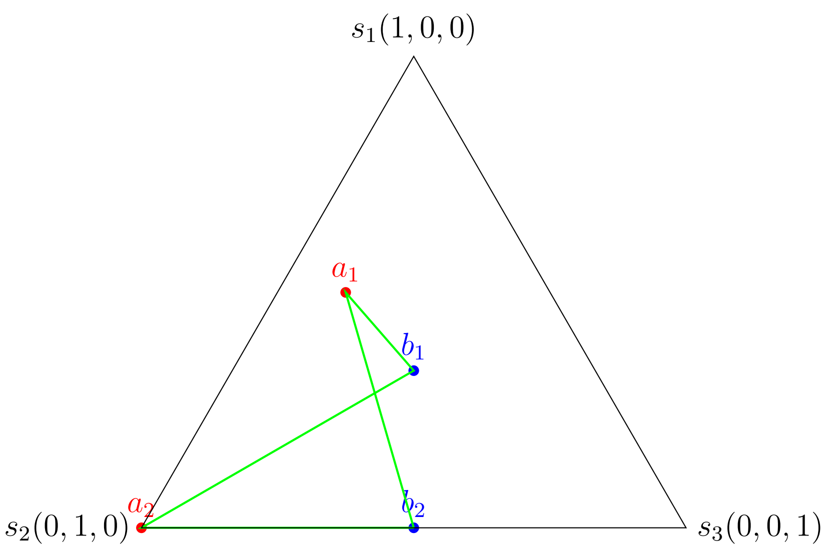


Figure 1: All points on the green lines form the mixture of $A = \{a_1, a_2\}$ and $B = \{b_1, b_2\}$.

If $\gamma \in (0,1]$, then for any $x \in A$, $y \in B$, and type θ, Independence implies

$$x \in R_\theta(A) \iff \gamma x + (1-\gamma)y \in R_\theta(\gamma A + (1-\gamma)B) \text{ for some } y \in B. \tag{2}$$

Therefore, for any $x \in A$, the set $\mathcal{T}_x(A)$ is partitioned by

$$\mathcal{T}_x(A) = \bigcup_{y \in [\gamma x + (1-\gamma)B]} \mathcal{T}_y(\gamma A + (1-\gamma)B). \tag{3}$$

This union holds by (2). It is disjoint because $R_\theta(\gamma A + (1-\gamma)B)$ is a singleton for any type $\theta \in \mathcal{T}$.

2.1 Expected Utility Maximization

Let $\mathcal{U} = \{u, v, \dots\}$ be the set of all non-constant functions $u : X \to \mathbb{R}$ such that for all $x, y \in X$ and $\gamma \in [0,1]$,

$$u(\gamma x + (1-\gamma)y) = \gamma u(x) + (1-\gamma)u(y). \tag{4}$$

If $X = \Delta^n$ is a simplex of lotteries, then $\mathcal{U}$ consists of all *expected utility* functions on X. More broadly, functions $u \in \mathcal{U}$ may have interpretations other than expected utility, but we still refer to maximization of such functions as expected utility maximization.

A type θ is *null* if $x\theta y$ for all $x, y \in X$; otherwise, θ is *non-null*. A type θ is *mixture continuous* if for all $x, y, z \in X$,

$$\{\gamma \in [0,1] : z\theta[\gamma x + (1-\gamma)y]\} \text{ and } \{\gamma \in [0,1] : [\gamma x + (1-\gamma)y]\theta z\}$$

are closed sets. Otherwise, θ is *discontinuous*.

Let $\mathcal{C}$ be the class of all types that are mixture continuous and non-null. Herstein and Milnor's (1953) mixture space theorem asserts that $\theta \in \mathcal{C}$ if and only if θ can be represented by some $u \in \mathcal{U}$: for all $x, y \in X$,

$$x\theta y \iff u(x) \geq u(y).$$

Moreover, such u is unique up to a positive linear transformation: if θ is represented by another $v \in \mathcal{U}$, then $v = \alpha u + \beta$ for some $\alpha > 0$ and $\beta \in \mathbb{R}$.

However, this representation cannot be derived for total types that are usually discontinuous. Indeed, let X be a convex subset of a line $\{x + \alpha d : \alpha \in \mathbb{R}\}$ for some $x, d \in L$ such that $d \neq 0$. Call such X *collinear*. Then there are only two non-null types θ_+ and θ_-

that are represented by $u(x+\alpha d)=\alpha$ and $u(x+\alpha d)=-\alpha$ respectively. Both are total and mixture continuous.

In the generic *non-collinear* case, X is not a subset of any line. Then all total types are discontinuous, that is,

$$\mathcal{T}\cap\mathcal{C}=\varnothing.$$

To show this, take $x,y,z\in X$ such that $x-y$ and $y-z$ are linearly independent. Let $\theta\in\mathcal{C}$. Then it is represented by some $u\in\mathcal{U}$. Without loss in generality, $u(x)\geq u(y)\geq u(z)$. Then either $u(x)=u(z)$, or $u(y)=u(\gamma x+(1-\gamma)z)$ where $\gamma=\frac{u(y)-u(z)}{u(x)-u(z)}$. As $x\neq z$ and $y\neq\gamma x+(1-\gamma)z$, then $\theta\notin\mathcal{T}$ in either case.

For any type θ, the mixture space theorem can be applied to its *closure* θ^*. Define θ^* as a binary relation on X such that for any $x,y\in X$, $x\theta^*y$ if and only if there are $z_x,z_y\in X$ such that for all $\gamma\in[0,1)$,

$$[\gamma x+(1-\gamma)z_x]\theta[\gamma y+(1-\gamma)z_y]. \tag{5}$$

In particular, $x\theta y$ implies $x\theta^*y$ because (5) holds for $z_x=x$ and $z_y=y$.

Theorem 1. *For any type θ, its closure θ^* is a mixture continuous type. The closure θ^* is represented by $u\in\mathcal{U}$ if and only if for all $x,y\in Z$,*

$$x\theta y \implies u(x)\geq u(y). \tag{6}$$

Moreover, if θ is non-null and X is finite dimensional, then θ^* is non-null.

The proof is in the appendix. Together with the mixture space theorem, this result delivers two possible cases for any non-null type θ. If its closure θ^* is non-null, then the one-way representation (6) holds for some $u\in\mathcal{U}$ that is unique up to a positive linear transformation. If θ^* is null, then (6) does not hold for any $u\in\mathcal{U}$. The latter case never happens when X is finite dimensional. Note that (6) holds trivially if u is constant, but constant functions do not belong to $\mathcal{U}$.

Representation (6) can be refined further by combining the maximization of $u\in\mathcal{U}$ with a tie-breaking rule that is used when $u(x)=u(y)$. See Theorem 5 below.

3 Main Representation Results

Let Ω be the set of all pairs (x,A) such that $A\in\mathcal{M}$ and $x\in A$, that is, x is a feasible element in a menu A. Such pairs are called *trials*.

A function $\rho : \Omega \to [0,1]$ is called a *stochastic choice rule* (scr) if

$$\sum_{x \in A} \rho(x, A) = 1 \text{ for all } A \in \mathcal{M}. \tag{7}$$

Here, the probability $\rho(x, A)$ of any trial $(x, A) \in \Omega$ can be interpreted as the likelihood of x being chosen when the menu A is feasible.

A function $\mu : 2^{\mathcal{T}} \to \mathbb{R}_+$ is called a *discrete distribution* if

(i) the set $\Theta = \{\theta \in \mathcal{T} : \mu(\theta) > 0\}$ is finite,

(ii) for all sets $\Psi \subset \mathcal{T}$, $\mu(\Psi) = \sum_{\theta \in \Theta \cap \Psi} \mu(\theta)$,

(iii) $\mu(\mathcal{T}) = \mu(\Theta) = 1$.

The set Θ is called the *support* of μ. Obviously, μ is a countably additive[1] probability measure, and its support Θ is determined uniquely.

Let $\mathcal{D}$ be the set of all discrete distributions. Say that a stochastic choice rule ρ is *represented* by a discrete distribution $\mu \in \mathcal{D}$ if for any trial $(x, A) \in \Omega$,

$$\rho(x, A) = \mu(\mathcal{T}_x(A)) \tag{8}$$

is the likelihood that μ assigns to all total types for which x is rational in A. By Theorem 1, total types $\theta \in \mathcal{T}$ with non-null closures can be interpreted in terms of expected utility maximization. Thus we refer to representation (8) as *discrete random expected utility* (DREU).

To get some observable implications of DREU, take any $(x, A) \in \Omega$, $B \in \mathcal{M}$, and $\gamma \in (0, 1]$. Let

$$\rho(x, A, \gamma, B) = \sum_{y \in [\gamma x + (1-\gamma) B]} \rho(y, \gamma A + (1-\gamma) B)$$

be the probability of all choices in the menu $\gamma A + (1-\gamma)B$ that belong to its submenu $\gamma x + (1-\gamma)B$.

Axiom 1 (Additivity)**.** $\rho(x, A) = \rho(x, A, \gamma, B)$.

This equality follows from (8) that applies the discrete distribution μ to both sides of the partition (3). In particular, Additivity requires

Axiom 2 (Monotonicity)**.** $\rho(x, A) \geq \rho(x, A \cup B)$.

[1]For discrete distributions, countable and finite additivity are equivalent.

Monotonicity is assumed by Block and Marschak for their general random utility model. It holds in all of its refinements that take the distribution of types μ to be invariant across menus A. It holds in DREU as well, but need not be assumed separately in the presence of Additivity, which implies

$$\rho(x, A) = \rho(x, A, 0.5, A \cup B) \geq \rho(x, A \cup B, 0.5, A) = \rho(x, A \cup B).$$

Here the middle inequality holds because $0.5x + 0.5(A \cup B) \supset 0.5x + 0.5A$.

As μ is discrete, it has a finite support $\Theta \subset \mathcal{T}$, and by (8),

$$\rho(x, A) > 0 \iff x = R_\theta(A) \text{ for some } \theta \in \Theta. \tag{9}$$

Let $k_\rho(A) = |\{x \in A : \rho(x, A) > 0\}|$ be the number of elements $x \in A$ that have positive probability to be chosen in A. By (9), $k_\rho(A) \leq |\Theta|$.

Axiom 3 (Discreteness)**.** *There is $n \in \mathbb{N}$ such that $k_\rho(A) \leq n$ for all $A \in \mathcal{M}$.*

Discreteness is a convenient technical assumption that cannot be rejected by finitely many observations.

Theorem 2. *A stochastic choice rule ρ satisfies Additivity and Discreteness if and only if ρ is represented by some discrete distribution $\mu \in \mathcal{D}$. Such μ is unique.*

The proof is constructive. By Discreteness, there is a maximal menu $B = \{b_1, \ldots, b_k\}$ such that $\rho(b_i, B) > 0$ for all i. For any $x, y \in X$ and $i = 1, \ldots, k$, let

$$x\theta_i y \iff \rho(\alpha x + (1-\alpha)b_i, \alpha\{x, y\} + (1-\alpha)B) > 0, \tag{10}$$

for sufficiently small $\alpha > 0$. Then Additivity delivers (8) where

$$\Theta = \{\theta_1, \ldots, \theta_k\} \subset \mathcal{T} \quad \text{and} \quad \mu(\theta_i) = \rho(b_i, B) \text{ for all } i = 1, \ldots k.$$

All details are in the appendix.

Note that any function $\rho : \Omega \to \mathbb{R}_+$ that is represented by some discrete distribution $\mu \in \mathcal{D}$ via (8) must be a stochastic choice rule because for any menu $A \in \mathcal{M}$, $\mathcal{T} = \bigcup_{x \in A} \mathcal{T}_x(A)$ is a partition and hence,

$$\sum_{x \in A} \rho(x, A) = \sum_{x \in A} \mu(\mathcal{T}_x(A)) = \mu(\mathcal{T}) = 1.$$

Thus the uniqueness claim in Theorem 2 establishes a bijection between $\mathcal{D}$ and the class of all stochastic choice rules that are represented by discrete distributions.

To illustrate the role of total types in Theorem 2, consider another definition. For any $\theta \in \mathcal{T}$ and $(x, A) \in \Omega$, let

$$\rho_\theta(x, A) = \begin{cases} 1 & \text{if } x = R_\theta(A) \\ 0 & \text{if } x \neq R_\theta(A). \end{cases}$$

Obviously, the function $\rho_\theta : \Omega \to [0, 1]$ is a stochastic choice rule, and representation (8) asserts that

$$\rho = \sum_{\theta \in \Theta} \mu(\theta) \rho_\theta$$

is a convex combination of ρ_θ across $\theta \in \Theta$. By contrast, there is no convenient way to specify ρ_θ when $\theta \in \mathcal{C}$ because of tie-breaking. For example, the uniform tie-breaking rule violates both Discreteness and Additivity. Indeed, let X be non-collinear and $\theta \in \mathcal{C}$. Then there are distinct $x, y \in X$ such that $u(x) = u(y)$. Let $B = \{x, y\}$, $z = 0.5x + 0.5y$, and $C = 0.5B + 0.5B = \{x, y, z\}$. The uniform tie-breaking implies that

$$\begin{aligned} &\rho_\theta(x, B) = \rho_\theta(y, B) = \tfrac{1}{2}, \\ &\rho_\theta(x, C) = \rho_\theta(y, C) = \rho_\theta(z, C) = \tfrac{1}{3}, \end{aligned}$$

which violates Additivity because $\theta(x, B) \neq \theta(x, B, 0.5, B)$. Next, for any $n \in \mathbb{N}$, consider a menu $A = \{a_1, \ldots, a_n\}$ where $a_i = \frac{i}{n}x + \frac{n-1}{n}y$. The uniform tie-breaking implies that $\rho(a_i, A) = \frac{1}{n}$ for each i. As n is arbitrary, then Discreteness fails.

In Section 4, we compare DREU and Theorem 2 with GP's REU and their representation results.

3.1 Context Dependence

Additivity and representation (8) can be problematic if the distribution of types in Θ is context dependent. For example, adding extra options to a menu A can increase the probability of choosing $x \in A$ due to reason-based heuristics (e.g., Shafir, Simonson, and Tversky (1993)) and other behavioral reasons. Thus Monotonicity and a fortiori, Additivity can be violated. To accommodate various kinds of context dependence, discrete distributions can be assigned contingently on the feasible menu A.

Metrize $\mathcal{D}$ and $\mathcal{M}$ by the uniform and Hausdorff metrics respectively so that such that for all $\mu, \pi \in \mathcal{D}$ and $A, B \in \mathcal{M}$,

$$d_u(\mu, \pi) = \max_{\theta \in \mathcal{T}} |\mu(\theta) - \pi(\theta)|$$
$$d_h(A, B) = \max \left\{ \max_{x \in A} \min_{y \in B} ||x - y||, \max_{y \in B} \min_{x \in A} ||x - y|| \right\}.$$

Consider a function $\mu : \mathcal{M} \to \mathcal{D}$ that assigns discrete distributions $\mu_A \in \mathcal{D}$ contingently on the menu A. Note that the value of μ at A is written as μ_A rather than $\mu(A)$.

Say that $\mu : \mathcal{M} \to \mathcal{D}$ is a *discrete map* if μ is continuous, and for all $A, B \in \mathcal{M}$, μ_A and μ_B have the same support

$$\Theta = \{\theta \in \mathcal{T} : \mu_A(\theta) > 0\} = \{\theta \in \mathcal{T} : \mu_B(\theta) > 0\}.$$

Let $\mathcal{D}(\mathcal{M})$ be the set of all discrete maps. With some abuse of notation, any discrete distribution $\mu \in \mathcal{D}$ is identified with the constant discrete map $\mu \in \mathcal{D}(\mathcal{M})$.

Say that a stochastic choice rule ρ is represented by a discrete map $\mu \in \mathcal{D}(\mathcal{M})$ if for all trials $(x, A) \in \Omega$,

$$\rho(x, A) = \mu_A(\mathcal{T}_x(A)) \tag{11}$$

is the likelihood that is assigned by μ_A to all total types θ for which x is rational in A. Obviously, (8) is a special case of (11) where the discrete map μ is constant.

Call the general representation (11) *discrete-map expected utility* (DMEU). This model violates Additivity because it applies distinct distributions μ_A and $\mu_{\gamma A+(1-\gamma)B}$ on the two sides of the partition (3). Similarly, Monotonicity can fail because the distributions μ_A and $\mu_{A\cup B}$ need not be the same.

Instead, DMEU implies a weaker form of Additivity.

Axiom 4 (Boundary Additivity (BA))**.** *For all $(x, A) \in \Omega$, $B \in \mathcal{M}$, and $\gamma \in (0, 1]$,*

$$\rho(x, A) = 0 \iff \rho(x, A, \gamma, B) = 0.$$

This condition imposes Additivity only in the boundary case when either side in the equation $\rho(x, A) = \rho(x, A, \gamma, B)$ is zero. Either of these cases is equivalent to the claim that x is not rational in A for any type $\theta \in \Theta$ in the support of the discrete map μ.

The continuity of the discrete map μ in DMEU translates into

Axiom 5 (Asymptotic Additivity (AA))**.** *For all $(x, A) \in \Omega$ and converging sequences*

$\{\gamma_m\}_{m=1}^{\infty} \in (0,1]$ *and* $\{B_n\}_{n=1}^{\infty} \in \mathcal{M}$,

$$\rho(x,A) > 0 \implies \lim_{m,n\to\infty} \rho(x,A,\gamma_m,B_n) > 0$$

Here the convergence $\lim_{m,n\to\infty} \rho(x,A,\gamma_m,B_n) = \alpha > 0$ is defined as a double limit: for any $\varepsilon > 0$, there is $k \in \mathbb{N}$ such that for all $m,n \geq k$,

$$|\rho(x,A,\gamma_m,B_n) - \alpha| < \varepsilon.$$

If Additivity holds, then AA is trivial because $\rho(x,A,\gamma_m,B_n) = \rho(x,A)$ for all m,n.

Theorem 3. *A stochastic choice rule ρ satisfies BA, AA, and Discreteness if and only if ρ is represented by a discrete map $\mu \in \mathcal{D}(\mathcal{M})$. Such μ is unique, and its support Θ satisfies*

$$|\Theta| = \max_{A\in\mathcal{M}} k_\rho(A).$$

This result characterizes DMEU and establishes the uniqueness of the discrete map μ in this model. Moreover, it asserts that the size $|\Theta|$ of the support Θ can be achieved as the maximal number of elements chosen with positive probability in some menu.

Similarly to DREU, any function $\rho : \Omega \to \mathbb{R}_+$ that is represented by some discrete map $\mu \in \mathcal{D}(\mathcal{M})$ via (11) must be a stochastic choice rule because for any menu $A \in \mathcal{M}$,

$$\sum_{x\in A} \rho(x,A) = \sum_{x\in A} \mu_A(\mathcal{T}_x(A)) = \mu_A(\mathcal{T}) = 1.$$

Thus the uniqueness claim in Theorem 3 establishes a bijection between $\mathcal{D}(\mathcal{M})$ and the class of all stochastic choice rules that are represented by discrete maps.

3.2 Monotonicity and Self-Selection

Self-selection is a plausible form of context dependence for stochastic choices: types can increase their participation across distinct menus when their rational choices are improved. For example, it should be reasonable for gamblers to bet more actively with bookmakers who offer them more favorable odds.

Unlike the behavioral heuristics—such as reason-based choice or extremeness aversion—self-selection should be consistent with Monotonicity. This condition appears particularly important for stochastic choice models because it is convenient for empirical tests and normative interpretations.

A discrete map $\mu \in \mathcal{D}(\mathcal{M})$ is called *selective* if for all menus $A, B \in \mathcal{M}$ and types $\theta \in \mathcal{T}$,

$$R_\theta(A) = R_\theta(A \cup B) \implies \mu_A(\theta) \geq \mu_{A\cup B}(\theta). \tag{12}$$

Condition (12) asserts that all types in Θ that preserve their rational choice when A is increased to the larger menu $A \cup B$ should not become more likely in $A \cup B$. This condition is plausible under self-selection because such types should not increase their participation in $A \cup B$, while other types should not decrease their participation in $A \cup B$ because their rational choices can never deteriorate when A is replaced with $A \cup B$.

Theorem 4. *A stochastic choice rule ρ satisfies Monotonicity, BA, AA, and Discreteness if and only if ρ is represented by a selective discrete map μ.*

This result asserts that given Monotonicity, all representations by discrete maps can be interpreted in terms of self-selection. As a special case of DMEU, the uniqueness and non-redundancy claims from Theorem 3 apply as is in the self-selective case.

4 Discussion

Our results can be naturally compared with GP's random expected utility. In their model, $X = \Delta^n$ is a finite-dimensional simplex, and for all $(x, A) \in \Omega$,

$$\rho(x, A) = \mu(\mathcal{C}_x(A)), \tag{13}$$

where μ is a finitely additive probability measure over $\mathcal{C}$, and

$$\mathcal{C}_x(A) = \{\theta \in \mathcal{C} : x \text{ is rational for } \theta \text{ in } A\}.$$

Note that $\mathcal{C} = \bigcup_{x\in A} \mathcal{C}_x(A)$, but this union need not be disjoint because a mixture continuous type $\theta \in \mathcal{C}$ can have several rational elements in A. To make representation (13) consistent with normalization (7), it must be assumed that the probability of ties is zero in each menu A, that is,

$$\mu(\mathcal{C}_x(A) \cap \mathcal{C}_y(A)) = 0 \tag{14}$$

for all distinct $x, y \in A$. GP calls such distributions μ *regular*. Regularity prohibits μ to have a finite support, except for the case when X is collinear and there are only two possible types.

By contrast, our representation (8) uses total rather than mixture continuous types, and takes μ to be discrete. This change in mathematical structure allows to identify μ

more directly via the algebraic formula (10) rather than GP's analytical arguments. Our approach is more flexible about the consumption domain: X is convex, but need not be closed, or bounded, or finite dimensional. For example, X may consist of monetary gambles with payoffs on the real line, or X can be the non-negative orthant of consumption bundles. We use such specifications later to discuss random risk attitudes and random Cobb-Douglas utility.

GP's REU implies Additivity. By (2), for all $(x, A) \in \Omega$, $B \in \mathcal{M}$, and $\gamma \in (0, 1)$.

$$\mathcal{C}_x(A) = \bigcup_{y \in B} \mathcal{C}_{\gamma x+(1-\gamma)y}(\gamma A + (1-\gamma)B),$$

but this union need not be disjoint because rational elements are not unique for mixture continuous types. Additivity still follows from representation (13) that applies μ to both sides of the above union and assumes the no-tie constraint (14).

On the other hand, our DREU implies all four conditions in GP's Theorem 2. Monotonicity has been shown. A special case of of Additivity where $B = y$ is

Axiom 6 (Linearity)**.** $\rho(x, A) = \rho(x, A, \gamma, y)$.

The constant function $\rho(x, A, \gamma, B)$ is obviously continuous.

Axiom 7 (Mixture Continuity (MC))**.** $\rho(x, A, \gamma, B)$ *is continuous over* $\gamma \in (0, 1]$.

Note that $\gamma x + (1-\gamma)y$ is never rational for a total type when x and y are also feasible. Thus, DREU implies

Axiom 8 (Extremeness)**.** *For all* $\gamma \in (0, 1)$*,*

$$\rho\left(\gamma x + (1-\gamma)y, \{x, y, \gamma x + (1-\gamma)y\}\right) = 0.$$

Therefore, if $X = \Delta^n$, and Discreteness holds, then the lists of axioms in our Theorem 2 and GP's Theorem 2 become equivalent. By contrast, their Theorem 3 requires that μ in (13) is countably additive, and ρ satisfies a continuity axiom that is stronger than mixture continuity. This assumption is incompatible with Discreteness, except when $X = \Delta^2$.

More generally, Additivity can be derived from GP's axioms even if X is not finite dimensional, but proving this claim requires more effort than proving Theorem 2 from scratch and appears tangential to our desiderata. So we skip it. Besides making the proofs and the statement of Theorem 2 more parsimonious, Additivity is also conveniently relaxed to BA and AA in Theorems 3 and 4.

Note also that Additivity can be derived from its subadditivity part: for all $(x, A) \in \Omega$, $B \in \mathcal{M}$, and $\gamma \in (0, 1)$,

$$\rho(x, A) \geq \rho(x, A, \gamma, B).$$

Indeed, if any of these inequalities is strict, then by (7),

$$1 = \sum_{x \in A} \rho(x, A) > \sum_{x \in A} \sum_{y \in B} \rho(\gamma x + (1 - \gamma)y, \gamma A + (1 - \gamma)B) \geq 1.$$

4.1 Self-Selection and Random Utility

Assume a selective environment as in condition (12). With the addition of x to a menu A, agents that participated in A would either make the same choice as in A if x is not the best in $A \cup x$, or switch to x otherwise. This implies that no agent would switch to a different option in A. Agents that did not participate in A, but are participating in $A \cup x$ would only choose x. Therefore, no pre-existing option in A would be chosen with a higher probability, which implies that self-selection does not violate Monotonicity.

In fact, Block-Marschak polynomials constructed in such an environment would be nonnegative. Using representation 11, a Block-Marschak Polynomial—interpreted as the probability that x is the best in A but worse than all items in $X \setminus A$—can be expressed as:

$$q(x, A) \equiv \sum_{A \subseteq S \subseteq X} (-1)^{|S \setminus A|} \mu_S(\mathcal{T}_x(S)) = \sum_{B \subseteq A \setminus \{x\}} (-1)^{|B|} \mu_{A \setminus B}(\mathcal{T}_x(A \setminus B)). \tag{15}$$

Let $\theta \in \mathcal{T}_x(A)$ be a type that finds x maximizing in A and let $B \subseteq A \backslash \{x\}$. Since $A \backslash B \subseteq A$, it must be that $\theta \in \mathcal{T}_x(A \backslash B)$ as only worse items are removed for type θ. The selective property ensures that $\mu_{A \backslash B}(\theta) \geq \mu_A(\theta)$ for any of such θ. Therefore, the set function $M \to \mu_M(\theta)$ is completely monotone decreasing over the lattice of subsets, in the sense that adding items weakly decreases the measure. The Block-Marshak polynomial is exactly the Möbius transform of $\mu_A(\mathcal{T}_x(A))$ over the Boolean lattice of subsets (Billot and Thisse, 2005), and complete monotonicity guarantees nonnegative Möbius coefficients (Grabisch and Miranda, 2015). Hence, by the properties of Möbius inversion, all Block-Marschak polynomials are nonnegative. Falmagne (1978) stated that such a condition is equivalent to having a random utility representation.

Consider the following example. Let the world consist of two types: $a\theta b\theta c$ and $c\psi b\psi a$, each taking up half of the true population. Assume a type would participate for sure if their favorite option is available, and with probability 50% otherwise. In menu $A = \{a, b\}$, all type θ agents will show up whereas half of type ψ agents will show up. θ agents will choose a and ψ agents will choose b, and the market will observe a stochastic choice rule of $\rho(a, A) = \frac{2}{3}$

and $\rho(b, A) = \frac{1}{3}$.[2]

However, researchers only observe the stochastic choice rule, and are not privy to the true distribution of types in the world. The following distribution would also provide the same ρ:[3]

$$\hat{\mu}(\theta) = \frac{5}{9},\ \ \hat{\mu}(\psi) = \frac{1}{3},\ \ \hat{\mu}(\xi) = \frac{1}{9}$$

where $b\xi c\xi a$. This is due to the non-uniqueness of Block-Marschak polynomials (Fishburn, 1998). As a result, the distribution of types can be misspecified in a random utility model where there is no parametric form, i.e., where agents preferences are arbitrary orders on the consumption space.

4.2 Random Risk Attitudes

Olschewski et al. (2022) discovered that the random utility model does not precisely differentiate between risk-averse and risk-seeking agents. This motivates us to next extend the discussion of menu dependence to risk attitudes.

First, if X is interpreted as a set of monetary lotteries, then the discrete type space Θ can capture heterogeneity of risk attitudes and/or beliefs over an exogenous state space S.

Payoffs $q \in Q$ and lotteries $x \in X$ are called *monetary* if $Q \subseteq \mathbb{R}$ consists of monetary rewards. In this case, let $\mathcal{T}_+ \subseteq \mathcal{T}$ bet the set of all types such that $q > r$ implies $q\theta r$ for all $q, r \in Q$. Each type $\theta \in \mathcal{T}_+$ can be interpreted in terms of the *risk attitude* that is revealed by its increasing utility index u_θ, and some endogenous tie-breaking rule.

If $Q \subseteq \mathbb{R}$ is monetary, then less money should never be chosen over more money.

Axiom 9 (No Free Lunch (NFL)). *For all $q, r \in Q$, if $q > r$, then $\rho(r, \{q, r\}) = 0$.*

An easy corollary of Theorem 3 is that ρ satisfies BA, AA, Discreteness and NFL if and only if ρ is represented by DREU with a discrete type space $\Theta \subseteq \mathcal{T}_+$. Indeed, NFL is obvious if $\Theta \subseteq \mathcal{T}_+$. Conversely, suppose that Φ satisfies these axioms. By Theorem 3, Φ is represented by some non-redundant pair (Θ, μ). Suppose that there is $\theta \in \Theta$ that does not belong to $\mathcal{T}_+$. Then $r\theta q$ for some $q > r$. By DREU, $\rho(r, \{q, r\}) > 0$, which contradicts NFL.

Peng (2025) explores the effects of a certain option (degenerate lottery) on risk behavior in an experimental setting. Let there be two types: a risk averse type θ with a Bernoulli

[2] $\rho(a, A) = \dfrac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2} \times \frac{1}{2}}, \rho(b, A) = \dfrac{\frac{1}{2} \times \frac{1}{2}}{\frac{1}{2} + \frac{1}{2} \times \frac{1}{2}}$

[3] $\rho(a, A) = \dfrac{\frac{5}{9}}{\frac{5}{9} + \frac{1}{3} \times \frac{1}{2} + \frac{1}{9}} = \dfrac{2}{3}, \rho(b, A) = \dfrac{\frac{1}{3} \times \frac{1}{2} + \frac{1}{9}}{\frac{5}{9} + \frac{1}{3} \times \frac{1}{2} + \frac{1}{9}} = \dfrac{1}{3}.$

utility index of $u_\theta(x) = \sqrt{x}$ and a risk neutral type n that always chooses the option with a higher expected value. Assume the following observations on two menus:

$$A = \left\{ r \equiv \begin{pmatrix} \$10 & \text{w.p.} & 0.5 \\ \$0 & \text{w.p.} & 0.5 \end{pmatrix}, l \equiv \begin{pmatrix} \$8 & \text{w.p.} & 0.5 \\ \$1 & \text{w.p.} & 0.5 \end{pmatrix} \right\} \text{ and } A' = A \cup \{c \equiv (\$3 \text{ w.p. } 1)\}.$$

A risk averse type will choose l in both cases whereas a risk neutral type will chose r. Let the data show that $\rho(r, A) = \rho(l, A) = 0.5$ but $\rho(r, A') = 0.25$ and $\rho(l, A') = 0.75$. This would suggest that in menu A, the population consists of half risk neutral agents and half risk averse agents. However, in menu A', risk neutral agents only takes up a quarter of the population while the rest are risk averse. This is a violation of Monotonicity as l is chosen with a higher likelihood in a superset menu. Hence, GP's standard REU would fail to model this phenomenon, and menu-dependence is needed.

We can explain this phenomenon by treating the certain option as a decoy, where its presence will make some agents more risk averse. Another working paper with a similar design (Lim and R., 2021) confirms that adding a safer option increases risk aversion. They also observed that adding a riskier option does not have a noticeable change in risk attitude. The latter was also noticed by Chen et al. (2024).

This can be extended to more types. Specifically, we can use discrete Arrow-Pratt indices by Baillon and L'Haridon (2021) to capture the finite nature of the type space. A discrete lognormal distribution may be appropriate for $\mu(\cdot, A)$ as CRRA, in the format of $u(x) = \frac{x^r}{1-r}$, indices are positive and skewed right, with most agents exhibiting some level of risk aversion. The mode of $\mu(\cdot, A)$ will move left (towards 0) when a certain option is present.

4.3 Random Cobb-Douglas

Representations from Theorems can be adapted to characterize random Cobb-Douglas utility in the setting where $X = \mathbb{R}^n_+$ is the orthant of consumption bundles.

4.4 Tie-Breakers and Non-Redundant Types

First, we clarify the structure of total types that appear in DREU.

For any $u \in \mathcal{U}$, let

$$L_u = \{\alpha(x - y) : \alpha \geq 0 \text{ and } x, y \in X \text{ such that } u(x) = u(y)\}.$$

As X is convex, then L_u is a linear space.[4]

[4]Indeed, take any $\alpha \geq 0$ and $x, y \in X$ such that $u(x) = u(y)$. Then $\gamma\alpha(x - y) \in L_u$ for all $\gamma \in \mathbb{R}$ because

Theorem 5. *A type θ has a non-null closure θ^* if and only if there is a function $u \in \mathcal{U}$ and a type ψ in L_u such that for all $x, y \in X$,*

$$x\theta y \iff \begin{array}{l} u(x) > u(y), \text{ or} \\ u(x) = u(y) \text{ and } (x-y)\psi 0. \end{array} \tag{16}$$

If (16) *holds with another pair (u', ψ'), then $u' = \alpha u + \beta$ for some $\alpha > 0$ and $\beta \in \mathbb{R}$, $L_u = L_{u'}$, and $\psi = \psi'$.*

The proof is in the appendix. Representation (16) asserts that any stable type θ maximizes some utility function $u \in \mathcal{U}$, which is unique up a positive linear transformation, and then uses the unique residual type ψ on the linear subspace L_u to compare any $x, y \in X$ such that $u(x) = u(y)$.

Representation (16) in Theorem 5 implies that for any stable type θ,

$$\theta \text{ is total in } X \iff \psi \text{ is total in } L_u.$$

Indeed, if $a \sim_\psi b$ for some distinct elements $a, b \in L_u$, then by Independence, $\frac{a-b}{2} \sim_\psi 0$. Take $x, y \in X$ and $\alpha \geq 0$ such that $\alpha(x-y) = \frac{a-b}{2} \neq 0$ and $u(x) = u(y)$. By Independence, $x - y \sim_\psi 0$. By (16), $x \sim_\theta y$, and θ is not total.

For any $u \in \mathcal{U}$, let $\mathcal{T}_u$ be the set of all total types on L_u. Such types are called *tie-breakers*. To construct tie-breakers, take any *basis* H in the linear space L_u. By definition, $H \subset L_u$ is a linearly independent system of vectors that spans L_u. This definition is standard if L_u is finite dimensional. More generally, such H is known as a *Hamel basis* when L_u is infinite dimensional. The existence of a Hamel basis in any linear space is guaranteed by Zorn's lemma (see Aliprantis and Border (1999, Theorem 1.6)). For any $x \in L_u$ and $h \in H$, let $x(h)$ be the coefficient (coordinate) of the vector h in the unique linear combination of basis vectors that equals x. Endow H with any total order $\geq$. Define a binary relation ψ_H for all $x, y \in L_u$ via

$$x\psi_H y \iff x = y \text{ or } x(h(x,y)) \geq y(h(x,y)), \tag{17}$$

where $h(x,y) \in H$ is the first vector in H such that $x(h) \neq y(h)$. Then ψ_H compares the coordinates of the vectors x and y *lexicographically* in the basis H. By definition, ψ_H is a tie-breaker: it is complete, total, transitive, and satisfies Independence.

---

if $\gamma < 0$, then $\gamma\alpha(x-y) = (-\gamma\alpha)(y-x)$. For any $\beta \geq 0$ and $x', y' \in X$ such that $u(x') = u(y')$, the sum

$$\alpha(x-y) + \beta(x'-y') = (\alpha+\beta)[\left(\tfrac{\alpha}{\alpha+\beta}x + \tfrac{\beta}{\alpha+\beta}x'\right) - \left(\tfrac{\alpha}{\alpha+\beta}y + \tfrac{\beta}{\alpha+\beta}y'\right)$$

also belongs to L_u because $u\left(\frac{\alpha}{\alpha+\beta}x + \frac{\beta}{\alpha+\beta}x'\right) = u\left(\frac{\alpha}{\alpha+\beta}y + \frac{\beta}{\alpha+\beta}y'\right)$.

Thus a rich family of total types can be generated by representations (16) and (17). These representations can be further refined (e.g. Fishburn (1982), Myerson (1986), Blume, Branderburger, and Dekel (1991)) by imposing more structure on the consumption domain X, the functions $u \in \mathcal{U}$, and the basis H.

Next, Discreteness can be strengthened to interpret the support Θ of the discrete map μ in terms of distinct expected utility functions.

Take any non-singleton menu A. Say that $x \in A$ is a *stable choice* in A if for any $y \in A \backslash x$ and $z_x, z_y \in X$, there is $\varepsilon > 0$ such that for all $\gamma \in [0, \varepsilon)$

$$\rho(\gamma z_x + (1-\gamma)x, \{\gamma z_x + (1-\gamma)x, \gamma z_y + (1-\gamma)y\} \cup (A \setminus x)\} > 0$$

Let $s_\rho(A)$ be the number of stable choices in A. If $\gamma = 0$, then the definition of a stable choice implies $\rho(x, A) > 0$. Thus $s_\rho(A) \leq k_\rho(A)$.

Axiom 10 (Stable Discreteness (SD))**.** *There is $A \in \mathcal{M}$ such that $s_\rho(A) \geq k_\rho(B)$ for all $B \in \mathcal{M}$.*

Clearly, SD implies Discreteness for $n = s_\rho(A)$. SD requires that the maximal number of chosen elements in an arbitrary menu B can be achieved via stable choices in some A.

A finite collection of total types $\Theta \subset \mathcal{T}$ is called *non-redundant* if for all types $\theta_1, \theta_2 \in \Theta$, their closures θ_1^* and θ_2^* are non-null, and

$$\theta_1^* = \theta_2^* \implies \theta_1 = \theta_2.$$

In other words, non-redundancy requires that distinct types should maximize distinct utility functions $u_1, u_2 \in \mathcal{U}$ rather than differ only in tie-breaking rules.

Theorem 6. *If ρ is represented by a discrete map $\mu \in \mathcal{M}(\mathcal{D})$, then the support Θ is non-redundant if and only if ρ satisfies SD.*

A APPENDIX: PROOFS

First, we show Theorem 3. Theorems 2 and 4 are then derived as corollaries.

Lemma A.1. *For any finite set $\Theta = \{\theta_1, \ldots, \theta_k\} \subset \mathcal{T}$, there is a menu $B = \{b_1, \ldots, b_k\}$ such that for all $i = 1, \ldots, k$,*

$$b_i = R_{\theta_i}(B).$$

Proof. Let I be the set of all pairs $i, j \in \{1, \dots, k\}$ such that $i > j$. For each pair $(i, j) \in I$, the types $\theta_i, \theta_j \in \mathcal{D}$ are distinct and total. Thus there are distinct elements $x_{ij}, x_{ji} \in X$ such that $x_{ij}\theta_i x_j$ and $x_{ji}\theta_j x_i$. For any $n \in \{1, \dots, k\}$, let

$$b_n = \sum_{(i,j)\in I} \tfrac{2}{k(k-1)} x^n_{ij},$$

where $x^n_{ij} = x_{ij}$ if $x_{ij}\theta_n x_{ji}$ and $x^n_{ij} = x_{ji}$ if $x_{ji}\theta_n x_{ij}$. For any $(m, n) \in I$, the types θ_n, θ_m satisfy Independence and hence,

$$\begin{aligned} b_m &= \left(\sum_{(i,j)\in I} \tfrac{2}{k(k-1)} x^m_{ij}\right) \theta_m \left(\sum_{(i,j)\in I} \tfrac{2}{k(k-1)} x^n_{ij}\right) = b_n \\ b_n &= \left(\sum_{(i,j)\in I} \tfrac{2}{k(k-1)} x^n_{ij}\right) \theta_n \left(\sum_{(i,j)\in I} \tfrac{2}{k(k-1)} x^m_{ij}\right) = b_m \end{aligned} \tag{18}$$

because $x^n_{ij}\theta_n x^m_{ij}$ and $x^m_{ij}\theta_m x^n_{ij}$ for all $(i, j) \in I$. Moreover, when $i = m$ and $j = n$, then these comparisons are strict, and hence b_m, b_n are distinct elements.

Let $B = \{b_1, \dots, b_k\}$. By (18), $b_i\theta_i b_j$ for all i, j and hence, $b_i = R_{\theta_i}(B)$. □

Suppose that ρ is represented by a discrete map μ. Let $\Theta \subset \mathcal{T}$ be the support of all distributions μ_A contingent on menus A.

Check BA. By (11) and (2), for all $(x, A) \in \Omega$, $B \in \mathcal{M}$ and $\gamma \in (0, 1)$,

$$\begin{gathered} \rho(x, A) = 0 \iff x \neq R_\theta(A) \text{ for all } \theta \in \Theta \iff \\ \gamma x + (1-\gamma)y \neq R_\theta(A + (1-\gamma)B) \text{ for all } \theta \in \Theta \text{ and } y \in B \iff \\ \rho(x, A, \gamma, B) = 0. \end{gathered}$$

Check AA. Take any $(x, A) \in \Omega$ and converging sequences $\{\gamma_m\}_{m=1}^\infty \in (0, 1]$ and $\{B_n\}_{n=1}^\infty \in \mathcal{M}$. Let $\gamma = \lim_{m\to\infty} \gamma_m$ and $B = \lim_{n\to\infty} B_n$. By (11),

$$\rho(x, A) > 0 \implies \mu(\mathcal{T}_x(A)) > 0 \implies \mu_{\gamma A+(1-\gamma)B}(\mathcal{T}_x(A)) > 0.$$

The Hausdorff metric implies that

$$\lim_{m,n\to\infty} d_h(\gamma_m A + (1-\gamma_m)B_n, \gamma A + (1-\gamma)B) = 0$$

As μ is continuous, then

$$\lim_{m,n\to\infty} \rho(x, A, \gamma_m, B_n) = \lim_{m,n\to\infty} \mu_{\gamma_m A+(1-\gamma_m)B_n}(\mathcal{T}_x(A)) = \mu_{\gamma A+(1-\gamma)B}(\mathcal{T}_x(A)) > 0.$$

Conversely, suppose that ρ satisfies Discreteness, BA, and AA. For all menus $A \in \mathcal{M}$, let

$$\Phi_\rho(A) = \{x \in A : \rho(x, A) > 0\}.$$

Show that for all $A \in \mathcal{M}$,

$$\Phi_\rho(A) = \bigcup_{\theta\in\Theta} R_\theta(A) \text{ for all } A \in \mathcal{M}. \tag{19}$$

Thus $\Phi_\rho(A)$ consists of all elements that are rational for at least one type $\theta \in \Theta$ in the menu A. Representation refines the *multiple utility* model proposed by Aizerman and Malishevski Aizerman and Malishevski (1981) by restricting the type space Θ to $\mathcal{T}$ rather than to arbitrary total orders.

Take $B \in \mathcal{M}$ such that $|\Phi(B)| = k$. Let $C = \Phi(B) = \{c_1, \ldots, c_k\}$. By Chernoff, $\Phi(C) = C$ because $C \subset B$. Let $R = \max_{i=1,\ldots,k} ||x_i||$. Take $\alpha \in (0, 1)$ such that

$$2\alpha R < (1 - \alpha) \min_{1\leq i<j\leq k} ||x_i - x_j||.$$

Then for any $\alpha \in (0, \alpha_{xy})$, the menu $\alpha\{x, y\} + (1 - \alpha)B$ has $2k$ distinct elements. To show this, it is enough to check that for any distinct $i, j \in \{1, \ldots, k\}$, the menu

$$\alpha\{x, y\} + (1 - \alpha)\{x_i, x_j\}$$

has four distinct elements. Here the inequalities

$$\alpha x + (1 - \alpha)x_i \neq \alpha y + (1 - \alpha)x_i \neq \alpha y + (1 - \alpha)x_j \neq \alpha x + (1 - \alpha)x_j \neq \alpha x + (1 - \alpha)x_i$$

are obvious, the other two inequalities $\alpha x+(1-\alpha)x_i \neq \alpha y+(1-\alpha)x_j$ and $\alpha x+(1-\alpha)x_j \neq \alpha y + (1 - \alpha)x_i$ follow from

$$(1 - \alpha)||x_i - x_j|| - \alpha||x - y|| > 0.$$

As $x_i \in \Phi(B)$ for all i, then by Type Independence, $\Phi(\alpha\{x, y\} + (1 - \alpha)B)$ overlaps with each of the k distinct sets

$$A_i = \alpha\{x, y\} + (1 - \alpha)x_i.$$

By (**??**), each of these k overlaps contains precisely one element, $\alpha x + (1-\alpha)x_i$ or $\alpha y + (1-\alpha)x_i$.

A.1 Proof of Theorem 5

Next, we will show the tie-breaking Theorem 5 and derive Theorem 1 as its corollary.

Prove Theorem 5. In this proof, write any type θ as a preference relation $\succeq$ with asymmetric and symmetric parts $\succ$ and $\sim$ respectively. Suppose that $\succeq$ is represented by (16) for some $u \in \mathcal{U}$ and type ψ in L_u. Completeness is obvious. To show transitivity, take any $x, y, z \in X$ such that $x \succeq y \succeq z$. Then $u(x) \geq u(y) \geq u(z)$. If $u(x) > u(z)$, then $x \succeq z$ by (16). If $u(x) = u(z)$, then $x\psi y\psi z$ and hence, $x\psi z$. By (16), $x \succeq z$. Next, θ satisfies Independence because u is mixture linear, and the tie-breaker ψ obeys Independence. Stability has been already derived from representation (6) and a fortiori, from (16).

Conversely, take any stable type $\succeq$ and derive representation (16). Take x, y that satisfy the stability condition (5). Then $x \succ y$ because $y \succeq x$ would imply by Independence that for any $z_x \prec z_y$ and $\alpha \in (0,1)$, $\alpha z_x + (1-\alpha)x \prec \alpha z_y + (1-\alpha)y$. Thus (5) would not hold.

Take any menu $A = \{a_1, a_2, \ldots, a_n\}$ and index so that $a_1 \succeq a_2 \succeq \cdots \succeq a_n$. Let Z be the convex hull of A in L.

Show that for all $z \in Z$,

$$a_1 \succeq z \succeq a_n. \tag{20}$$

The proof is by induction with respect to $n \geq 2$. If $n = 2$, then $z = \gamma a_1 + (1-\gamma)a_2$ for some $\gamma \in (0,1)$, and (20) follows from $a_1 \succ a_2$ by Independence. Let $n > 2$ and Z' be the convex hull of $A \setminus a_n$. Take any $z \in X$ and let $\gamma = z(a_n)$. Then $z = \gamma a_n + (1-\gamma)y$ for some $y \in X'$. By the inductive assumption, $a_1 \succeq y \succeq a_{n-1} \succeq a_n$. By Independence,

$$a_1 \succeq \gamma a_n + (1-\gamma)a_1 \succeq \gamma a_n + (1-\gamma)y \succeq \gamma a_n + (1-\gamma)a_n = a_n$$

and hence, (20) holds.

For each $\alpha \in [0,1]$, let $z_\alpha = \alpha a_1 + (1-\alpha)a_n$. Show that for all $\alpha, \beta \in [0,1]$,

$$\alpha \geq \beta \iff z_\alpha \succeq z_\beta. \tag{21}$$

By Independence,

$$\alpha > \beta \implies z_\alpha = \alpha a_1 + (1-\alpha)a_n \succ a_n \implies$$
$$z_\alpha = \tfrac{\beta}{\alpha} z_\alpha + \tfrac{\alpha-\beta}{\alpha} z_\alpha \succ \tfrac{\beta}{\alpha} z_\alpha + \tfrac{\alpha-\beta}{\alpha} a_n = z_\beta.$$

If $\alpha \geq \beta$, then either $z_\alpha = z_\beta$ or $z_\alpha \succ z_\beta$. Thus (21) holds. For all $x \in X$, let

$$u_\theta(x) = \inf\{\alpha \in [0,1] : z_\alpha \succeq x\}$$
$$v_\theta(x) = \sup\{\beta \in [0,1] : x \succeq z_\beta\}.$$

By (20), these values are well-defined. Show that $u_\theta = v_\theta$. Indeed, $z_\alpha \succeq x \succeq z_\beta$ implies $\alpha \geq \beta$ and hence, $u_\theta(x) \geq v_\theta(x)$. Suppose that $u_\theta(x) > v_\theta(x)$. Then there is γ such that $u_\theta(x) > \gamma > v_\theta(x)$. By definition of u_θ, $u_\theta(x) > \gamma$ implies $x \succ z_\gamma$. By definition of v_θ, $\gamma > v_\theta(x)$ implies $z_\gamma \succ x$. By contradiction, $u_\theta(x) = v_\theta(x)$.

Show that u_θ satisfies the mixture linearity (**??**). By Independence, for any $x, y \in X$ and $\alpha, \beta, \gamma \in [0,1]$ such that $z_\alpha \succeq x$ and $z_\beta \succeq y$,

$$z_{\gamma\alpha+(1-\gamma)\beta} = \gamma z_\alpha + (1-\gamma) z_\beta \succeq \gamma x + (1-\gamma) z_\beta \succeq \gamma x + (1-\gamma) y.$$

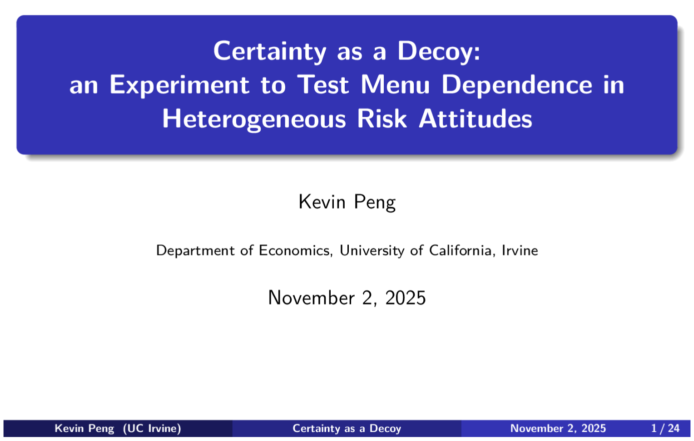
Thus $u_\theta(\gamma x + (1-\gamma)y) \leq \gamma\alpha + (1-\gamma)\beta$. By definition of u_θ, α and β can be arbitrarily close to $u_\theta(x)$ and $u_\theta(y)$ respectively. Thus

$$u_\theta(\gamma x + (1-\gamma)y) \leq \gamma u_\theta(x) + (1-\gamma) u_\theta(y).$$

Analogously, $v_\theta(\gamma x + (1-\gamma)y) \geq \gamma v_\theta(x) + (1-\gamma) v_\theta(y)$. As $u_\theta = v_\theta$, then u_θ is mixture linear.

Suppose that x does not maximize u_θ in A. Take $y \in A$ such that $u_\theta(y) > u_\theta(x)$. Take γ such that $u_\theta(y) = v_\theta(y) > \gamma > u_\theta(x)$. Then $y \succ z_\gamma \succ x$ and hence, x does not maximize $\succeq$ in A. Thus $\Phi_\theta(A) \in \Phi_{u_\theta}(A)$.

B Select Slides for Experiment



Research Question and Contribution

On the aggregate level, *the distribution of risk attitudes can vary depending on what options are presented to the decision makers.*

- *Harrison, et al., 12*: individuals are risk neutral when lotteries are presented in a binary format $L = \begin{pmatrix} \$x & \text{w.p. } \pi \\ \$y & \text{w.p. } 1-\pi \end{pmatrix}$.
- *Hermanns & Kokot, 23*: framing can affect levels of risk aversion.

Question: does the presence of a certain option in the menu make agents on average more risk averse?

- Most literature focuses on **framing** effects (*Crosta, et al., 23*).
- *Chen, et al., 24*: introducing a high-risk-high-reward lottery does not decrease risk aversion.

$\rightarrow$ I designed the menus so a no-risk-low-reward is introduced.

$\rightarrow$ **Risk aversion is observed to increase**.

Monotonicity

When additional options are added to a menu, the probability of choosing a **pre-existing** option should **not increase**: for all $x \in A$, A, y,

$$\rho(x, A) \geq \rho(x, A \cup \{y\})$$

- A stochastic choice rule that has a random utility representation is **monotone**.
- However, if agents become more risk averse with the expansion of a menu, *they may switch to a pre-existing option that is of lower risk.*
- **Goal**: violations of monotonicity in some observations.

$\Longrightarrow$ a menu-dependent version of random utility is needed to model the choice data.

Hypotheses

Introduce a no-risk-low-reward (certain) option:

$$A \equiv \{\text{low-risk } (l), \text{high-risk } (h)\} \to \{l, h, \text{certain } (c)\} \equiv A'.$$

***H1*:** the proportion of agents that choose h will decrease.
***H2*:** the proportion of agents that choose l will increase.
***H3*:** the relative proportion of agents that chose l will increase over h:

$$\rho(l, A) \leq \frac{\rho(l, A')}{\rho(l, A') + \rho(h, A')}.$$

Note that $H2 \implies H3 \implies H1$.

→ **Decoy effects** and **extemeness aversion** suggest at least *H1*.

***H4*:** exposure to c has an extended effect, even after it is removed.

→ *Hedgcock, Rao & Chen, 16*: decoy effects can last beyond removal.

→ *Pettibone & Wedell, 07*: phantom decoys (not selectable).

Menu Without a Certain Option

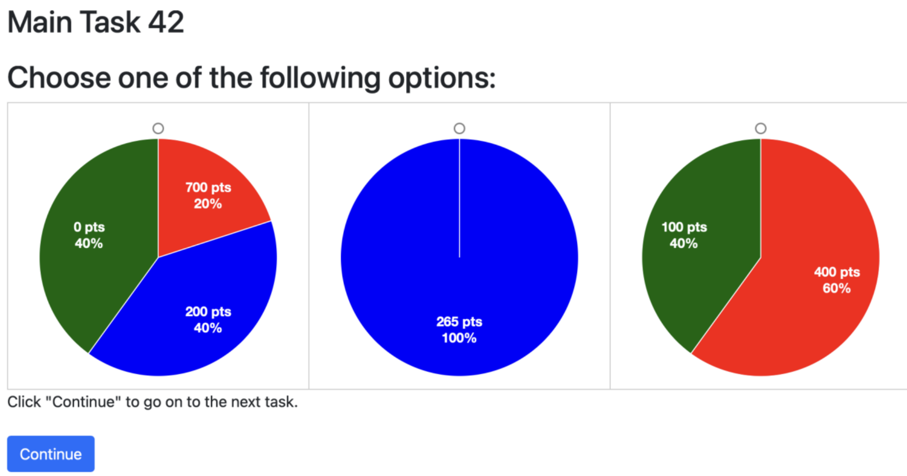
Main Task 2

Choose one of the following options:

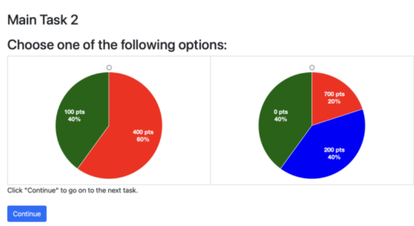
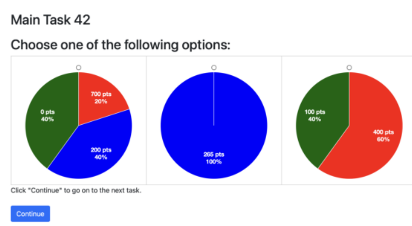
Click "Continue" to go on to the next task.

Continue

Menu With a Certain Option



Main Tasks and Treatments

Menu A	Menu A'
	

- Each Menu A has a corresponding Menu $A' \supseteq A$.
- There are 26 pairs, with 52 total main tasks.

***T1*:** Subjects face the 26 A menus first, and then move on to 26 A' menus.

***T2*:** Subjects face the 26 A' menus first, and then move on to 26 A menus.

***T3*:** Subjects face alternating menus $A \to A' \to A \to \cdots$.

Pilot Study

23 test subjects (UCI economics graduate students).
48 choices each, with a total of 1104 observations.

	high-risk (h) option		low-risk (l) option		certain
(#)	no c	with c	no c	with c	option
T1 (8)	95 (49%)	77 (40%)	97 (51%)	84 (43%)	31 (16%)
T2 (7)	77 (46%)	58 (35%)	91 (54%)	102 (61%)	8 (5%)
T3 (8)	72 (38%)	70 (36%)	120 (62%)	92 (48%)	30 (16%)
Pool	244 (44%)	205 (37%)	308 (56%)	278 (50%)	69 (13%)
Cond.	44%	43%	56%	57%	

- *H1* (decrease in high-risk choice) holds here.

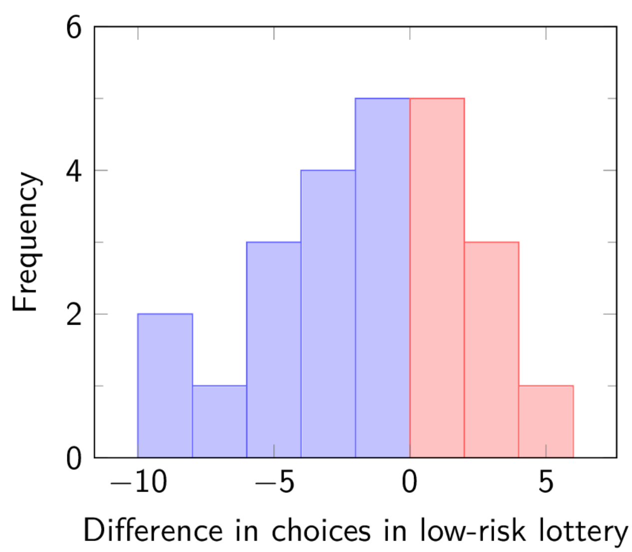
$\rightarrow$ On average, agents become more risk averse with the presence of c.

- *H3* does not seem statistically significant. (*H2* does not hold.)

$\rightarrow$ Some people switch from l to c.

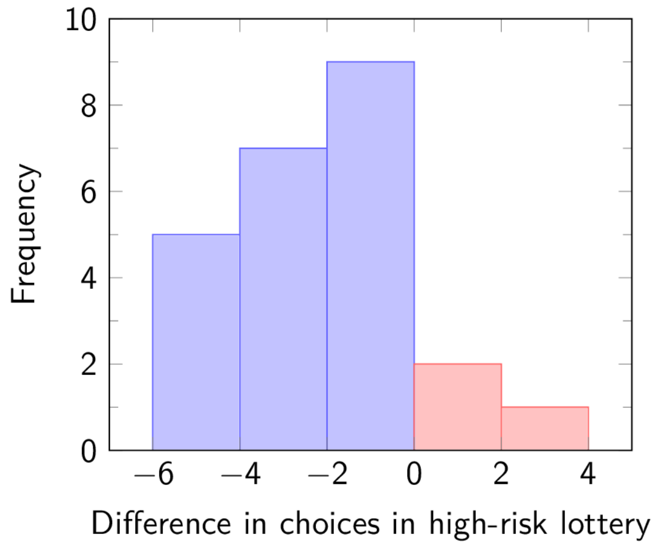
Monotonicity: Low-Risk Option

When separately discussing each menu pair, monotonicity is violated in 9 of the 24 pairs when comparing the number of people that chose the low-risk option with or without the certain option (c).



Monotonicity: High-Risk Option

Monotonicity is violated in 3 of the 24 pairs when comparing the number of people that chose the *high*-risk option with or without the certain option.



Logistic Regression with Random Effects

$$\log\left(\frac{P(Y_{ij}=1)}{1-P(Y_{ij}=1)}\right) = \beta_0 + \beta_1 threeHigh_{ij} + \beta_2 threeLow_{ij} + \beta_3 switch_{ij} + \beta_4 certain_{ij} + b_i$$

- Y_{ij}: whether subject i chooses the riskier lottery in task j.
- $threeHigh_{ij}$: whether the high-risk lottery in the task has three states.
- $threeLow_{ij}$: whether the low-risk lottery in the task has three states.
- $switch_{ij}$: the level of risk aversion (category) that would switch from high- to low-risk.
- $certain_{ij}$: whether a certain lottery is present in the task.
- $b_i \sim \mathcal{N}(0, \sigma_b^2)$: random intercept for subject i.

Estimation

	Estimate	Std. Error	z value	Pr($>|z|$)
(Intercept)	-3.562	0.319	-11.150	$<$ 2e-16***
threeHigh	-0.165	0.140	-1.174	0.240
threeLow	0.049	0.140	0.353	0.724
switch	0.603	0.046	13.155	$<$ 2e-16***
certain	-0.382	0.141	-2.719	0.007**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 1: Fixed effect estimates from mixed-effects logistic regression predicting whether a subject will choose the high-risk lottery.

Persistent Effect of Exposure (*H4*)

There should be higher level of risk aversion in the A portion (no c) for subjects faced with A' (with c) first (*T2*) than those with A first (*T1*).

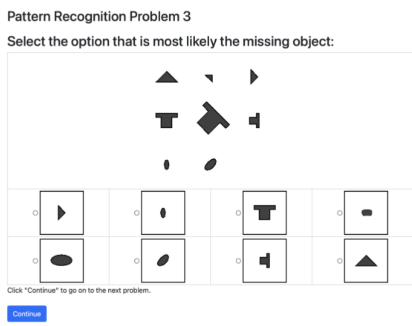
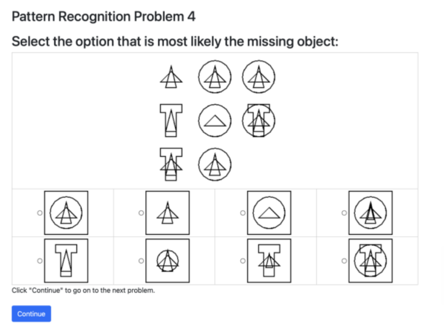
	high-risk option		low-risk option		certain
	no c	with c	no c	with c	option
T1	95 (**49%**)	77 (40%)	97 (51%)	84 (43%)	31 (16%)
T2	77 (**46%**)	58 (35%)	91 (54%)	102 (61%)	8 (5%)

- *T3*'s level of risk aversion should increase over time. The starting level should be similar to *T1* and end similar to *T2*.

$\rightarrow$ Plan: compare first 10 tasks with last 10 tasks.

Cognitive Test: Pattern Recognition

Matzen, et al., 10: a non-branded version of Raven's Progressive Matrices.

Patterns	Logic gates
	

- 8 problems with progressive difficulty.
- Plan: find relationship between cognition and menu-dependent risk attitudes. Cognizant agents less likely to be affected by the decoy?

Conclusion

Findings in the pilot study:

- Agents are on average **more risk averse** when a certain option is presented.
- *Monotonicity* is **violated** in a third of the scenarios.

To dos for the main study (around 270 subjects, to be collected in ESSL):

- Observe what types of menus are more likely to violate monotonicity.
- Analysis on persistence of decoy effects (*H4*).
- Relate cognition to sensitivity to decoy effects.

References

M. Agranov and P. Ortoleva. Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125:40–68, 2017.

D. S. Ahn and H. Ergin. Framing contingencies. 78(2):655–695, 2010.

M. Aizerman and A. Malishevski. General theory of best variants choice. *IEEE Trans. Automat. Control*, pages 1030–1041, 1981.

C. Aliprantis and K. Border. *Infinite Dimensional Analysis*. Springer, 1999.

F. Anscombe and R. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.

J. Apesteguia, M. Ballester, and J. Lu. Single-crossing random utility models. *Econometrica*, 85:661–674, 2017.

A. Baillon and O. L'Haridon. Discrete arrow-pratt indexes for risk and uncertainty. *Economic Theory*, 72(4):1375–1393, 2021.

A. Billot and J. F. Thisse. Stochastic rationality and möbius inverse. *International Journal of Economic Theory*, (3):211–217, 2005.

H. Block and J. Marschak. Random orderings and stochastic theories of response. In Olkin, Ghurye, Hoeffding, Madow, and Mann, editors, *Contributions to Probability and Statistics*, pages 97–132. Stanford University Press, Palo Alto, 1960.

L. Blume, A. Branderburger, and E. Dekel. Lexicographic probabilities and choice under uncertainty. *Econometrica*, pages 61–79, 1991.

Z. Chen, R. Golman, and J. Somerville. Menu-dependent risk attitudes: Theory and evidence. *Journal of Risk and Uncertainty*, 68(1):77–105, 2024.

A. Daly and S. Zachery. Improved multiple choice models. In D. Hensher and Q. Dalvi, editors, *Identifying and Measuring the Determinants of Mode Choice*. London: Teakfield, 1979.

J. Falmagne. A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, 18:52–72, 1978.

P. C. Fishburn. *The Foundations of Expected Utility*. D. Reidel Publishing Company, Dordrecht, 1982.

P. C. Fishburn. Stochastic utility. In S. Barbera, P. J. Hammond, and C. Seidl, editors, *Handbook of Utility Theory*. Springer, 1998.

E. E. Garnett, A. Balmford, C. Sandbrook, M. A. Pilling, and T. M. Marteau. Impact of increasing vegetarian availability on meal selection and sales in cafeterias. *Proceedings of the National Academy of Sciences*, 116(42):20923–20929, 2019.

M. Grabisch and P. Miranda. Exact bounds of the mÃ¶bius inverse of monotone set functions. *Discrete Applied Mathematics*, 186:7–12, 2015.

F. Gul and W. Pesendorfer. Random expected utility. *Econometrica*, 74:121–146, 2006.

G. W. Harrison, J. MartÃnez-Correa, and J. T. Swarthout. Inducing risk neutral preferences with binary lotteries: A reconsideration. *Journal of Economic Behavior & Organization*, 94:145–159, 2013.

I. Herstein and J. Milnor. An axiomatic approach to measurable utility. *Econometrica*, 21: 291–297, 1953.

N. Hing, M. Smith, M. Rockloff, H. Thorne, A. M. T. Russell, N. A. Dowling, and B. Breen. How structural changes in online gambling are shaping the contemporary experiences and behaviours of online gamblers: an interview study. *BMC Public Health*, 22(1620), 2022.

X. Lim and S. R. Avoiding risk in the lab: An experiment on avoidable risk. Working paper, 2021. URL `https://silvioravaioli.github.io/website-documents/Research/Avoidable%20Risk/LimRavaioli2021AvoidableRisk.pdf`.

J. Lu. Random ambiguity. *Theoretical Economics*, 16(2):539–570, 2021.

D. McFadden. Conditional logit analysis of qualitative choice behavior. In P. Zarembka, editor, *Frontiers in Economics*. Academic Press, New York, 1973.

D. McFadden. Economic choices. *American Economic Review*, 91(3):351–378, 2001. doi: 10.1257/aer.91.3.351.

R. Myerson. Multistage games with communication. *Econometrica*, pages 323–358, 1986.

S. Olschewski, P. Sirotkin, and J. Rieskamp. Empirical underidentification in estimating random utility models: The role of choice sets and standardizations. *British Journal of Mathematical and Statistical Psychology*, 75(2):252–292, 2022.

K. Peng. Certainty as a decoy: an experiment to test menu dependence in heterogeneous risk attitudes. Working paper, 2025.

E. Piermont. Disentangling strict and weak choice in random expected utility models. *Journal of Economic Theory*, 202:201–230, 2022.

M. Rabin. Psychology and economics. *Journal of Economic Literature*, pages 11–46, 1998.

E. Shafir, I. Simonson, and A. Tversky. Reason-based choice. *Cognition*, 49:11–36, 1993.

K. M. Sharpe, R. Staelin, and J. Huber. Using extremeness aversion to fight obesity: Policy implications of context dependent demand. *Journal of Consumer Research*, 35(3):406–422, 2008.

K. E. Train. Recreation demand models with taste differences over people. *Land economics*, pages 230–239, 1998.

H. Williams. On the formulation of travel demand models and economic measures of user benefit. *Environment and Planning A*, 9:285–344, 1977.

E. Yang and I. Kopylov. Random quasi-linear utility. *Journal of Economic Theory*, 209, 2023. doi: 10.1016/j.jet.2023.105650.