A Conditional Probability Hierarchy for Stochastic Choice*

Erya Yang †   Adam Brandenburger ‡

September 6, 2025

Abstract

We develop a hierarchy of families of stochastic choice rules (SCR's) based on the concept of a conditional probability space (CPS) due to Rényi (38, 1955; 39, 1956). The CPS machinery allows for specializations and generalizations that yield a natural sequence of families of SCR's. Level 1 of our hierarchy consists of point CPS's (PCPS's), which are CPS's each of whose component probability measures concentrate on a point. Level 2 consists of CPS's. Level 3 is the set of probabilistic mixtures of PCPS's and Level 4 is the set of signed mixtures of PCPS's. This work aims at two unifications of results in stochastic choice. The first is that we establish our hierarchy at a general measure-theoretic level, where the underlying set of elementary choice objects is any measurable space and the sub-family of choice sets is finitary (i.e., the ideal of finite subsets). At this same level of generality, we characterize Levels 1-3 by the Weak Axiom of Stochastic Revealed Preference, a generalized Independence from Irrelevant Alternatives (IIA) axiom, and a generalized Block-Marschak condition, respectively. If the set of elementary choices is finite, Level 4 corresponds to no restriction (the infinite case appears to be open). The second unification we offer is a treatment of four canonical examples of non-classical stochastic behavior – specifically, the Similarity, Compromise, Attraction, and Repulsion Effects – in terms of a single idea, which is mixing over our basic stochastic choice rule of PCPS's. These examples also serve to establish that the levels in our hierarchy are strictly nested.

1 Introduction

The field of stochastic choice has deep roots in psychology (Luce, 29, 1959), discrete choice (McFadden, 33, 1974), and behavorial decision theory (Gabaix 20, 2019). The starting point of theoretical development in this field is the classic Independence from Irrelevant Alternatives Axiom (IIA) due to Luce (29, 1959). But, very quickly, the empirical validity of this axiom was called into question (Debreu, 14) and a vast and rich literature has since been built on the basis of alternative axioms and choice rules. (See Strzalecki 46, 2022 for a comprehensive survey.)

In this paper, we go back to the beginning and IIA, and we make this the starting point for a new architecture of stochastic choice. Our beginning observation is that a stochastic choice rule (SCR) satisfying IIA is formally equivalent to a concept in probability theory called a conditional probability space (CPS) due to Rényi (38, 1955; 39, 1956). (This equivalence was briefly noted by Luce 29, 1959, for strictly positive choice probabilities, and has been generalized by Cerreia-Vioglio et al. 7, 2021.) A CPS is a family of probability measures indexed by a family of observable events. The key condition on a CPS is a chain rule, which disciplines the way probabilities are updated on nested sequences of events. Clearly, a CPS can also carry a choice interpretation. The observable events become choice sets and the associated probabilities become choice probabilities.

The import of the equivalence between IIA choice rules and CPS's is that the CPS concept admits natural specializations and generalizations. Thus, we define the concept of a point CPS (PCPS), where all probability measures that appear are concentrated (degenerate) on single points. We also consider probabilistic mixtures of PCPS's and signed mixtures of PCPS's. Including the basic CPS concept, these

*We thank Samson Abramsky, Victor Aguiar, Inacio Bo, Stefan Bucher, Emiliano Catonini, Vered Kurtz-David, Igor Kopylov, Kenway Louie, Jay Lu, Geralt Tserenjigmid, and Stuart Zoble for important input, and Kai Steverson for early inspiration for this project. We also thank participants in workshops at University of Macau, the Harbin Institute of Technology (Shenzhen), the 2025 Barcelona School of Economics Summer Forum, the Society for the Advancement of Economic Theory 2025 Conference (SAET 2025), the 14th International Symposium on Imprecise Probabilities: Theories and Applications (ISIPTA 2025), and the 13th World Congress of the Econometric Society (ESWC 2025) for comments. Brandenburger acknowledges financial support from NYU Stern School of Business, NYU Shanghai, and J.P. Valles. Yang acknowledges financial support from National Natural Science Foundation of China (NSFC), China, Award No. 72403259.

†School of Business, Sun Yat-sen University; Research Center for Innovation, Entrepreneurship, and Technology Finance, Sun Yat-sen University, Shenzhen, 518107, China, yangery@mail.sysu.edu.cn; corresponding author

‡Stern School of Business, Tandon School of Engineering, NYU Shanghai, New York University, New York, NY 10012, U.S.A., adam.brandenburger@stern.nyu.edu

four objects establish a well-grounded hierarchy of SCR's, built in a precise and progressive manner from the classic starting point of IIA. See Figure 1 for a summary of our CPS hierarchy.

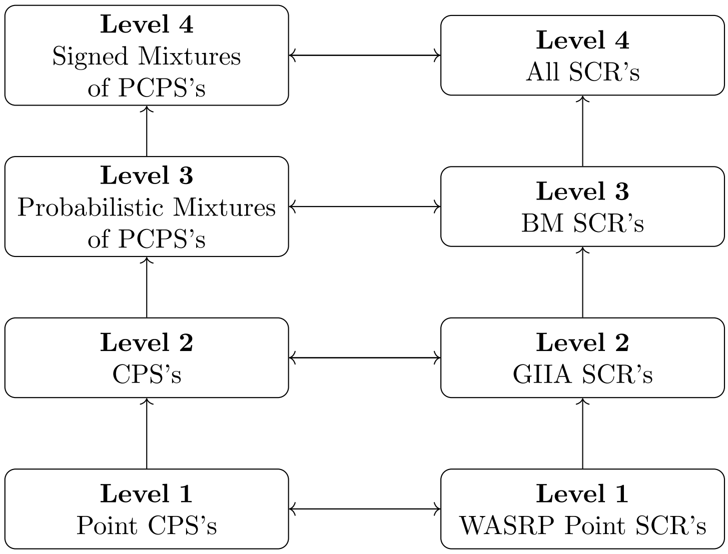

Figure 1: A Conditional Probability Hierarchy and Associated SCRs

This paper aims at a unification of results in stochastic choice for a broad setting. To this end, we establish our choice hierarchy at a general measure-theoretic level, where the underlying set of elementary choice objects is any measurable space and the sub-family of choice sets is finitary (i.e., the ideal of finite subsets). This direction appears to have been initiated by Hildebrand (23, 1971), who studied a general-equilibrium model with random preferences. McFadden and Richter (35, 1971; 36, 1991) and McFadden (34, 2005, Theorem 5.3) obtain a characterization of random utility in the case where the underlying space is Polish and the choice sets are compact. Clark (8, 1996) provides a very general characterization result for random utility based on the de Finetti probability-coherency condition (de Finetti 15, 1937). Closest to our work is Cohen (9, 1980), who proves an equivalence between finitely additive random choice and random utility in the finitary case. We will make use of techniques in this last paper. To the best of our knowledge, our paper is new in offering a unified treatment of a full hierarchy for stochastic choice in a general measure-theoretic (although finitary) setting. In particular, our focus is not on random-utility representations but instead on systematically building a full choice hierarchy from PCPS's, which are a very basic stochastic choice rule.

Our general hierarchy is characterized at each level by existing or generalized axioms for stochastic choice. Level 1 is characterized by the Weak Axiom of Stochastic Revealed Preference or WASRP (Bandyopadhyay, Dasgupta, and Pattanaik, 2, 1999, 3, 2002; Dasgupta and Pattanaik 12, 2007). (Level 1 concerns point SCR's, which are degenerate on each choice set.) Level 2 is characterized by a Generalized Independence of Irrelevant Alternatives (GIIA) axiom we define. Level 3 is characterized by a generalized Block-Marschak (BM) condition (Block and Marschak 4, 1960; Cohen 9, 1980). If the underlying set of choices is finite, then Level 4 corresponds to no restriction. This follows from Dogan and Yildiz 16, 2023, Theorem 2). A general characterization of Level 4 appears to be open.

Each level of our hierarchy is nested within the next one. This is true by definition for Level 1 vs. Level 2 and Level 3 vs. Level 4. We establish that Level 2 is nested within Level 3 by proving that our GIIA axiom implies a generalized BM condition. This proof involves some new techniques involving dimensionally-ordered systems of measures (Rényi 39, 1956) which we believe may be of independent interest.

The inclusions in our hierarchy are strict. This is immediate for Level 1 vs. Level 2. Level 3 strictly nests Level 2 because we can cast the Similarity Effect (Debreu 14, 1960; Tversky 47, 1972, 48, 1972; McFadden 35, 1974) and the Compromise Effect (Simonson 42, 1989; Simonson and Tversky 44, 1992) as probabilistic mixtures of PCPS's that violate IIA. In a similar fashion, we place the Attraction Effect (Huber, Payne, and Puto 24, 1982; Simonson 42, 1989) and the Repulsion Effect (Aaker 1, 1991) in Level 4, by casting them in terms of signed mixtures of PCPS's. As is well known, these last two effects violate the Regularity Axiom, which is implied by BM, so they strictly separate our Level 4 from our Level 3. We use these behavioral effects – Similarity, Compromise, Attraction, and Repulsion – as a convenient way to establish strict inclusions. But we hope that our treatment of them is a second way that our paper unifies some aspects of stochastic choice. These four effects are usually described quite differently in behavioral terms (as we review later). By viewing all of them as arising from mixtures over our basic PCPS object, we aim to bring out their common architecture.

The organization of the rest of the paper is as follows. Section 2 introduces stochastic choice rules and conditional probability spaces in our general setting. Section 3 defines probabilistic and signed mixtures of PCPS's. Section 4 proves the equivalence between CPS's and SCR's satisfying GIIA (Level 2) and between probabilistic mixtures of PCPS's and BM (Level 3). The main result of this section is the proof that if an SCR satisfies GIIA, then it satisfies BM. Section 5 completes the axiomatization of levels of our choice hierarchy by proving the equivalence between PCPS's and point SCR's satisfying WASRP. It then shows that Level 4 corresponds to no restrictions on SCR's in the finite case. Section 6 covers the four behavioral effects we mentioned earlier. Section 7 explains why our choice hierarchy differs from one built from total orders (the short reason is that mixing and the chain rule for CPS's do not commute) and contains discussion of some literature that adds finer detail relative to our hierarchy and of an alternative stochastic transitivity hierarchy. Short proofs of theorems are in the main text, while longer proofs can be found in two appendices.

2 SCR's and CPS's

In this section, we define the objects that occupy Levels 1 and 2 of our hierarchy. Given a measurable space $(\Omega, \mathcal{F})$, we write $\Delta(\Omega)$ for the set of all probability measures on $(\Omega, \mathcal{F})$. Let $\mathcal{G}$ be a sub-family of $\mathcal{F}$. Throughout, we will assume that $\emptyset \notin \mathcal{G}$.

Definition 1. *A **stochastic choice rule (SCR)** (relative to $\mathcal{G}$) is a map $c : \mathcal{G} \to \Delta(\Omega)$, which we write as $G \mapsto c_G(\cdot)$ for $G \in \mathcal{G}$, satisfying:*

1. *$c_G(G) = 1$ for every $G \in \mathcal{G}$.*

Definition 2. *A **conditional probability space (CPS)** (relative to $\mathcal{G}$) is a map $p : \mathcal{G} \to \Delta(\Omega)$, which we write as $G \mapsto p_G(\cdot)$ for $G \in \mathcal{G}$, satisfying:*

1. *$p_G(G) = 1$ for every $G \in \mathcal{G}$;*
2. *$p_G(E) = p_G(F)p_F(E)$ for every $E \subseteq F \subseteq G$ with $E \in \mathcal{F}$ and $F, G \in \mathcal{G}$.*

Thus, a CPS is an SCR satisfying an extra chain rule requirement, namely, Condition 2 of Definition 2). (In Rényi 38, 1955, Condition 2 takes the form: If $E, F \in \mathcal{F}$, $G \in \mathcal{G}$, and $F \cap G \in \mathcal{G}$, then $p_G(E \cap F) = p_G(F)p_{F \cap G}(E)$. This is readily seen to be equivalent to our Condition 2.)

Luce (29, 1959) already noted the equivalence between CPS's and SCR's satisfying IIA, for the case of finite choice sets and strictly positive probabilities. Cerreia-Vioglio et al. (7, 2021) extend this equivalence to the finitary case and non-negative probabilities.

Next, let $s : \mathcal{G} \to \Omega$ be a selection on $\mathcal{G}$, that is, for each $G \in \mathcal{G}$, $s(G) \in G$.

Definition 3. *Fix a SCR $c : \mathcal{G} \to \Delta(\Omega)$. Suppose there is a selection s such that, for each $G \in \mathcal{G}$ and $F \in \mathcal{F}$:*

$$c_G(F) = \mathbf{1}_F(s(G)). \tag{1}$$

*That is, $c_G(\cdot)$ is the Dirac measure concentrated on $s(G)$. We call such an SCR a **point SCR (PSCR)**.*

Definition 4. *Fix a CPS $p : \mathcal{G} \to \Delta(\Omega)$. Suppose there is a selection s such that, for each $G \in \mathcal{G}$ and $F \in \mathcal{F}$:*

$$p_G(F) = \mathbf{1}_F(s(G)). \tag{2}$$

*That is, $p_G(\cdot)$ is the Dirac measure concentrated on $s(G)$. We call such a CPS a **point CPS (PCPS)**.*

3 Probabilistic and Signed Mixtures of PCPS's

To define Level 3 of our hierarchy, we need to build a measurable space of PCPS's. We do so by adapting methods from Cohen (9, 1980). Fix an arbitrary (infinite) set Ω. Let the family $\mathcal{G}$ of choice sets consist of all finite nonempty subsets of Ω, and let $\mathcal{F}$ be the field generated by $\mathcal{G}$. We will call $\mathcal{G}$ finitary. Note, in particular, that $\mathcal{F}$ contains all singletons $\{\omega\}$. Let Π be the set of all PCPS's (relative to the fixed family $\mathcal{G}$) on $(\Omega, \mathcal{F})$.

Next, define the set:

$$\mathcal{A} = \{(\alpha, G) : \alpha \in G \text{ and } G \in \mathcal{G}\}, \tag{3}$$

and, for each $(\alpha, G) \in \mathcal{A}$:

$$[\alpha, G]_\Pi = \{p \in \Pi : p_G(\alpha) = 1\}. \tag{4}$$

Let $\mathcal{D}_\Pi$ be a field on Ω containing the sets $[\alpha, G]_\Pi$. (We choose a specific such field in the next section.) Fix a finitely additive probability measure Q on $(\Pi, \mathcal{D}_\Pi)$. We now have the apparatus needed to define Level 3. Under mixing over PCPS's, the probability that the choice α is made from the choice set G is exactly $Q([\alpha, G]_\Pi)$. With this, we can say what it means for an SCR to be realized in this fashion.

Definition 5. *An SCR $G \mapsto c_G(\cdot)$ is **realizable as a Π-based probability mixture** if there is a finitely additive probability measure Q on $(\Pi, \mathcal{D}_\Pi)$ such that for all $(\alpha, G) \in \mathcal{A}$:*

$$c_G(\{\alpha\}) = Q([\alpha, G]_\Pi). \tag{5}$$

We restrict the measure theory here to the finitely additive case because later we will make use of a finitely-additive characterization result in Cohen (9, 1980). The extension of Level 3 – and Level 4 below – of our choice hierarchy to the countably additive case is, we believe, open.

To define Level 4, we employ the same measurable structure as for probabilistic mixtures, and simply replace unsigned with signed measures.

Definition 6. *An SCR $G \mapsto c_G(\cdot)$ is **realizable as a Π-based signed probability mixture** if there a finitely additive signed probability measure Q on $(\Pi, \mathcal{D}_\Pi)$ such that for all $(\alpha, G) \in \mathcal{A}$:*

$$c_G(\{\alpha\}) = Q([\alpha, G]_\Pi). \tag{6}$$

The introduction of negativity in this definition can be seen as a formal move, or can be given a behavioral interpretation (Dogan and Yildiz 16, 2023; Brandenburger et al. 5, 2025), as we will see later.

4 Nesting of Levels

Two of the three relationships between levels of our hierarchy in Figure 1 are immediate from the definitions. Level 2 nests Level 1 by relaxing degeneracy of the component probability measures in a PCPS. Level 4 nests Level 3 by relaxing non-negativity of the mixing probability measure Q. This leaves the relationship between Levels 2 and 3, which is not immediate. We establish that Level 3 indeed nests Level 2 by making use of an axiomatic characterization of each level and then relating the axioms. We are not aware of a direct proof of this nesting, at least in the general setting we consider. That said, our indirect route employs some techniques – in particular, a theorem on dimensionally ordered systems of measures – that may find other uses in the field of stochastic choice. We begin with a definition for Level 2.

Definition 7. *A stochastic choice rule $c : \mathcal{G} \to \Delta(\Omega)$ satisfies **Generalized Independence of Irrelevant Alternatives (GIIA)** if for every $G, H \in \mathcal{G}$, and every $E, F \in \mathcal{F}$ with $E, F \subseteq G \cap H$, we have:*

$$c_G(E) \times c_H(F) = c_H(E) \times c_G(F). \tag{7}$$

It is clear that GIIA implies the usual statement of IIA in terms of probability ratios, in the case that $c_H(E) \times c_H(F) \neq 0$. The following result is essentially already in Cerreia-Vioglio et al. (7, 2021). We state it here under the assumption that $\mathcal{G}$ is closed under finite intersections.

Theorem 1. *Fix a measurable space $(\Omega, \mathcal{F})$ and a sub-family $\mathcal{G}$ of $\mathcal{F}$ which is a π-system. The family of CPS's $G \mapsto p_G(\cdot)$ coincides with the family of SCR's $G \mapsto c_G(\cdot)$ satisfying Generalized Independence of Irrelevant Alternatives (GIIA).*

Proof. Fix a CPS $G \mapsto p_G(\cdot)$. Also fix $G, H \in \mathcal{G}$ and $E, F \in \mathcal{F}$, where $E, F \subseteq G \cap H$. Using the equivalent Rényi (38, 1955) form of Condition 2 of a CPS, we can write:

$$p_G(E) = p_G(E \cap H) = p_G(H) \times p_{G \cap H}(E), \tag{8}$$
$$p_G(F) = p_G(F \cap H) = p_G(H) \times p_{G \cap H}(F), \tag{9}$$
$$p_H(E) = p_H(E \cap G) = p_H(G) \times p_{G \cap H}(E), \tag{10}$$
$$p_H(F) = p_H(F \cap G) = p_H(G) \times p_{G \cap H}(F). \tag{11}$$

Multiplying the left sides of Equations 8 and 11, and the left sides of Equations 9 and 10, we obtain:

$$p_G(E) \times p_H(F) = p_H(E) \times p_G(F), \tag{12}$$

establishing the forward direction of the proof. (This direction extends the proof of Theorem 6 in Rényi, 38, 1955.) For the reverse direction, suppose that Equation 12 holds and $H \subseteq G$, and set $F = H$. We obtain

$$c_G(E) \times c_H(H) = c_H(E) \times c_G(H), \tag{13}$$

from which, since $c_H(H) = 1$, our Condition 2 of a CPS is satisfied. □

Turning to Level 3, we obtain an axiomatic characterization in terms of the classic Block and Marschak (4, 1960) condition, as generalized by Cohen (9, 1980, Definition 2.4).

Definition 8. *The* ***Block-Marschak (BM) condition*** *associated with an SCR $G \mapsto c_G(\cdot)$ is the requirement that the function $q : \Omega \times \mathcal{G} \times \mathcal{G} \to \mathbb{R}$ defined by:*

$$q(\alpha, G, H) = \sum_{j=0}^{|G|} (-1)^{|G|-j} \sum_{\{E \subseteq G : |E| = j\}} c_{H \backslash E}(\{\alpha\}), \tag{14}$$

for $G \subsetneq H$ and $\alpha \in H \backslash G$, is everywhere non-negative.

Theorem 2. *Fix a set Ω and a finitary sub-family $\mathcal{G}$ of subsets of Ω. The family of SCR's $G \mapsto c_G(\cdot)$ realizable as Π-based probability mixtures coincides with the family of SCR's $G \mapsto c_G(\cdot)$ satisfying the Block-Marschak (BM) condition.*

Our method of proof is to transfer a characterization result in Cohen (9, 1980, Theorem 4.1) on total (complete, transitive, and antisymmetric) orders to PCPS's. (The Cohen result extends the classic equivalence result of Falmagne 17, 1978 to the infinite choice context.) Let Λ be the set of all total orders $\triangleright$ on Ω. For $(\alpha, G) \in \mathcal{A}$, we mirror Equation 4 and define the sets:

$$[\alpha, G]_\Lambda = \{\triangleright \in \Lambda : \alpha \triangleright G \backslash \{\alpha\}\}, \tag{15}$$

where $\alpha \triangleright G \backslash \{\alpha\}$ means $\alpha \triangleright \beta$ for all $\beta \in G \backslash \{\alpha\}$. Cohen (9, 1980, Section 4) builds a field $\mathcal{B}_\Lambda$ on Λ with the property that all such sets $[\alpha, G]_\Lambda$ lie in this field. We want to build a field $\mathcal{D}_\Pi$ on Π that contains all sets $[\alpha, G]_\Pi$ and on which we can define a finitely additive probability measure Q from a finitely additive probability measure P on $(\Lambda, \mathcal{B}_\Lambda)$. To do so, we make use of the theory of charges (i.e., finitely additive measures). We begin with a definition (Rao and Rao 37 1983, Definition 3.2.2).

Definition 9. *Fix an arbitrary set Ω and let $\mathcal{E}$ be a collection of subsets of Ω. A function $\mu : \mathcal{E} \to \mathbb{R}^+$ is a* ***positive real partial charge on*** $(\Omega, \mathcal{E})$ *if:*

$$\sum_{i=1}^{M} \mathbf{1}_{C_i} \leq \sum_{j=1}^{N} \mathbf{1}_{D_j} \tag{16}$$

implies:

$$\sum_{i=1}^{M} \mu(C_i) \leq \sum_{j=1}^{N} \mu(D_j), \tag{17}$$

for any $C_1, \ldots, C_M, D_1, \ldots, D_N \in \mathcal{E}$.

To use this theory, consider the following collection of subsets of Π:

$$\mathcal{H} = \{[\alpha, G]_\Pi : (\alpha, G) \in \mathcal{A}\} \cup \{\Pi\}, \tag{18}$$

and define a function $Q : \mathcal{H} \to \mathbb{R}^+$ by:

$$Q([\alpha, G]_\Pi) = P([\alpha, G]_\Lambda) \text{ and } Q(\Pi) = 1. \tag{19}$$

The next result is key. It is proved in Appendix A.

Theorem 3. *The function Q is a positive real partial charge on $\mathcal{H}$.*

We now appeal to an extension theorem for charges (Rao and Rao 37, 1983, Theorem 3.2.10).

Theorem 4. *Fix an arbitrary set Ω, let $\mathcal{E}$ be a collection of subsets of Ω with $\Omega \in \mathcal{E}$, and suppose μ is a positive real partial charge on $(\Omega, \mathcal{E})$. Then, for any field $\mathcal{D}$ containing $\mathcal{E}$, there is a finitely additive probability measure $\tilde{\mu}$ on $(\Omega, \mathcal{D})$ that extends μ.*

To apply this theorem, let $\mathcal{D}_\Pi$ denote the field generated by the collection of sets $\mathcal{H}$ in Equation 18. (In the statement to follow, we do not distinguish the extension of Q from Q itself. No confusion should result.)

Theorem 5. *Fix the finitely additive probability space $(\Lambda, \mathcal{B}_\Lambda, P)$. There is a finitely additive probability measure Q on $(\Pi, \mathcal{D}_\Pi)$ such that:*

$$Q([\alpha, G]_\Pi) = P([\alpha, G]_\Lambda), \tag{20}$$

for all $(\alpha, G) \in \mathcal{A}$.

Proof. The proof is immediate from putting together Theorems 3, Theorem 4, and the definition of Q in Equation 19. $\square$

We can now transfer the characterization result from Cohen (9, 1980).

Definition 10. *An SCR $G \mapsto c_G(\cdot)$ is **realizable as a Λ-based probability mixture** if there a finitely additive probability measure P on $(\Lambda, \mathcal{B}_\Lambda)$ such that for all $(\alpha, G) \in \mathcal{A}$:*

$$c_G(\{\alpha\}) = P([\alpha, G]_\Lambda). \tag{21}$$

Cohen (9, Theorem 4.1) proves that an SCR is realizable in this way if and only if the generalized BM condition (our Definition 8) is satisfied. This result together with our Theorem 5 immediately establishes our Theorem 2.

Now that we have axiomatizations – via GIIA and BM, respectively – of Levels 2 and 3, we can prove nestedness via the following result.

Theorem 6. *Fix a measurable space $(\Omega, \mathcal{F})$ and a finitary sub-family $\mathcal{G}$ of $\mathcal{F}$. If an SCR $G \mapsto c_G(\cdot)$ satisfies GIIA, then it satisfies BM.*

The proof of this theorem relies on a representation of CPS's given by Rényi (39, 1956).

Definition 11. *Fix a measurable space $(\Omega, \mathcal{F})$, and a sub-family $\mathcal{G}$ of $\mathcal{F}$ with $\emptyset \notin \mathcal{G}$. Let $(\Gamma, \prec)$ denote an arbitrary totally ordered index set, and suppose that for every $\gamma \in \Gamma$ there is an associated measure μ_γ on $(\Omega, \mathcal{F})$. Call the system of measures **dimensionally ordered** if:*

1. *for every $G \in \mathcal{G}$, there is a $\gamma \in \Gamma$ such that $0 < \mu_\gamma(G) < +\infty$;*
2. *if $\mu_\alpha(G) < +\infty$ and $\alpha \prec \beta$, then $\mu_\beta(G) = 0$.*

Note that for every $G \in \mathcal{G}$, there is a unique $\gamma \in \Gamma$ such that $0 < \mu_\gamma(G) < +\infty$. Also, if $\beta \prec \gamma$, then $\mu_\beta(G) = +\infty$. Now, for $G \in \mathcal{G}$ and $F \in \mathcal{F}$, we define:

$$p_G(F) = \frac{\mu_\gamma(F \cap G)}{\mu_\gamma(G)}, \tag{22}$$

where γ is this unique index for G. The next result is implied by the proof of Theorem 1 in Rényi (39, 1956). We provide a self-contained proof in Appendix B.

Theorem 7. *The family of maps $G \mapsto p_G(\cdot)$ defined in Equation 22 is a CPS relative to $\mathcal{G}$.*

The following representation result is Theorem 1 in Rényi (39, 1956). It gives a sufficient condition for a CPS to be generated by a dimensionally ordered system of measures. Appendix B gives an outline of the proof (which is an extended exercise in measure theory).

Theorem 8. *Fix a CPS $G \mapsto p_G(\cdot)$ where $\mathcal{G}$ is closed under finite unions. Then there is a dimensionally ordered system of measures $\{\mu_\gamma : \gamma \in \Gamma\}$ such that each $p_G(\cdot)$ is obtained via Equation 22.*

The proof of our Theorem 6 first observes that, by our Theorem 1, if an SCR satisfies GIIA, then it is a CPS. If $\mathcal{G}$ is finitary, then it is closed under finite unions. Therefore, by Theorem 8, this CPS is generated by a dimensionally ordered system of measures. From this, and using Equation 22, we can write, for all $(\alpha, G) \in \mathcal{A}$:

$$c_G(\{\alpha\}) = \frac{\mu_{\gamma(G)}(\{\alpha\})}{\mu_{\gamma(G)}(G)}, \tag{23}$$

where, for any $F \in \mathcal{G}$, $\gamma(F)$ is the unique index with $0 < \mu_\gamma(F) < +\infty$. Turning to the BM condition of Equation 14, we use this formula for the terms on the right-hand side. We are then able to establish that $q(\alpha, G, H) \geq 0$ when G is a singleton, and we go on to treat the general case by successive substitution. The full proof is in Appendix B.

With this result, all nested relationships indicated in Figure 1 are verified.

5 Axiomatization of Levels Contd.

We have already stated axiomatic characterizations of Levels 2 and 3 of our choice hierarchy, in order to establish nestedness (Theorem 6). We now go back to Level 1 and axiomatize this level. To do so, we make use of a stochastic analog to the Weak Axiom of Revealed Preference (WARP) of classic demand theory introduced by Bandyopadhyay, Dasgupta, and Pattanaik (2, 1999, 3, 2002) and Dasgupta and Pattanaik (12, 2007).

Definition 12. *A stochastic choice rule $G \mapsto c_G(\cdot)$ satisfies the* ***Weak Axiom of Stochastic Revealed Preference (WASRP)*** *if for every $G, H \in \mathcal{G}$ and every $F \in \mathcal{F}$ with $F \subseteq G \cap H$:*

$$c_H(F) - c_G(F) \leq c_G(G \backslash H). \tag{24}$$

We next establish our characterization of Level 1 and then discuss the meaning of WASRP. Our result is stated for the assumption (weaker than finitarity) that $\mathcal{G}$ is closed under finite intersections and set differences. Recall the meaning of a point SCR (PSCR) from Definition 3.

Theorem 9. *Fix a measurable space $(\Omega, \mathcal{F})$ and a sub-family $\mathcal{G}$ of $\mathcal{F}$ which is a semiring. The set of PCPS's $G \mapsto p_G(\cdot)$ coincides with the set of PSCR's $G \mapsto c_G(\cdot)$ satisfying WASRP.*

Proof. Fix a PCPS $p_G(\cdot)$ relative to $\mathcal{G}$. We need to show that $p_G(\cdot)$, viewed as an SCR, is a PSCR and satisfies WASRP. It is immediate that $p_G(\cdot)$ is a PSCR, so we show that WASRP is satisfied. Fix $G, H \in \mathcal{G}$ and $F \in \mathcal{F}$ with $F \in G \cap H$. Case (i): $s(G) \in G \backslash H$. Then $p_G(G \backslash H) = 1$, from which Inequality 24 is satisfied. Case (ii): $s(G) \in F$. Then $p_G(F) = 1$, from which Inequality 24 is again satisfied. Case (iii): $s(G) \in (G \cap H) \backslash F$. Then $p_G(F) = 0$ and $p_G(G \backslash H) = 0$, so, in light of Inequality 24, we need to show that $p_H(F) = 0$. Suppose not, so that $s(H) \in F$ and $p_H(F) = 1$. Then $p_H(G \cap H) = 1$. Using $p_H(F) = p_H(G \cap H)\, p_{G \cap H}(F)$, we find $p_{G \cap H}(F) = 1$. Using $p_G(F) = p_G(G \cap H)\, p_{G \cap H}(F)$, $p_G(F) = 0$ (as already shown), and $p_G(G \cap H) = 1$ (since $s(G) \in G \cap H$), we find $p_{G \cap H}(F) = 0$, a contradiction. This establishes one direction of the proof.

For the reverse direction, fix an SCR $c_G(\cdot)$ relative to $\mathcal{G}$ satisfying WASRP. We need to show that the chain rule holds: $c_G(E) = c_G(F) c_F(E)$ for every $E \subseteq F \subseteq G$ with $E \in \mathcal{F}$ and $F, G \in \mathcal{G}$. Write WASRP in the form:

$$c_G(E) - c_F(E) \leq c_F(F \backslash G). \tag{25}$$

Case (i): $s(G) \in E$ and $s(F) \in E$. Then $c_G(E) = 1$, and therefore $c_G(F) = 1$, and also $c_F(E) = 1$, so that the chain rule is satisfied. Case (ii): $s(G) \in E$ and $s(F) \notin E$. Then $c_G(E) = 1$ and $c_F(E) = 0$. But from Inequality 25, using $c_F(F \backslash G) = c_F(\emptyset) = 0$, we then get $1 - 0 \leq 0$, a contradiction. Case (iii): $s(G) \notin F$. Then $c_G(E) = 0$ and $c_G(F) = 0$, so that the chain rule is again satisfied. □

To explain WASRP, observe that the axiom does not require equality of the choice probabilities $c_H(F)$ and $c_G(F)$. This might be the obvious extension of WARP but would be too restrictive. Instead, WASRP limits how big the difference in probabilities can be, which must lie in the range $[-c_H(H \backslash G), +c_G(G \backslash H)]$. Bandyopadhyay, Dasgupta, and Pattanaik (2, 1999) and Dasgupta and Pattanaik (12, 2007) provide the following argument for the upper bound. When the choice set changes from G to H, this rules out "competing" choices in $G \backslash H$ to the alternatives in F. This effect may raise the probability $c_H(F)$ that the chosen alternative lies in F. But this increase, viz. the difference $c_H(F) - c_G(F)$, should be bounded above by the original choice probability for $G \backslash H$, viz. $c_G(G \backslash H)$. The argument for the lower bound is analogous.

It remains to characterize Level 4 of our stochastic choice hierarchy. We proceed by transferring a result from Dogan and Yildiz (16, 2023). Fix finite Ω and let $\mathcal{F} = 2^\Omega$ and $\mathcal{G} = 2^\Omega \backslash \{\emptyset\}$. Theorem 2 in Dogan and Yildiz (16, 2023) states that every SCR on Ω is realizable as as a Λ-based signed probability mixture. Here, the definition of realizability is the same as in our Definition 10 except that P is allowed to be signed. This result can be transferred to our setting in the same way we transferred Cohen (9, 1980, Theorem 4.1) earlier. Recalling Definition 6, we have a characterization of Level 4 – namely, that all SCR's are now included.

Theorem 10. *Suppose the choice domain Ω is finite, $\mathcal{F} = 2^\Omega$, and $\mathcal{G} = 2^\Omega \backslash \emptyset$. Then every SCR $G \mapsto c_G(\cdot)$ is realizable as a Π-based signed probability mixture.*

Saito (41, 2018) proves a more general result on realization of SCR's as signed mixtures of orders, where the family $\mathcal{G}$ of choice sets can be an arbitrary (nonempty) subset of $2^\Omega \backslash \{\emptyset\}$. Saito (41, 2018, Corollary 6(ii) and Footnote 22) demonstrates how Theorem 2 in Dogan and Yildiz (16, 2023) follows.

The proof of the Dogan-Yildiz results works via a (significant) extension of the Ford-Fulkerson Theorem from combinatorial matrix theory (Ford and Fulkerson 18, 2015) to allow for negative row and column sums. The generalization of this method to an infinite choice domain Ω is, we believe, open.

6 Strict Nesting

Challenges to the IIA axiom as an accurate description of stochastic behavior go back to Debreu (14, 1960). Starting then, many behavioral phenomena have been identified that violate this and other choice axioms. In this section, we cast four classic such phenomena in the language of our framework. (Davis-Stober et al. (13 2023) is a recent survey of the relevant theoretical and empirical literature.) This is a convenient way to demonstrate strict nesting of the levels of our hierarchy. We will also make a conceptual observation that emerge from our treatment. We begin with the separation between Level 2 and Level 3. (The separation between Level 1 and Level 2 is immediate. It is witnessed by any non-point CPS.)

6.1 The Similarity Effect

Debreu (14, 1960) introduced the scenario represented by the Similarity Effect, which was named this way by Tversky (47, 1972, 48, 1972). We review the effect couched in the McFadden (33, 1974) "red bus-blue bus" language.

The underlying set of choices is $\Omega = \{\text{red bus}, \text{blue bus}, \text{taxi}\}$. The decision maker (DM) is indifferent between taking a bus or a taxi and flips a coin to decide between these two modes of travel. Within the bus category, the DM does not care about the color of the bus and therefore flips a coin to decide on red vs. blue if both choices are available. Set $\mathcal{G} = \{\Omega, \{\text{blue bus}, \text{taxi}\}\}$ and consider the following three PCPS's p^{r}, p^{b}, p^{t}:

$$p^r_\Omega(\{\text{red bus}\}) = 1,\ p^r_{\{\text{blue bus,taxi}\}}(\{\text{blue bus}\}) = 1, \tag{26}$$

$$p^b_\Omega(\{\text{blue bus}\}) = 1,\ p^b_{\{\text{blue bus,taxi}\}}(\{\text{blue bus}\}) = 1, \tag{27}$$

$$p^t_\Omega(\{\text{taxi}\}) = 1,\ p^t_{\{\text{blue bus,taxi}\}}(\{\text{taxi}\}) = 1. \tag{28}$$

The DM puts probability $\frac{1}{4}$ on each of the first and second PCPS's p^r and p^b, and probability $\frac{1}{2}$ on the third PCPS p^t. If the choice set is Ω, then the DM averages over the PCPS's to arrive at probability $\frac{1}{2}$ on choosing a bus and probability $\frac{1}{2}$ on choosing a taxi. If the choice set is $\{\text{blue bus}, \text{taxi}\}$, then the DM averages to get again probability $\frac{1}{2}$ on a bus and probability $\frac{1}{2}$ on a taxi. This is the intuitive set of choice probabilities. But, in the first case, the probability ratio of blue bus to taxi is $1:2$, while in the second case it is $1:1$. This violates (G)IIA. By Theorem 1, the SCR just defined is not a CPS. We conclude that Level 3 of our choice hierarchy strictly nests Level 2.

6.2 The Compromise Effect

We next show how to encompass the Compromise Effect (Simonson 42, 1989; Simonson and Tversky 44, 1992) in the same manner. The underlying set of choices is $\Omega = \{\text{l-camera, m-camera, h-camera}\}$, where l-camera is the low-quality/low-price option, m-camera is the medium-quality/medium-price option, and h-camera is the high-quality/high-price option. Set $\mathcal{G} = \{\{\text{l-camera, m-camera}\}, \Omega\}$. Faced with the choice set $\{\text{l-camera, m-camera}\}$, the DM selects each camera with equal probability. Faced with the choice set Ω, the DM selects each of the l-camera and the h-camera with probability $\frac{1}{4}$, and selects the m-camera with probability $\frac{1}{2}$. This again is a violation of (G)IIA. The usual story is that the addition of the h-camera to the choice set emphasizes the inferiority of the l-camera, which in turn makes the m-camera stand out as a good compromise between the low quality of the l-camera and the high price of the h-camera.

This scenario can be represented via a probabilistic mixture of PCPS's:

$$p^l_\Omega(\{\text{l-camera}\}) = 1,\ p^l_{\{\text{l-camera, m-camera}\}}(\{\text{l-camera}\}) = 1, \tag{29}$$

$$p^h_\Omega(\{\text{h-camera}\}) = 1,\ p^h_{\{\text{l-camera, m-camera}\}}(\{\text{l-camera}\}) = 1, \tag{30}$$

$$p^m_\Omega(\{\text{m-camera}\}) = 1,\ p^m_{\{\text{l-camera, m-camera}\}}(\{\text{m-camera}\}) = 1. \tag{31}$$

The DM puts probability $\frac{1}{4}$ on each of the first and second PCPS's p^l and p^h, and probability $\frac{1}{2}$ on the third PCPS p^m. The key here is that when the choice set expands to include the h-camera, then, with probability $\frac{1}{4}$, the DM will switch from the l-camera to the h-camera, reflecting the inferiority of the former.

We point out that the Similarity and Compromise Effects are usually described quite differently in behavioral terms – as the words we chose to describe are meant to reflect. Despite this, the two effects have the same formal structure within our framework. While, arguably, some descriptive detail is lost under our approach, we think that this unification is helpful in bringing out the common underlying architecture of these behaviors.

6.3 The Attraction and Repulsion Effects

We now turn to the treatment of the Attraction Effect (Huber, Payne, and Puto 24, 1982; Simonson 42, 1989) and the Repulsion Effect (Aaker 1) in our framework. To proceed, we review the Regularity axiom of stochastic choice, which says that the probability of making a choice that lies in a given set is non-increasing in the size of the overall choice set. We then check in our framework the well-known fact that BM implies REG, which will set up our treatment of the two behavioral effects.

Definition 13. *A stochastic choice rule $G \mapsto c_G(\cdot)$ satisfies **Regularity (REG)** if for every $G, H \in \mathcal{G}$ with $G \subseteq H$, and every $F \in \mathcal{F}$ with $F \subseteq G$, we have $c_H(F) \leq c_G(F)$.*

Theorem 11. *Fix a measurable space $(\Omega, \mathcal{F})$ and a finitary sub-family $\mathcal{G}$ of $\mathcal{F}$. If an SCR $G \mapsto c_G(\cdot)$ satisfies BM, then it satisfies REG.*

Proof. Consider $G, H \in \mathcal{G}$ with $G \subseteq H$, and an SCR $G \mapsto c_G(\cdot)$ satisfying the BM condition. We want to show that $c_H(\{\alpha\}) \leq c_H(\{\alpha\})$. Fix a PCPS $G \mapsto p_G(\cdot)$ and some $\alpha \in G$. The chain rule gives $p_H(\{\alpha\}) = p_H(G) \times p_G(\{\alpha\})$. It follows that if $p_H(\{\alpha\}) = 1$, then both terms on the right side must be 1 In particular, we infer $p_G(\{\alpha\}) = 1$. From this:

$$[\alpha, H]_\Pi = \{p \in \Pi : p_H(\{\alpha\}) = 1\} \subseteq \{p \in \Pi : p_G(\{\alpha\}) = 1\} = [\alpha, G]_\Pi. \tag{32}$$

Using Theorem 2, we write $c_G(\alpha) = Q([\alpha, G]_\Pi)$ and $c_H(\alpha) = Q([\alpha, H]_\Pi)$. Monotonicity of Q yields $c_H(\alpha) \leq c_G(\alpha)$, as required. □

We now lay out the Attraction and Repulsion Effects in our framework. The underlying set of choices is $\Omega = \{x, y, z\}\}$ and $\mathcal{G} = \{\Omega, \{x, y\}\}$. We consider an SCR satisfying:

$$c_{\{x,y\}}(\{x\}) < c_\Omega(\{x\}), \tag{33}$$

which is a violation of Regularity. In one scenario for the Attraction Effect (Simonson and Tversky 44, 1992), item x is a nice pen, item y is a certain sum of money, and item z is a plain pen. The addition of item z highlights the attractiveness of item x, which is then chosen with higher probability. In a scenario for the Repulsion Effect (Simonson 43, 2014, Kruis et al. 28, 2020), item x is candy, item y is an orange, and item z is a spoiled clementine. The addition of item z casts doubt on the freshness of item y, so that, again, item x is chosen with higher probability.

We now show how to produce the the Attraction and Repulsion Effects via a signed mixture of PCPS's. Specifically, consider four PCPS's relative to $\mathcal{G}$:

$$p^1_{\{x,y\}}(\{x\}) = 1,\, p^1_{\{x,y,z\}}(\{x\}) = 1, \tag{34}$$

$$p^2_{\{x,y\}}(\{y\}) = 1,\, p^2_{\{x,y,z\}}(\{y\}) = 1, \tag{35}$$

$$p^3_{\{x,y\}}(\{x\}) = 1,\, p^3_{\{x,y,z\}}(\{z\}) = 1, \tag{36}$$

$$p^4_{\{x,y\}}(\{y\}) = 1,\, p^4_{\{x,y,z\}}(\{z\}) = 1. \tag{37}$$

We put a signed probability measure (q^1, q^2, q^3, q^4), where each $q^i \in \mathbb{R}$ and $\sum_{i=1}^4 q_i = 1$, on these four PCPS's and obtain:

$$c_{\{x,y\}}(\{x\}) = q^1 + q^3, \tag{38}$$

$$c_{\{x,y,z\}}(\{x\}) = q^1. \tag{39}$$

In order to create a violation of REG, we must have $q^3 < 0$. This makes intuitive sense. The PCPS p^3 selects item x – the nice pen or the candy – from the choice set $\{x, y\}$. But it selects item z – the plain pen or spoiled clementine – from the larger set $\{x, y, z\}$. We expect the DM to want to avoid choosing according to this PCPS. (We do not rule out that $q^4 < 0$ as well.)

Our analysis here is similar to the treatment of the Attraction Effect by Dogan and Yildiz (16, 2023, Example 2) in terms of orders. A story to go with the example is that the DM is a principal and there are four agents, each with one of the four PCPS's above. (This might be a multiple-selves story.) The principal has preferences over which agent gets to make choices and dislikes the event that agent 3 – with PCPS p^3 – is the one to choose. Formally, the DM has a negative willingness-to-bet on this event, which is an interpretation of a negative subjective probability from Brandenburger et al. 5, 2025).

At the formal level, this violation of Regularity via signed probability measures, together with Theorem 11, establishes that Level 4 of our choice hierarchy strictly nests Level 3. At the conceptual level, we note the unified treatment of all four behavioral effects – Similarity, Compromise, Attraction, and Repulsion. In our framework, these effects are all obtained in the same way, namely, by mixing over our basic PCPS object. As usual, mixing can, in turn, be understood as reflecting uncertainty over which specific basic choice rule is at work.

7 Conclusion

In this section, we make a conceptual point, cover some literature that refines our hierarchy, and mention an alternative hierarchy based on stochastic transitivity.

a. Non-Commutativity It might seem that our choice hierarchy is built entirely from PCPS's. If this were so, then it would be possible to use the language of total orders instead of starting from CPS's, as we do. Indeed, we built a map from total orders to PCPS's for two of our results. (In Section 3, we used Cohen 9, 1980 to define a finitely additive probability measure on a field of PCPS's on Ω. In Section 5, we likewise transferred a result on total orders from Dogan and Yildiz 16, 2023.)

But a careful look reveals that our hierarchy has a different fundamental character. The reason in short is that two operations – probabilistic mixing and imposing the chain rule – do not commute. Start with a point SCR. If we impose the chain rule, we obtain a PCPS. Mixing then (in general) yields an object in Level 3. Next, start again with a point SCR. If we mix, we obtain a general SCR. Imposing the chain rule then yields an object in Level 2, namely, a CPS. We know that Levels 2 and 3 are distinct. The conclusion is that the two operations are non-commutative and, for exactly this reason, our hierarchy is not made up only of mixing operations on PCPS's.

b. Related Literature Our four-level hierarchy for SCR's arose organically from the starting point of a CPS. Naturally, there are important finer distinctions between SCR's that can be made by adding more structure. Here, we comment on two such refinements.

First, one can distinguish between Level-4 SCR's that do or do not satisfy REG. A well-known example of an SCR that lies above BM but below REG is the additive perturbed utility (APU) model due to Fudenberg, Iijima, and Strzalecki (19, 2015). This model family can enccompass non-expected utility preferences arising from implementation costs and ambiguity aversion. A theory that is the convex dual to APU is the variational preference model of Maccheroni, Marinacci, and Rustichini (30, 2006), interpretable as a game against malevolent nature. Under restrictions on cost functions, these models all satisfy REG.

Lying above REG is the deliberately stochastic choice model of Cerreia-Vioglio et al. (6, 2019), where convex preferences over outcomes induce REG violations. Information processing architectures can also lead to departures from REG, as in the leading analysis by Matejka and McKay (32, 2015). The divisive normalization model of Steverson, Brandenburger, and Glimcher (45, 2019) employs a neuroscience-inspired information cost that likewise yields REG violations. Finally, two-stage models involving choice of consideration set and then choice of alternative can readily accommodate non-REG behavior (Kashaev and Aguiar 26, 2022).

Next, we locate some models in our Level 3 that carry additional structure. Galichon 21, 2022) shows how to write a nested logit model as a mixture of Luce models so that it accommodates the Compromise and Attraction Effects (but not the Attraction Effect in the form that violates REG). Behavioral models such as the the attribute model (Gul, Natenzon, and Pesendorfer 22, 2014) and the focal Luce model (Kovach and Tserenjigmid 27, 2022) also sit at this level.

c. Stochastic Transitivity Hierarchy A well-known hierarchy for stochastic choice is based on stochastic transitivity, which comes in three forms. An SCR $G \mapsto c_G(\cdot)$ satisfies strong stochastic transitivity (SST), medium stochastic transitivity (MST), or weak stochastic transitivity (WST), if $c_{\{x,y\}}(\{x\}) \geq 1/2$ and $c_{\{y,z\}}(\{y\}) \geq 1/2$ imply:

$$c_{\{z,x\}}(\{x\}) \geq \begin{cases} \max\{c_{\{x,y\}}(\{x\}), c_{\{y,z\}}(\{y\})\}, \\ \min\{c_{\{x,y\}}(\{x\}), c_{\{y,z\}}(\{y\})\}, \\ \frac{1}{2}, \end{cases} \tag{40}$$

respectively. Rieskamp, Busemeyer, and Mellers (40, 2006) propose an increasing "rationality" hierarchy going from WST through IIA to SST. Junnan and Natenzon (25, 2024) use MST to characterize Fechnerian utility functions. It can be shown that our GIIA axiom implies SST. But BM does not imply SST (Strzalecki 46, 2022, Example 3.17) and MST does not imply BM (Cohen and Falmagne 10, 1990, Figure 1).

Returning to Level-4 SCR's, the APU model (Fudenberg, Iijima, and Strzalecki 19, 2015), in its weak form, satisfies WST. The divisive normalization model (Steverson, Brandenburger, and Glimcher 45, 2019), even though it lies above APU, satisfies SST. Overall, the stochastic transitivity hierarchy branches off from our CPS-based hierarchy at Level 2 (GIIA) and appears to be a distinct organization of SCR's.

Appendix A: Proof of Theorem 3

We begin by building a map $f : \Lambda \to \Pi$. To do so, start with a total order $\rhd \in \Lambda$ and define a selection $s : \mathcal{G} \to \Omega$ by:

$$s(G) = !\, \alpha \text{ such that } \alpha \rhd G \backslash \{\alpha\}, \tag{A.1}$$

where $\alpha \rhd G\backslash\{\alpha\}$ means $\alpha \rhd \beta$ for all $\beta \in G\backslash\{\alpha\}$.

Theorem A.1. *The selection s defines a PCPS $G \mapsto p_G(\cdot)$ on $(\Omega, \mathcal{F}$.*

Proof. We need to show that if $E \subseteq F \subseteq G$ with $E \in \mathcal{F}$ and $F, G \in \mathcal{G}$, then $p_G(E) = p_G(F)p_F(E)$. Case (i): $p_G(E) = 1$, that is, $s(G) \in E$. Then $p_G(F) = 1$ by monotonicity, and since $F \subseteq G$, we certainly have:

$$s(G) \rhd \beta \;\forall\; \beta \in F \backslash \{s(G)\}, \tag{A.2}$$

so that $p_F(E) = 1$. Case (ii): $p_G(E) = 0$, that is, $s(G) \in F \backslash E$ or $s(G) \in G \backslash F$. In the first case, we find $p_F(E) = 0$, and in the second case we find $p_G(F) = 0$. □

Let $f : \Lambda \to \Pi$ be the map just constructed. Observe that, by construction, if $\rhd \in [\alpha, G]_\Lambda$, then $f(\rhd) \in [\alpha, G]_\Pi$.

Theorem A.2. *Suppose:*

$$\sum_{i=1}^{M} \mathbf{1}_{[\alpha_i, G_i]_\Pi} \leq \sum_{j=1}^{N} \mathbf{1}_{[\alpha_j, G_j]_\Pi}. \tag{A.3}$$

for integers M and N. Then:

$$\sum_{i=1}^{M} \mathbf{1}_{[\alpha_i, G_i]_\Lambda} \leq \sum_{j=1}^{N} \mathbf{1}_{[\alpha_j, G_j]_\Lambda}. \tag{A.4}$$

Proof. Fix an order $\rhd \in \Lambda$ such that (the numbering is without loss of generality):

$$\rhd \in [\alpha_i, G_i]_\Lambda \text{ for } i \leq m, \tag{A.5}$$

$$\rhd \notin [\alpha_i, G_i]_\Lambda \text{ for } i > m, \tag{A.6}$$

for some $m \leq M$. We need to show that there is a subset $\{j_1, j_2, \ldots, j_m\}$ of the index set $\{1, 2, \ldots, N\}$ such that:

$$\rhd \in [\alpha_{j_i}, G_{j_i}]_\Lambda \text{ for } i \leq m. \tag{A.7}$$

From Equation A.1 and the definition of f, the PCPS $p = f(\rhd)$ satisfies:

$$p \in [\alpha_i, G_i]_\Pi \text{ for } i \leq m. \tag{A.8}$$

By Inequality A.3, there are indices $j_1, j_2, \ldots, j_m$ (there is no loss of generality in using the same indices as above) such that:

$$p \in [\alpha_{j_i}, G_{j_i}]_\Pi \text{ for } i \leq m. \tag{A.9}$$

By the construction of p from $\rhd$, the relation A.9 holds if and only if:

$$\alpha_{j_i} \rhd G_{j_i} \backslash \{\alpha_{j_i}\}, \tag{A.10}$$

for each $i \leq m$, from which we have:

$$\rhd \in [\alpha_{j_i}, G_{j_i}]_\Pi \text{ for } i \leq m. \tag{A.11}$$

We conclude that $\rhd$ lies in at least m sets on the right-hand side of Inequality A.3, which establishes the theorem. □

The next observation is part of Rao and Rao (37, 1983, Theorem 3.1.9).

Theorem A.3. *Let $\mathcal{E}$ be a field and let μ be a finitely additive probability measure on $(\Omega, \mathcal{E})$. Then μ is a positive real charge on $(\Omega, \mathcal{E})$.*

Now fix a finitely additive probability measure P on $(\Lambda, \mathcal{B}_\Lambda)$, the collection $\mathcal{H}$ of subsets of Π defined by Equation 18 in the main text, and the function $Q : \mathcal{H} \to \mathbb{R}^+$ defined by Equation 19 in the main text. We restate Theorem 3 here and provide the proof.

Theorem A.4. *The function Q is a positive real partial charge on $\mathcal{H}$.*

Proof. Suppose Inequality A.3 holds. Then, by Theorem A.2, Inequality A.4 holds. By Theorem A.3, we know that P is a positive real partial charge on $(\Lambda, \mathcal{B}_\Lambda)$. Therefore, applying Inequality 17 in the main text:

$$\sum_{i=1}^{M} P([\alpha_i, G_i]_\Lambda) \leq \sum_{j=1}^{N} P([\alpha_j, G_j]_\Lambda), \tag{A.12}$$

from which

$$\sum_{i=1}^{M} Q([\alpha_i, G_i]_\Pi) \leq \sum_{j=1}^{N} Q([\alpha_j, G_j]_\Pi). \tag{A.13}$$

This establishes that Q is a positive real partial charge on $\mathcal{H}$, except if, instead of Inequality A.3, we have:

$$\mathbf{1}_\Pi + \sum_{i=1}^{M} \mathbf{1}_{[\alpha_i, G_i]_\Pi} \leq \sum_{j=1}^{N} \mathbf{1}_{[\alpha_j, G_j]_\Pi}, \tag{A.14}$$

or:

$$\sum_{i=1}^{M} \mathbf{1}_{[\alpha_i, G_i]_\Pi} \leq \sum_{j=1}^{N} \mathbf{1}_{[\alpha_j, G_j]_\Pi} + \mathbf{1}_\Pi. \tag{A.15}$$

But then the proof of Theorem A.2 can be re-run to show that:

$$\mathbf{1}_\Lambda + \sum_{j=1}^{N} \mathbf{1}_{[\alpha_j, G_j]_\Pi} \leq \sum_{j=1}^{N} \mathbf{1}_{[\alpha_j, G_j]_\Lambda}, \tag{A.16}$$

or:

$$\sum_{j=1}^{N} \mathbf{1}_{[\alpha_j, G_j]_\Pi} \leq \sum_{j=1}^{N} \mathbf{1}_{[\alpha_j, G_j]_\Lambda} + \mathbf{1}_\Lambda. \tag{A.17}$$

The rest of the argument proceeds as before, with Inequality A.16 or A.17 replacing Inequality A.4. If $\mathbf{1}_\Pi$ is added to both sides of Inequality A.3, this term cancels out. $\square$

Appendix B: Proof of Theorem 6

We begin by restating Theorem 7 in the main text and providing a proof.

Theorem B.1. *The family of maps $G \mapsto p_G(\cdot)$ defined in Equation 22 is a CPS relative to $\mathcal{G}$.*

Proof. The proof of Condition 1 of Definition 2 is immediate. For Condition 2, we need to show:

$$\frac{\mu_\gamma(E \cap G)}{\mu_\gamma(G)} = \frac{\mu_\gamma(F \cap G)}{\mu_\gamma(G)} \times \frac{\mu_\delta(E \cap F)}{\mu_\delta(F)}, \tag{B.1}$$

where $\gamma \in \Gamma$ is the unique index such that $0 < \mu_\gamma(G) < +\infty$ and $\delta \in \Gamma$ the unique index such that $0 < \mu_\delta(F) < +\infty$. Rewrite Equation B.1 as:

$$\frac{\mu_\gamma(E)}{\mu_\gamma(G)} = \frac{\mu_\gamma(F)}{\mu_\gamma(G)} \times \frac{\mu_\delta(E)}{\mu_\delta(F)}, \tag{B.2}$$

or

$$\mu_\gamma(E) \times \mu_\delta(F) = \mu_\gamma(F) \times \mu_\delta(E). \tag{B.3}$$

Now note that since $F \subseteq G$, we have $\mu_\delta(G) > 0$ and therefore $\delta \preceq \gamma$. Case (i): If $\delta = \gamma$, then Equation B.3 is immediately satisfied. Case (ii): If $\delta \prec \gamma$, then $\mu_\gamma(F) = 0$. But $\mu_\gamma(E) = 0$, since $E \subseteq F$, and then both sides of Equation B.3 are equal to 0, completing the proof. $\square$

We next outline the proof of Theorem 8 in the main text, which is Theorem 1 in Rényi (39, 1956). Given two events $G, H \in \mathcal{G}$, define the indicator of the two events by:

$$i(G, H) = \frac{p_{G \cup H}(G)}{p_{G \cup H}(H)}. \tag{B.4}$$

If $p_{G \cup H}(H) = 0$, set $i(G, H) = +\infty$. Also, note that nominator and denominator cannot both be 0. Define a relation on events in $\mathcal{G}$ by setting $G \sim H$ if $0 < i(G, H) < +\infty$, and say G and H are of the same dimension. Lemma 2 in Rényi (1956 (39)) establishes that $\sim$ is an equivalence relation, which is used to partition $\mathcal{G}$ into mutually disjoint classes $\mathcal{G}_\gamma$. For two distinct indices β and γ, write $\beta \prec \gamma$ or $\gamma \prec \beta$ if $i(G, H) = 0$ or $+\infty$, respectively, for $G \in \mathcal{G}_\beta$ and $H \in \mathcal{G}_\gamma$. (This definition can be shown to be independent of the representatives G and H chosen.) The relation $\prec$ is a total order.

The next step in the proof is to build measures μ_γ on $(\Omega, \mathcal{F})$, one for each equivalence class $\mathcal{G}_\gamma$, so that, if $G \in \mathcal{G}_\gamma$, then Equation 22 is satisfied for this μ_γ. The final step is to show that the system of measures $\{\mu_\gamma\}$ is dimensionally ordered, i.e., satisfies our Definition 11.

Csíszár (11, 1955) weakens the closure condition in Theorem 8 to the requirement that if $G, H \in \mathcal{G}$, then there is $J \in \mathcal{G}$ such that $G \cup H \subseteq J$ and $p_J(G) + p_J(H) > 0$. We do not make use of this condition in the current paper.

It is instructive to see how Theorem 8 fails if $\mathcal{G}$ is not closed under (finite) unions. Let $\Omega = \{x, y, z\}$, $\mathcal{G} = \{\{x, y\}, \{y, z\}, \{z, w\}\}$, and $p_{\{x,y\}}(\{y\}) = p_{\{y,z\}}(\{z\}) = p_{\{z,x\}}(\{x\}) = 1$. Then:

$$\frac{\mu_\alpha(\{y\})}{\mu_\alpha(\{x, y\})} = \frac{\mu_\beta(\{z\})}{\mu_\beta(\{y, z\})} = \frac{\mu_\gamma(\{x\})}{\mu_\gamma(\{z, x\})} = 1, \tag{B.5}$$

from which:

$$\mu_\alpha(\{y\}) > 0,\ \mu_\beta(\{z\}) > 0,\ \mu_\gamma(\{x\}) > 0, \tag{B.6}$$
$$\mu_\alpha(\{x\}) = 0,\ \mu_\beta(\{y\}) = 0,\ \mu_\gamma(\{z\}) = 0. \tag{B.7}$$

But then, by the definition of dimensional ordering, we have $\alpha < \beta$, $\beta < \gamma$, and $\gamma < \alpha$, contradicting total order.

We need one additional preliminary result. Recall the definition of the function $q : \Omega \times \mathcal{G} \times \mathcal{G} \to \mathbb{R}$ from Definition 8 in the main text. Using finitarity of our set-up, we can directly apply Theorem 3 in Falmagne (17, 1978) to obtain the following identity.

Theorem B.2. *For any $H \in \mathcal{G}$ and $\emptyset \neq G \subseteq H$:*

$$\sum_{\alpha \in H\backslash G} q(\alpha, G, H) = \sum_{\beta \in G} q(\beta, G\backslash\{\beta\}, H). \tag{B.8}$$

We now restate and prove Theorem 6.

Theorem B.3. *Fix a measurable space $(\Omega, \mathcal{F})$ and a finitary sub-family $\mathcal{G}$ of $\mathcal{F}$. If an SCR $G \mapsto c_G(\cdot)$ satisfies GIIA, then it satisfies BM.*

Proof. We first treat the case that G is a singleton and then treat the general case by successive substitution. Suppose $G = \{\omega\}$, so that:

$$q(\alpha, \{\omega\}, H) = (-1) \times c_H(\{\alpha\}) + (+1) \times c_{H\backslash\{\omega\}}(\{\alpha\}). \tag{B.9}$$

Since $c_H(\cdot)$ satisfies GIIA, the reverse direction of Theorem 1 implies that it is a CPS. Since $\mathcal{G}$ is closed under finite unions, Theorem 8 says that $c_H(\cdot)$ is dimensionally ordered. Therefore, form Equation 22:

$$c_H(\{\alpha\}) = \frac{\mu_{\gamma(H)}(\{\alpha\})}{\mu_{\gamma(H)}(H)}, \tag{B.10}$$

where, for any $F \in \mathcal{G}$, $\gamma(F)$ is the unique index for which $0 < \mu_\gamma(F) < +\infty$. Similarly:

$$c_{H\backslash\{\omega\}}(\{\alpha\}) = \frac{\mu_{\gamma(H\backslash\{\omega\})}(\{\alpha\})}{\mu_{\gamma(H\backslash\{\omega\})}(H\backslash\{\omega\})}. \tag{B.11}$$

Case (i): $\gamma(H) = \gamma(H\backslash\{\omega\}) = \gamma$. Then:

$$q(\alpha, \{\omega\}, H) = (-1) \times \frac{\mu_\gamma(\{\alpha\})}{\mu_\gamma(H)} + (+1) \times \frac{\mu_\gamma(\{\alpha\})}{\mu_\gamma(H\backslash\{\omega\})} \geq 0, \tag{B.12}$$

since $\mu_\gamma(H) \geq \mu_\gamma(H\backslash\{\omega\})$ by monotonicity. Case (ii): $\gamma(H\backslash\{\omega\}) \prec \gamma(H)$. Then:

$$\mu_{\gamma(H)}(H\backslash\{\omega\}) = 0, \tag{B.13}$$

so that, again using monotonicity, $\mu_{\gamma(H)}(\{\alpha\}) = 0$ and therefore $q(\alpha, \{\omega\}, H) \geq 0$, as required.

Now consider a general (finite) G. We write:

$$q(\alpha, G, H) = \sum_{j=0}^{|G|} (-1)^{|G|-j} \sum_{\{E \subseteq G : |E| = j\}} c_{H\backslash E}(\{\alpha\}) \tag{B.14}$$
$$= \sum_{j=0}^{|G|} (-1)^{|G|-j} \sum_{\{E \subseteq G : |E| = j\}} \frac{\mu_{\gamma(H\backslash E)}(\{\alpha\})}{\mu_{\gamma(H\backslash E)}(H\backslash E)}. \tag{B.15}$$

Observe that since $\{\alpha\} \subseteq H\backslash E$, we have $\gamma(\{\alpha\}) \preceq \gamma(H\backslash E)$. If $\gamma(\{\alpha\}) \prec \gamma(H\backslash E)$ for some E, then $\mu_{\gamma(H\backslash E)}(\{\alpha\}) = 0$. This means we can re-write Equation B.15 as:

$$q(\alpha, G, H) = \mu_{\gamma(\{\alpha\})}(\{\alpha\}) \times \sum_{j=0}^{|G|}(-1)^{|G|-j} \sum_{\substack{\{E\subseteq G:|E|=j\} \\ \{E\subseteq G:\gamma(H\backslash E)=\gamma(\{\alpha\})\}}} \frac{1}{\mu_{\gamma(\alpha)}(H\backslash E)}. \tag{B.16}$$

From this, we calculate:

$$\sum_{\alpha\in H\backslash G} q(\alpha, G, H) = \sum_{\alpha\in H\backslash G} \mu_{\gamma(\{\alpha\})}(\{\alpha\}) \times \sum_{j=0}^{|G|}(-1)^{|G|-j} \sum_{\substack{\{E\subseteq G:|E|=j\} \\ \{E\subseteq G:\gamma(H\backslash E)=\gamma(\{\alpha\})\}}} \frac{1}{\mu_{\gamma(\alpha)}(H\backslash E)} \tag{B.17}$$

$$= \mu_{\gamma(\{\alpha\})}(H\backslash G) \times \sum_{j=0}^{|G|}(-1)^{|G|-j} \sum_{\substack{\{E\subseteq G:|E|=j\} \\ \{E\subseteq G:\gamma(H\backslash E)=\gamma(\{\alpha\})\}}} \frac{1}{\mu_{\gamma(\alpha)}(H\backslash E)}. \tag{B.18}$$

If $q(\alpha, G, H) = 0$, there is nothing to prove, so we can assume $q(\alpha, G, H) \neq 0$. This tells us that the double sum on the right side of Equation B.16 is nonzero, so that, using $\mu_{\gamma(\{\alpha\})}(H\backslash G) > 0$ and Equation B.18, we know that $\sum_{\alpha\in H\backslash G} q(\alpha, G, H) \neq 0$. We can therefore divide Equation B.16 by Equation B.18 to get:

$$\frac{q(\alpha, G, H)}{\sum_{\alpha\in H\backslash G} q(\alpha, G, H)} = \frac{\mu_{\gamma(\{\alpha\})}(\{\alpha\})}{\mu_{\gamma(\{\alpha\})}(H\backslash G)}. \tag{B.19}$$

From this:

$$q(\alpha, G, H) = \frac{\mu_{\gamma(\{\alpha\})}((\{\alpha\})}{\mu_{\gamma(\{\alpha\})}(H\backslash G)} \times \sum_{\alpha\in H\backslash G} q(\alpha, G, H) \tag{B.20}$$

$$= \frac{\mu_{\gamma(\{\alpha\})}(\{\alpha\})}{\mu_{\gamma(\{\alpha\})}(H\backslash G)} \times \sum_{\beta\in G} q(\beta, G\backslash\{\beta\}, H) \tag{B.21}$$

$$= \frac{\mu_{\gamma(\{\alpha\})}(\{\alpha\})}{\mu_{\gamma(\{\alpha\})}(H\backslash G)} \times \sum_{\beta\in G}\Big[\frac{\mu_{\gamma(\{\beta\})}(\{\beta\})}{\mu_{\gamma(\{\beta\})}(H\backslash(G\backslash\{\beta\}))} \times \sum_{\beta\in H\backslash(G\backslash\{\beta\})} q(\beta, G\backslash\{\beta\}, H)\Big] \tag{B.22}$$

$$= \frac{\mu_{\gamma(\{\alpha\})}(\{\alpha\})}{\mu_{\gamma(\{\alpha\})}(H\backslash G)} \times \sum_{\beta\in G}\Big[\frac{\mu_{\gamma(\{\beta\})}(\{\beta\})}{\mu_{\gamma(\{\beta\})}(H\backslash(G\ \backslash\{\beta\}))} \times \sum_{\theta\in G\backslash\{\beta\}} q(\theta, G\backslash\{\beta,\theta\}, H)\Big] \tag{B.23}$$

$$= \cdots . \tag{B.24}$$

Here, Equality B.21 uses Theorem B.2. Equality B.22 comes from substituting in the analog, using β and $G\backslash\{\beta\}$ in place of α and G, to Equation B.20. Equality B.23 is another appeal to Theorem B.2. Noting that G is finite, we repeat this substitution process a finite number of times until the only term involving $q(\cdot,\cdot,\cdot)$ on the right side is of the form $q(\xi, \{\omega\}, H)$ for some singleton $\{\omega\} \subseteq G$ and $\xi \in H\backslash\{\omega\}$. We know from the first part of the proof that this term is non-negative. The rest of the right side will reduce to nested weights and sums, where all weights are positive, so we can conclude that $q(\alpha, G, H) \geq 0$, as required. □

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work the authors used ChatGPT o3 for copy-editing and bibliographic search. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

[1] Aaker, J., "The Negative Attraction Effect? A Study of the Attraction Effect Under Judgment and Choice," *Advances in Consumer Research* 18, 1991, 462-469.

[2] Bandyopadhyay, T., I. Dasgupta, and P. Pattanaik, "Stochastic Revealed Preference and the Theory of Demand," *Journal of Economic Theory*, 84, 1999, 95-110.

[3] Bandyopadhyay, T., I. Dasgupta, and P. Pattanaik, "Demand Aggregation and the Weak Axiom of Stochastic Revealed Preference," *Journal of Economic Theory*, 107, 2002, 483-489.

[4] Block, H., and J. Marschak, "Random Orderings and Stochastic Theories of Responses," in Olkin, I., S. Ghurye, W. Hoeffding, W. Madow, and H. Mann (eds.), *Contributions to Probability and Statistics*, Stanford University Press, 1960, 97-132.

[5] Brandenburger, A., P. Ghirardato, D. Pennesi, and L. Stanca, "Event Valence and Subjective Probability," 2025, at https://adambrandenburger.com/aux/material/evsp-03-10-25.pdf.

[6] Cerreia-Vioglio, S., D. Dillenberger, P. Ortoleva, and G. Riella, "Deliberately Stochastic", *American Economic Review*, 109, 2019, 2425-2445.

[7] Cerreia-Vioglio, S., P.O. Lindberg, F. Maccheroni, and M. Marinacci, "A Canon of Probabilistic Rationality," *Journal of Economic Theory*, 196, 2021, 105289.

[8] Clark, S., "The Random Utility Model with an Infinite Choice Space," *Economic Theory*, 7, 1996, 179-189.

[9] Cohen, M., "Random Utility – The Infinite Case," *Journal of Mathematical Psychology*, 22, 1980, 1-23.

[10] Cohen, M., and J. Falmagne, "Random Utility Representation of Binary Choice Probabilities: A New Class of Necessary Conditions," *Journal of Mathematical Psychology*, 34, 1990, 88-94.

[11] Csíszár, Á., "Sur la Structure des Espaces de Probabiliteé Conditionnelle," *Acta Math. Academiae Scientiarum Hungarica*, 6, 1955, 337-361.

[12] Dasgupta, I., and P. Pattanaik, "'Regular' Choice and the Weak Axiom of Stochastic Revealed Preference," *Economic Theory*, 31, 2007, 35-50.

[13] Davis-Stober, C., A. Marley, W. McCausland, and B. Turner, "An Illustrated Guide to Context Effects," *Journal of Mathematical Psychology*, 115, 2023, 102790.

[14] Debreu, G., "Review of 'Individual Choice Behavior' by R.D. Luce," *American Economic Review*, 1960, 50, 186-188.

[15] de Finetti, B., "Foresight: Its Logical Laws, Its Subjective Sources," *Annales de l'Institut Henri Poincaré*, 7, 1937.

[16] Dogan, S., and K. Yildiz, "Every Choice Function Is Pro-Con Rationalizable," *Operations Research*, 71, 2023, 1857-1870.

[17] Falmagne, J., "A Representation Theorem for Finite Random Scale Systems," *Journal of Mathematical Psychology*, 18, 1978, 52-72.

[18] Ford, L., and and D. Fulkerson, *Flows in Networks*, Princeton Landmarks in Mathematics and Physics, 56, Princeton University Press, 2015.

[19] Fudenberg, D., R. Iijima, and T. Strzalecki, "Stochastic Choice and Revealed Perturbed Utility," *Econometrica*, 83, 2015, 2371-2409.

[20] Gabaix, X., "Behavioral Inattention," in Bernheim, B.D., S. DellaVigna, and D. Laibson (eds.), *Handbook of Behavioral Economics - Foundations and Applications*, Volume 2, Elsevier, 2019, Chapter 4.

[21] Galichon, A., "On the Representation of the Nested Logit Model," *Econometric Theory*, 38, 2022, 370-380.

[22] Gul, F., P. Natenzon, and W. Pesendorfer, "Random Choice as Behavioral Optimization," *Econometrica*, 82, 2014, 1873-1912.

[23] Hildebrand, W., "Random Preferences and Equilibrium Analysis," *Journal of Economic Theory*, 3, 1971, 414-429.

[24] Huber, J., J. Payne, and C. Puto, "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research*, 9, 1982, 90-98.

[25] Junnan, H., and P. Natenzon, "Moderate Utility," *American Economic Review: Insights*, 6, 2024, 176-195.

[26] Kashaev, N., and V. Aguiar. "A Random Attention and Utility Model," *Journal of Economic Theory 204*, 2022, 105487.

[27] Kovach, M., and G. Tserenjigmid, "The Focal Luce Model," *American Economic Journal: Microeconomics* 14, 2022, 378-413.

[28] Kruis, J., G. Maris, M. Marsman, M. Bolsinova, H. van der Maas, "Deviations of Rational Choice: An Integrative Explanation of the Endowment and Several Context Effects," *Scientific Reports*, 10, 2020, at https://doi.org/10.1038/s41598-020-73181-2.

[29] Luce, R.D., *Individual Choice Behavior: A Theoretical Analysis*, Wiley, 1959.

[30] Maccheroni, F., M. Marinacci, and A. Rustichini, "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 2006, 1447-1498.

[31] Marschak, J., "Binary Choice Constraints on Random Utility Indicators, in Arrow (ed.), *Stanford Symposium on Mathematical Methods in the Social Sciences*, Stanford University Press, 1960, 312-329.

[32] Matejka, F., and A. McKay, "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105.1 (2015): 272-298.

[33] McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in Zarembka, P. (ed.), *Frontiers in Econometrics*, Academic Press, 1974, 105-142.

[34] McFadden, D., "Revealed Stochastic Preference: A Synthesis," *Economic Theory*, 26, 2005, 245-264.

[35] McFadden, D., and M. Richter, "On the Extension of a Set Function to a Probability on the Boolean Algebra Generated by a Family of Events, with Applications," working Paper No. 14, Mathematical Social Science Board Workshop on the Theory of Markets under Uncertainty, 1971, Department of Economics, UC Berkeley, 1971.

[36] McFadden, D., and K. Richter, "Stochastic Rationality and Revealed Stochastic Preference," in Chipman, J., D. McFadden, and K. Richter (eds.), *Preferences, Uncertainty, and Rationality*, Westview Press, 1991, 161-186.

[37] Rao, B., and B. Rao, *Theory of Charges: A Study of Finitely Additive Measures*, Academic Press, 1983.

[38] Rényi, A., "On a New Axiomatic Theory of Probability," *Acta Math. Academiae Scientiarum Hungarica*, 6, 1955, 285-335.

[39] Rényi, A., "On Conditional Probability Spaces Generated by a Dimensionally Ordered Set of Measures," *Teor. Veroyatnost. i Primenen.*, 1956, 1, 61-71.

[40] Rieskamp, J., J. Busemeyer, and B. Mellers, "Extending the Bounds of Rationality: Evidence and Theories of Preferential Choice," *Journal of Economic Literature*, 44, 2006, 631-661.

[41] Saito, K., "Axiomatizations of the Mixed Logit Model," 2018, at https://saito.caltech.edu/documents/3687/revised_mixed_logit11.pdf.

[42] Simonson, I., "Choice Based on Reasons: The Case of Attraction and Compromise Effects," *Journal of Consumer Research*, 16, 1989, 158-174.

[43] Simonson, I., "Vices and Virtues of Misguided Replications: The Case of Asymmetric Dominance," *Journal of Marketing Research*, 51, 2014, 514-519.

[44] Simonson, I., and A. Tversky, "Choice in Context: Tradeoff Contrast and Extremeness Aversion," *Journal of Marketing Research*, 29, 1992, 281-295.

[45] Steverson, K. A. Brandenburger, and P. Glimcher, "Choice-Theoretic Foundations of the Divisive Normalization Model," *Journal of Economic Behavior and Organization*, 164, 2019, 148-165.

[46] Strzalecki, T., *Stochastic Choice Theory*, 2022, at https://scholar.harvard.edu/sites/scholar.harvard.edu/files/tomasz/files/manuscript_01.pdf.

[47] Tversky, A., "Elimination by Aspects: A Theory of Choice," *Psychological Review*, 79, 1972, 288-299.

[48] Tversky, A., "Choice by Elimination," *Journal of Mathematical Psychology*, 9, 1972, 341-367.