Random Choice with Framing Effects: a Bayesian Model*

Evgenii Safonov†

This version: April 28, 2022, Latest version available here

Abstract

An expected utility maximizer learns various aspects of her preferences in different decision frames; each frame is defined as a collection of observable features of the choice environment. We interpret each feature as a Blackwell experiment (signal structure) associated with a state space that describes uncertainty of the agent's preferences, and consider an agent who updates her beliefs using the Bayes rule. An analyst observes the resulting stochastic choice of a population of agents with heterogeneous state-dependent utility functions. We show that almost any stochastic choice that admits a random utility representation within each frame is consistent with such model. However, when the state space has limited cardinality, the resulting stochastic choice should satisfy additional constraints on the sums of Block and Marschak (1960) polynomials constructed from choice frequencies under each decision frame.

Keywords: Framing, Random Choice, Expected Utility.

JEL: D01, D80, D83.

*I am grateful to my advisor, Wolfgang Pesendorfer, and to Faruk Gul for their encouragement, support and continuing guidance in developing this project. I'd like to thank Leeat Yariv for careful reading of the paper and insightful comments. For helpful comments and suggestions I would like to thank David Dillinberger, Mira Frick, Ryota Iijima, Svetlana Kosterina, Lasse Mononen, Stephen Morris, Pietro Ortoleva, Franz Ostrizek, Pellumb Reshidi, Denis Shishkin, Rui Tang, Kemal Yildiz, Mu Zhang and seminar audience in Princeton University.

†Department of Economics, Princeton University.

1 Introduction

This paper interprets decision frames as emerging from different unobservable information structures that signal about the values of the alternatives in the decision problem faced by the agent. The resulting distribution over posterior beliefs induces a stochastic choice that admits a random utility representation within each frame. A question that we address is: is there a connection between choice frequencies *across* various decision frames?

We motivate our analysis by two interpretation of this informational model of framing effects. The first interpretation is that the relevant state space encodes uncertainty regarding the intrinsic preference of the decision maker. Imagine a person who does not know precisely some aspect of her true preferences, such as the extent of her fear of height. She contemplates between spending a weekend at the seashore or going to a mountain trip. Her decision might then depend on a seemingly irrelevant detail: whether the day before, the featured Wikipedia article has been "Parachuting" or not. In the first case, being a Wikipedia fan, the agent reads about parachuting. Although the article does not provide any objective information regarding the benefits or costs of going to a particular seashore or mountain trip, reading about parachuting triggers the agent to think more about her own attitude toward great heights. As a result, the distribution of beliefs regarding the extend of fear or love of heights in the population of agents would be different depending on whether such learning of own preferences has been triggered or not, resulting in different frequencies of vacation choices.

The second interpretation is that the relevant state space encodes some objective unobservable properties of the choice problem that matter for the decision making. For instance, suppose a salesman can be either honest or dishonest; the true value of the items offered by the salesman depends on his type. A buyer acquires information about the type of the salesman by observing fine details of his facial expression during a conversation. One decision frame in this example corresponds to buyer wearing her glasses, while the other—to buyer not wearing her glasses, and hence, having lower ability to differentiate between salesman's types.

The primitive of our model is an *extended random choice* that consists of choice frequencies from different menus of items under various decision frames. Each frame is characterized by a collection of observable features that we call *frame details*. Our model of Bayesian Agents with Uncertain Preferences (BAUP) introduces a state space with a commonly known prior and signal structures associated with each frame detail; the agents in the population have heterogeneous state-dependent utility functions. We say that a BAUP model represents an extended random choice if the probability that the agent from the population receives a signal that suggest to choose a certain item from the menu is equal to the corresponding observed frequency, and this holds for all decision frames.

Our results show that whether there is a connection between choice frequencies across different frames depends on the cardinality of the state space. We first analyse the case when the relevant state space could be arbitrary. Our first main result, Theorem 1, says that an extended random choice could be approximated by a BAUP model—meaning that there exists a model that induces the choice probabilities arbitrary close to the original extended random choice—if and only if the random choice within each frame admits a random utility representation. Thus, a BAUP model places almost no restrictions on choice frequencies observed under different decision frames. One corollary of this result is that the analyst cannot infer the relative informativeness of the decision frames from the observed extended random choice. Another application is to the dynamic random choice, where Theorem 1 replicates the "almost everything goes" result that could be inferred from the analysis made in Frick, Iijima, and Strzalecki (2019).

We next analyse the case when the relevant state space is fixed. As an example, consider a stylized experiment where a population of agents chooses from menus of alternatives that provide different consequences depending on the random binary state $s \in \{1,2\}$. There are three alternatives: x, y, and z. Item $x = (x_1, x_2)$ provides consequence x_1 in the first state of the world and x_2 in the second state, and similarly do items $y = (y_1, y_2)$ and $z = (z_1, z_2)$. There are 3 treatments in the experiment that differ in terms of information about the state revealed to the agents. An analyst observes choice frequencies of a population of agents from different subsets of items x, y, z. For each agent, first, Nature draws a state of the world $s \in \{1,2\}$, and then reveals a random signal informative about the state according to the treatment's signal structure, after which the agent makes a choice.

Let $\rho_i^A(a)$ be the frequency with which item a is chosen from menu A under treatment i, and suppose $\rho_1^{\{x,y\}}(x) = 0.8$, $\rho_1^{\{x,y,z\}}(x) = 0.1$, $\rho_2^{\{y,z\}}(y) = 0.7$, and $\rho_2^{\{x,y,z\}}(y) = 0$. Therefore, under treatment $i = 1$, the resulting random choice admits a random utility representation that places probability $\rho_1^{\{x,y\}}(x) - \rho_1^{\{x,y,z\}}(x) = 0.7$ that the agent prefers z to x to y (we write it as $z \succ x \succ y$). Under treatment $i=2$, a random utility representation should place probability $\rho_2^{\{y,z\}}(y)-\rho_2^{\{x,y,z\}}(y)=0.7$ that the agent's preference is $x \succ y \succ z$. Thus, at least some agents in the population exhibit preference reversal—placing item z to the top or to the bottom of the preference order depending on the treatment.

The stochastic choice behavior in this example can be consistent with expected utility maximization and Bayesian updating. For instance, suppose that utility values of items in states 1 and 2 are $x=(-1,9)$, $y=(0,0)$, and $z=(5,-15)$. In this case, depending on the belief $p=Pr(s=2)$, the preference is: $z \succ y \succ x$ if $p<0.1$, $z \succ x \succ y$ if $p \in (0.1,0.2)$, $x \succ z \succ y$ if $p \in (0.2,0.25)$, and $x \succ y \succ z$ if $p>0.25$. Let the prior be $\pi=0.233$, and treatments 1 and 2 induce distributions over posteriors $P_1=0.2\delta_0+0.7\delta_{0.19}+0.1\delta_1$ and $P_2=0.1\delta_0+0.2\delta_{0.22}+0.7\delta_{0.27}$. The resulting random utility model under treatment 1 places probability 0.2 on $z \succ y \succ x$, 0.7 on $z \succ x \succ y$, and 0.1 on $x \succ y \succ z$. Similarly, the resulting random utility model under treatment 2 places probability 0.1 on $z \succ y \succ x$, 0.2 on $x \succ z \succ y$, and 0.7 on $x \succ y \succ z$, consistent with the observed choice frequencies.

In fact, if we are concerned only with choice frequencies under treatments 1 and 2, almost any data could be explained by such informational model. However, our results show that, if the observed choice frequencies under the first two treatments are as in our example, and the agents in the population are expected utility maximizers, it must be the case that under treatment 3, $\rho_3^{\{x,z\}}(z)-\rho_3^{\{x,y,z\}}(z)\leq 0.6$. This restriction holds for a population of heterogeneous agents as well, and it does not depend even on the assumption of Bayesian updating. For this result, it is sufficient to know that the agents are expected utility maximizers, and the relevant state space is binary.

We provide two type of explicit constraints specific for the binary state space in Proposition 2. For instance, the constraint mentioned in the example above is given by

$$\rho_2^{\{y,z\}}(y)-\rho_2^{\{x,y,z\}}(y)+\rho_3^{\{x,z\}}(z)-\rho_3^{\{x,y,z\}}(z)+\rho_1^{\{x,y\}}(x)-\rho_1^{\{x,y,z\}}(x) \;\leq\; 2$$

Our next main results, Theorems 2 and 3, generalize Proposition 2 and formulate a way to calculate a set of constraints necessary for a BAUP representation with a given state space S. Theorem 2 does it indirectly, requiring the existence of a family of random utility representations coherent in a certain way across frames. Theorem 3, in contrast, formulates a set of explicit constraints. It requires to calculate, from the frequencies of choices within each frame, specific Block and Marschak (1960) polynomials, and restricts the values of certain sums of these polynomials. In Corollary 3, we show that for any fixed state space, if the set of alternatives and the set of frames are rich enough, these additional constraints do have a bite: there is an extended random choice that admits random utility representations within each frame, but fails to be approximated by a BAUP model with the given state space.

The rest of the paper is organized as follows. In the remaining part of the Introduction, we discuss the relevant literature. In Section 2, we develop a BAUP model and provide results for an unrestricted state space. In Section 3, we analyze a BAUP model with a fixed state space. All proofs are given in the Appendix.

1.1 Related Literature

This paper contributes to the literature on random utility models, random choice with unobservable information, and to the literature on framing effects.

A random utility model has been considered in Block and Marschak (1960) where the authors have introduced necessary conditions in terms of linear inequalities on the choice frequencies that later has been called the *Block-Marschak polynomials*. In Falmagne (1978), it has been proven that these conditions are also sufficient; this result has been then discussed in Economics literature in Barberá and Pattanaik (1986). In our model, the random choice frequencies are consistent with the random utility model within each frame, and the Block-Marschak polynomials play an important role in our analysis when the state space has fixed cardinality.

Tversky and Kahneman (1981) introduced a notion of framing that serves as a benchmark that enables the agents to identify some outcomes as *gains* and others as *losses*. In our paper, the framing triggers a specific learning process and is associated with some unobservable signal structure. Framing effects has been studied in a great number of papers, for instance in Huber, Payne, and Puto (1982), Tversky and Simonson (1993), and Salant and Rubinstein (2008); our notion of the *extended random choice* could be viewed as a stochastic counterpart of the Salant and Rubinstein (2008)'s "extended choice". The latter consists of deterministic choices from menus under various decision frames. Menu-dependent stochastic choice is considered in Manzini and Mariotti (2014).

Preference uncertainty has been discussed in Kreps (1979) and Dekel, Lipman, and Rustichini (2001) in the context of choice over menus and preference for flexibility; in Karni and Safra (2016) in the setup of "mental acts". In Ahn and Sarver (2013), the authors connect the model of Dekel et al. (2001) and the random expected utility of Gul and Pesendorfer (2006) to identify probabilities over subjective state space and utilities. In Dillenberger, Lleras, Sadowski, and Takeoka (2014), the authors show how the analyst can recover the unobservable private information from the preference relation on the set of menus of acts. In our model, the analyst has access to a much less rich choice data.

In Frick et al. (2019), the authors consider a dynamic random choice in presence of unobservable private information, where the agents choose from lotteries of today's consumption and the continuation menu. The most close to our paper is their analysis of the model of "Bayesian Evolving Beliefs" with restriction to "Atemporal consumption problems" with "Consumption inertia and learning".[1] There are several differences between their and our setup: first, in their model, the analyst observes correlations between choices in various periods, while in our model, only marginal distributions within each frame are observable. However, the proof of our Theorem 1, in fact, allows us to extend the theorem to the setup where the analyst observes the joint distribution of choices under various decision frames as well.[2] Second, in Frick et al. (2019), the lotteries over consumption (and future menus) are considered, while in our model, we focus on deterministic consumption.[3] Third, our setup allows us to consider a more general assumptions on how the agent's private information is connected across the decision frames; an application to a dynamic choice is a special case of our model. Finally, we analyse the implications of a known size of the relevant state space, while in Frick et al. (2019), the authors do not constraint the state space.

In Lu (2016), the author considers a stochastic choice from menus of Anscombe-Aumann acts defined on the objective state space. He assumes that the choice results from an unobservable private information and shows that the analyst can infer the agent's private information from her choices. In particular, the analyst can figure out if one agent is better informed than another.[4] In our model, the relevant state space is subjective, and the utilities of the alternatives depend on

[1] Sections 4.2 and 6.2 of Frick et al. (2019) correspondingly.

[2] We consider marginal probabilities in our model because it is a more natural primitive; in general, it is hard to justify the observability of choices of the same agent under different decision frames if the set of these frames within the model does not admit a linear Blackwell order.

[3] Thus, our model extends the random utility model of Falmagne (1978), while Frick et al. (2019) is extends the random expected utility model of Gul and Pesendorfer (2006).

[4] In the sense of the Blackwell information order.

the state. An application of our Theorem 1 shows that in our setup, in contrast to insights from Lu (2016), the analyst almost always cannot make any inference about the informativeness of the signals associated with the particular decision frame from the observed extended random choice.

In Natenzon (2019), the author considers a decision maker who learns her preferences when she encounters a given choice problem. In his setup, signals are in the form of an additive Gaussian noise associated with each of the "real" alternatives that can be potentially chosen and each of the "phantom" alternatives in the menu that cannot be chosen. Equivalently, each "real menu" consisting of the "real alternatives" can be encountered under one of the decision frames that correspond to different sets of "phantom" alternatives (decoys). Our paper complements Natenzon (2019)'s analysis by considering a general non-parametric setup and a more general set of frames at the expense of having less sharp predictions.

Several papers study the role of unobserved information structures in explaining the realized choices/deterministic preferences as oppose to the stochastic choice primitive, considered in our model. In the setup considered in Shmaya and Yariv (2016), the authors show that the Bayesian updating has a bite only when under all signal realizations, the agent chooses the same option; the counterpart of this result in the probabilistic world would say that the Bayesian updating put restriction only on the degenerate choice frequencies, which mirrors "almost everything goes" result in our Theorem 1. In Piermont (2017), the author studies a family of preference relations on the set of Anscombe-Aumann acts indexed by a "menu"; the "menu" serves as the set of potential consequences of the considered acts. He assumes that the agent has state-dependent utility function and that each menu is associated with an unobservable signal realization. Thus, in his model, the state of the world may have arbitrary correlation with the menu, which yields a version of "everything goes" result. In our model, in contrast, for any menu, the probability of the state of the world and the signal structures associated with each decision frame remain the same; thus, the decision frames are completely disentangled from the menus. Finally, a contemporary study Deb and Renou (2021) characterises, under various assumptions on the signal structures and state-dependent utilities, the set of possible realizations of the sequences of choices of a population of agents who receive the same unobservable signals regarding the state of the world.

2 Model

2.1 Primitive

There is a finite non-empty universal set of alternatives, X. A menu is a non-empty subset of X. Denote by $\mathcal{X}$ the set of all menus; thus, $\mathcal{X} = 2^X \backslash \{\varnothing\}$. A finite non-empty set D is a set of *frame details*. We interpret a frame detail as an observable by the analyst feature of the environment that may be present or absent in the particular choice problem—such as a certain article in Wikipedia being featured or not. The set D, hence, is the set of all such features. We call a *decision frame* a subset of frame details $f \subseteq D$ and use letters $f, g, \ldots$ for a generic frame. We denote by F the set of frames that the agent may face; thus, $F \subseteq 2^D$, $F \neq \varnothing$.

A primitive of the model is a family of probability distributions $\{\rho_f^A\}_{A\in\mathcal{X}, f\in F}$ that, for each menu $A \in \mathcal{X}$ and each frame $f \in F$, gives the probability $\rho_f^A(a)$ that the agent chooses alternative $a \in A$ from menu A under the decision frame f. Thus, $\text{supp}(\rho_f^A) \equiv \{x \in X | \rho_f^A(x) > 0\} \subseteq A$. We call ρ an *extended random choice*.

2.2 Bayesian Agents with Uncertain Preferences

A model of Bayesian Agents with Uncertain Preferences (BAUP) consists of a tuple (S, π, σ, μ), where S is a finite state space, π is a prior probability distribution on S, $\sigma : D \to \Sigma(S, \mathbb{M})$ is a mapping from the set of frame details D to the set $\Sigma(S, \mathbb{M})$ of Blackwell experiments (signal structures) with state space S and some finite message space $\mathbb{M}$, and μ is a finite-support[5] probability distribution over state-dependent utility functions $u : S \times X \to \mathbb{R}$. For convenience, we call σ a *framing function*.

Thus, there is a population of agents who share a common prior regarding the state and interpret the details of the framing as signals regarding the state. Different types of agents in the population have different state-dependent utility functions[6].

Denote by M_d the random signal, associated with the Blackwell experiment $\sigma(d)$, and by m_d

[5]Considering infinitely many population types does not change the analysis—this follows from the fact that the primitive is a finite data set.

[6]An alternative models would allow for a heterogeneity in the prior beliefs and/or in the signal structures as well. The corresponding analysis leads to similar results. We consider a less flexible model that still allows to represent a wide variety of primitives.

its generic realization. Denote by $\sigma(d)(s,m_d)$ the probability of signal realization m_d conditional on state $s \in S$; thus,

$$\sigma(d)(s,m) = Pr(M_d = m \mid s)$$

If the agent with state-dependent utility function u receives a collection of signals $m = \{M_d = m_d\}_{d\in D'}$ for some $D' \subseteq D$, she forms a posterior belief p using Bayes formula

$$p^m(s) = \frac{\pi(s)Pr\Big(\{M_d = m_d\}_{d\in D'} \mid s\Big)}{\sum_{t\in S}\pi(t)Pr\Big(\{M_d = m_d\}_{d\in D'} \mid t\Big)} \tag{1}$$

Denote by $u^m : X \to \mathbb{R}$ the corresponding expected utility function, that is:

$$u^m(x) \equiv \mathbb{E}[u(x) \mid m] = \sum_{s\in S} p^m(s)u_s(x) \tag{2}$$

We assume that signals associated with different frame details are independent.

CONDITION 1 (Conditional independence of signals):

$$Pr\Big(\{M_d = m_d\}_{d\in D} \mid s\Big) = \prod_{d\in D} Pr\Big(M_d = m_d \mid s\Big) \tag{3}$$

2.3 Representation

Consider a BAUP model (S,π,σ,μ). Given a subset of frame details $D' \subseteq D$, denote by $M^{D'} = \{M_d\}_{d\in D'}$ the collection of random signals associated with Blackwell experiments $\sigma(d)$ for $d \in D'$. We say that a BAUP model (S,π,σ,μ) *represents* an extended random choice ρ if for all $A \in \mathcal{X}$, $f \in F$, and $x \in A$,

$$\rho_f^A(x) = \sum_u \mu(u)\cdot Pr\Big(\Big\{u^{M(f)}(x) \geq u^{M(f)}(y)\ \forall y \in A\Big\}\Big) \tag{4}$$

where the summation is performed over all state-dependent utility functions $u : S \times X \to \mathbb{R}$ such that $\mu(u) > 0$, and $u^{M(f)}(\cdot)$ is a function given in eq. (2) where signal $M(f)$ associated with frame f is random and distributed according to

$$Pr\Big(M(f) = \{m_d\}_{d\in f}\Big) = \sum_{s\in S}\pi(s)\cdot Pr\Big(M(f) = \{m_d\}_{d\in f}\ \Big|\ s\Big) = \sum_{s\in S}\pi(s)\cdot\prod_{d\in D}\sigma(d)(s,m_d) \tag{5}$$

Note that the representation implies that the maximizer of the expected utility for each menu under each frame is unique with probability one; in other words, ties in the expected utility occur with probability zero.

We offer two interpretations of a BAUP model. In the first interpretation, each agent in the population does not know perfectly her own preferences regarding alternatives in X. We capture this uncertainty via a state $s \in S$; the agent knows her utility $u_s(x)$ of alternative x depending on s. Absent further information, she uses prior π regarding the state s. A frame detail d reveals some aspect of the agent's preferences in the form of a signal M_d with conditional distribution $Pr(M_d = m|s) = \sigma(d)(s,m)$; a frame consisting of several frame details reveals the corresponding aspects of the agent's preferences. Facing some menu under a particular frame, the agent chooses the best alternative from this menu, calculating expected utility according to her posterior belief about her preferences. The stochasticity of the observed choices comes from three sources. Two of them emerge from the population heterogeneity: a random state-dependent utility type u—known by the agent and distributed according to $\mu(\cdot)$—and a random sub-type $s \in S$—unknown to the agent and distributed according to $\pi(\cdot)$. The third source of stochasticity is a random signal $M(f)$ that induces a random posterior distribution $p^{M(f)}(\cdot)$ of agent's beliefs regarding her sub-type $s \in S$. Implicitly in the representation, we require coherency of beliefs regarding s, because prior π plays role both in the agent's updating rule (eq. 1) and in the calculation of the resulting extended random choice via formulas (4), (5).

In the second interpretation, the unknown state of the world $s \in S$ represents conditions that matter for the optimal choice; each choice problem, thus, belongs to one of $|S|$ categories, but agents do not know precisely what particular category they face. Each agent has a population type given by her state-dependent (category-dependent) utility function u. Frame details reveal information about the relevant category in the form of a random signal, and the agent evaluates an alternative by its expected utility across different categories of problems. Thus, the three sources of stochasticity in the representation are: first, a random category $s \in S$, second, a random agent type u, and third, a random signal $M(f)$ regarding the relevant category of the choice problem.

Following the literature, we say that a random choice that provides choice frequencies for various menus $A \in \mathcal{X}$ has a random utility representation if there is a distribution of linear orders[7] over X such that the probability that item $x \in A$ is the highest-rank alternative in menu A is equal to the probability with which x is chosen from A. It is straightforward to see that within each frame, a BAUP model reduces to a random utility model:

[7] A linear order over X is a complete transitive antisymmetric binary relation $\succeq \subseteq X \times X$. We use the symbol $\succ$ to denote the asymmetric part of $\succeq$.

Proposition 1. *If an extended random choice ρ has a BAUP representation, then for each frame $f \in F$, the random choice ρ_f has a random utility representation.*

In particular, in a special case when the state space is singleton ($|S| = 1$), a BAUP model reduces a single random utility model:

Corollary 1. *If $|S| = 1$, then an extended random choice ρ has a BAUP representation if and only if the following two conditions hold:*

(1) $\rho_f = \rho_g$ for all $f, g \in F$;

(2) The random choice ρ_f has a random utility representation for all $f \in F$.

Recall that a classical random utility model is characterized in terms of the system of linear inequalities constructed recursively from the observed random choice[8]. Given an extended random choice ρ_f^A, let us define the following *Block-Marschak* linear polynomials recursively:

$$q_f(x, \varnothing) = \rho_f^X(x) \ \ \forall x \in X, \qquad q_f(x, A) = \begin{cases} \rho_f^{X \setminus A}(x) - \sum\limits_{B \subsetneq A} q_f(x, B) & \text{if } x \notin A \\ 0 & \text{if } x \in A \end{cases} \tag{6}$$

Note that the summation in the formula above is performed over all subsets B of the set X that are not subsets of the set A. If ρ_f has a random utility representation, the Block-Marschak polynomial $q_f(x, A)$ gives the probability that set A is a strict upper counter set of item x; thus, if a BAUP model represents ρ, it must be that $q_f(x, A) \geq 0$ for all x, A, f. Our first main result says that these conditions are *almost* sufficient for the representation in the sense, that a BAUP model can approximate choice frequencies arbitrary close when these conditions hold.

DEFINITION 1. An extended random choice ρ has a BAUP approximation if for any $\epsilon > 0$ there exists an extended random choice τ such that τ is represented by a BAUP model, and $|\rho_f^A(x) - \tau_f^A(x)| < \epsilon$ for all $A \in \mathcal{X}$, $x \in X$, $f \in F$.

Theorem 1. *An extended random choice ρ has a BAUP approximation if and only if $q_f(x, A) \geq 0$ for all $f \in F$, $A \in \mathcal{X}$, $x \in A$, where q are Block-Marschak polynomials defined in eq. (6). Moreover, if $q_f(x, A) > 0$ for all $f \in F$, $A \in \mathcal{X}$, $x \notin A$, then ρ has a BAUP representation.*

[8] Falmagne (1978), Barberá and Pattanaik (1986)

Corollary 2. *An extended random choice ρ has a BAUP approximation if and only if for each $f \in F$, the random choice ρ_f has a random utility representation.*

Theorem 1 is a negative result: a BAUP, as a general informational model of framing, (almost) does not put restrictions on choice frequencies across different frames. Moreover, it is clear from the proof of Theorem 1, that it continues to hold in case if the primitive of the model is not a family of marginal distributions, but a joint distribution of choices across all frames instead. The generality of our primitive—in particular, the flexibility of the set of frame details D and the set of observable frames F—allows us to use Theorem 1 to formulate this negative result in several contexts.

First, suppose F consists of singleton frames $f = \{d_f\}$. Then we may interpret Theorem 1's message as the statement that hidden belief updating that starts from a common prior *almost* does not bring constraints on the resulted frequencies of random choices—in the sense that any random choice that admits a random utility model within one frame can be approximated by BAUP models.

Second, considering a set of frames $F = \{f_i\}_{i=1}^N$ that is nested: $f_i = \{1, 2, ..., i\}$, we can interpret an extended random choice as a dynamic random choice where the agent gradually learns some payoff-relevant information. Thus, a similar "almost everything goes" result holds for a dynamic random choice with unobservable gradual learning—a result that one could infer from Proposition 6 of Frick et al. (2019)[9] by restricting the menus to include only deterministic options and restricting the set of menus in each period to be determined in period 1 (period 0 in their model).

Third, assume the analyst knows that some frames are more informative than others. Thus, the analyst knows a transitive[10] binary relation $\succeq_{inf} \in F \times F$ on the set frames. It turns out, there is a structure of frame details that is coherent with $\succeq_{inf}$ in the sense that $f \succeq_{inf} g$ implies $g \subseteq f$. For instance, this holds when $D = \{d_f\}_{f \in F}$, and $f = \{d_f\} \cup \{d_g \in D | f \succeq_{inf} g\}$. In this context, Theorem 1 says that for a generic extended random choice consistent with random utility maximization within each frame, it is (almost) impossible to infer which frame is more informative than another—a conclusion, opposite to the results in the setup considered in Lu (2016). This follows from the fact that the set of constraints in Theorem 1 is invariant with respect to all permutations of the set of frames $F \to F$.

[9]Their primitive is dynamic stochastic choice from menus of lotteries of items and the continuation menu.

[10]Since the Blackwell order is transitive.

3 Fixed State Space

In this section, we consider the implications of the knowledge of the relevant state space. Most naturally, this corresponds to the interpretation of a BAUP model in the objective sense; that is, there are $|S|$ categories the represent various objective features of the choice situations that the agent may face. The second interpretation is subjective, in which case we are interested in the fixed set of the agent's states of the mind—for instance, whether the agent fears heights or not.

3.1 Binary State Space

Let us start our analysis with the binary state space, $S = \{1,2\}$, and let $\{S,\pi,\sigma,\mu\}$ be a BAUP model. As it turns out to be, once we restrict the state space to be binary, there are more conditions necessary for a BAUP representation (approximation) than those stated in Theorem 1. In this section, we first provide calculations that illustrate reasons behind the emergence of these new constraints that are specific to the binary state space, and then formalize these constraints in Proposition 2. In the next section, we discuss how this intuition generalizes to non-binary state spaces.

Consider some state-dependent utility u that occur with positive probability; that is, $\mu(u) > 0$. Consider items $x,y,z \in X$. Since we assume away ties in the BAUP model[11], it is without loss to focus only on the linear orders induced by the expected utility function. Let $U^{\lambda}(a) = (1-\lambda)u_1(a) + \lambda u_2(a)$ be the expected utility of item a when the agent belief is $p(2) = \lambda$. If it is not the case that one of x,y is always preferred to the other, there is some belief λ_{xy} that divides the set of beliefs $[0,1]$ to the range of beliefs for which x is preferred to y and the range of beliefs for which y is preferred to x.

Let λ_{xz} and λ_{yz} be analogous beliefs that render items x and z, and items y and z correspondingly indifferent for the agent. The three points λ_{xy}, λ_{xz}, λ_{yz} divide the space of beliefs $[0,1]$ by at most 4 distinct regions; for each of these regions, the expected utility U^{λ} induces one linear order over x,y,z; the total number of such linear orders is, thus, no more than 4. Let $O = \{xyz, xzy, yxz, yzx, zxy, zyx\}$ be the set of all linear orders over alternatives x,y,z, where term

[11] According to the representation given by eq. (4).

xyz means a linear order $x \succ y \succ z$, etc. Given a state-dependent utility u, let

$$O(u) \;=\; \left\{ \succeq \in O \;\middle|\; \exists \lambda \in [0,1]: \; a \succ b \iff U^{\lambda}(a) > U^{\lambda}(b) \; \forall a,b \in \{x,y,z\} \right\}$$

where the symbol $\succ$ denotes the asymmetric part of the linear order $\succeq$, and let

$$\Theta \;=\; \{\; O(u) \mid u \text{ is some state-dependent utility function } \}$$

One can figure out that Θ consists of the sets $\theta_1 = \{xyz, xzy, zxy, zyx\}$, $\theta_2 = \{xyz, yxz, yzx, zyx\}$, several subsets of θ_1, θ_2, and the sets that could be constructed from θ_1 and θ_2 and these subsets by permutations of x, y, z. The expected utilities that give rise to types θ_1 and θ_2 are illustrated in Figure 1. For instance, for θ_1, we have $0 < \lambda_{yz} < \lambda_{xz} < \lambda_{xy} < 1$, while for θ_2, we have $0 < \lambda_{xy} < \lambda_{xz} < \lambda_{yz} < 1$.

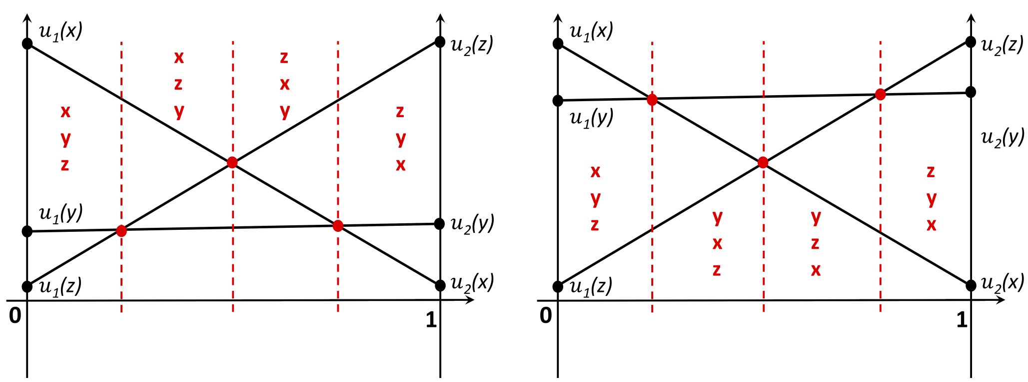

Figure 1: Subsets of linear orders $\theta_1 = \{xyz, xzy, zxy, zyx\}$ (left plot) and $\theta_2 = \{xyz, yxz, yzx, zyx\}$ (right plot) induced by the maximization of expected utility with beliefs $\lambda \in [0,1]$.

Thus, each type $\theta \in \Theta$ is a set of linear orders that could be exhibited by an agent with fixed state-dependent utility function under various frames. Let us now think about the extended random choice induced by a population of such agents. Take some set of linear orders $\psi \subseteq O$, and suppose that for all $\theta \in \Theta$, $\psi \not\subseteq \theta$. In other words, a subset of linear orders ψ is such that each possible type θ of agent misses at least one of the linear orders in ψ—that is, for all $\theta \in \Theta$, there is $\succeq \in \psi$ such that $\succeq \notin \theta$. One example of such set of linear orders is $\psi = \{xyz, yzx, zxy\}$. Indeed, one can see that $yzx \notin \theta_1$, $zxy \notin \theta_2$, and similarly for the other $\theta \in \Theta$.

Let us analyse further $\psi = \{xyz, yzx, zxy\}$. Consider an agent with state-dependent utility u such that $\mu(u) > 0$, and, for an arbitrary frame f, let p^u_f be the distribution of linear orders induced by the expected utility of this agent, that is, by a BAUP model (S,π,σ,δ_u) within this frame. Note that since $O(u)$ misses at least one of the linear orders xyz, yzx, zxy, it must be the case that at least one of the probabilities $p^u_f(xyz)$, $p^u_f(yzx)$, $p^u_f(zxy)$ is zero for all frames $f \in F$. It follows that for arbitrary frames $f,g,h \in F$,

$$p^u_f(xyz) + p^u_g(yzx) + p^u_h(zxy) \ \leq \ 2$$

Summing this inequality over all agents in the population with weights $\mu(u)$, we get

$$p_f(xyz) + p_g(yzx) + p_h(zxy) \ \leq \ 2$$

where p_f for $f \in F$ is a distribution of linear orders induced by a BAUP model (S,π,σ,μ). To formulate this constraint in terms of the observed extended random choice, note that $p_f(xyz) = \rho^{\{y,z\}}_f(y) - \rho^{\{x,y,z\}}_f(y)$, etc. For an arbitrary $f,g,h \in F$ and $x,y,z \in X$, denote by

$$W^{xyz}_{fgh} \ = \ \rho^{\{y,z\}}_f(y) - \rho^{\{x,y,z\}}_f(y) + \rho^{\{x,z\}}_g(z) - \rho^{\{x,y,z\}}_g(z) + \rho^{\{x,y\}}_h(x) - \rho^{\{x,y,z\}}_h(x) \tag{7}$$

then $W^{xyz}_{fgh} \leq 2$ is a necessary condition for a BAUP representation.

Continuing this analysis, one can proceed with $\psi = \{xzy, yxz, yzx, zxy\}$ and see that it also has the property that $\psi \not\subseteq \theta$ for all $\theta \in \Theta$. Repeating the same arguments, we consider, for arbitrary $f,g,h,j \in F$ and $x,y,z \in X$,

$$Q^{xyz}_{fghj} \ = \ \rho^{\{y,z\}}_f(z) - \rho^{\{x,y,z\}}_f(z) + \rho^{\{x,z\}}_g(x) - \rho^{\{x,y,z\}}_g(x) + \rho^{\{x,z\}}_h(z) - \rho^{\{x,y,z\}}_h(z) + \rho^{\{x,y\}}_j(x) - \rho^{\{x,y,z\}}_j(x) \tag{8}$$

Again, since for any type u in the population, at least one of the probabilities $p^u_f(xzy)$, $p^u_g(yxz)$, $p^u_h(yzx)$, $p^u_j(zxy)$ must be zero, then $Q^{xyz}_{fghj} \leq 3$. We formalize our analysis in the next proposition. Naturally, we say that an extended random choice has a BAUP approximation with a binary state space if each model of the sequence that approximates the random choice according to Definition 1 has a binary state space.

Proposition 2. *Suppose an extended random choice ρ has a BAUP approximation (representation) with a binary state space. Then for any frames $f,g,h,j \in F$ and any items $x,y,z \in X$, $W^{xyz}_{fgh} \leq 2$ and $Q^{xyz}_{fghj} \leq 3$, where W and Q are linear polynomials of choice frequencies defined in eq. (7) and (8).*

Note that constraints given in Proposition 2 emerge solely because the set of potential linear orders induced by an expected utility is restricted when the state space is binary; this argument does not involve consideration of the beliefs' updating. Therefore, this result is robust to the assumptions regarding the actual updating procedure—for instance, if the agent uses Bayes formula or not—as well as assumptions regarding common prior and common framing function. Whenever the agents in the population are expected utility maximizers with binary state space, the observed choice frequencies should satisfy conditions given in Proposition 2.

A continuation of analysis of linear orders over three items performed above does not bring new constraints that are not implied by $W^{xyz}_{fgh} \leq 2$ and $Q^{xyz}_{fghj} \leq 3$. However, consideration of linear orders over more than 3 items, most likely, brings new constraints necessary for a BAUP representation with binary state space. The corresponding general analysis is performed in the next section for an arbitrary size of the state space.

Finally, note that constraint $W^{xyz}_{fgh} \leq 2$ has a bite only when all 3 frames are distinct; indeed, if, say, $f = g$, then $p_f(xyz) + p_f(yzx) \leq 1$, hence $W^{xyz}_{ffh} = p_f(xyz) + p_f(yzx) + p_h(zxy) \leq 2$ always holds. Similarly, a constraint $Q^{xyz}_{fghj} \leq 3$ has a bite only when all 4 frames f, g, h, j are distinct.

3.2 General State Space

Assume now that the relevant state space is known—meaning that the analyst knows $|S|$—but it may have cardinality larger than 2. Building on the intuition developed in the previous section, let O be the set of linear orders over items in X. For each state-dependent utility function u, define

$$O(u) \;=\; \Big\{ \succeq \in O \;\Big|\; \exists \lambda \in \triangle(S):\; x \succ y \iff U^{\lambda}(x) > U^{\lambda}(y)\ \forall x, y \in X \Big\} \tag{9}$$

where

$$U^{\lambda}(x) \;=\; \sum_{s \in S} \lambda_s u_s(x) \qquad \forall x \in X$$

is the corresponding expected utility function for belief $\lambda \in \triangle(S)$. An important observation is that if the set of alternative X is large enough, the set $O(u)$ cannot contain all linear orders over alternatives in X.

Lemma 1. *For any size of the state space $|S|$, there is n such that if $|X| \geq n$, then $O(u) \neq O$ for any state-dependent utility function u.*

The intuition behind Lemma 1 generalizes the observation made by us for $|S| = 2$ that the beliefs λ_{xy}, λ_{xz}, λ_{yz} divide the probability simplex by at most 4 regions where the induced linear orders remain constant, and uses a Combinatorics result[12] regarding the slicing of the space $\mathbb{R}^{|S|-1}$ by a set of hyperplanes of dimension $|S|-2$.

We consider now a large enough set of alternatives. Define the collection of sets of linear orders Θ by

$$\Theta \ = \ \{ \ O(u) \mid u \text{ is some state-dependent utility function } \} \tag{10}$$

and for $k = 1, 2, ..., |X|!$, define

$$\Psi^k \ = \ \left\{ \psi \subseteq O \ \middle| \ |\psi \backslash \theta| \geq k \ \ \forall \theta \in \Theta \ \right\} \tag{11}$$

where $|\psi \backslash \theta|$ is the number of elements of the set ψ that are not elements of the set θ. Thus, Ψ^k consists of the sets of linear orders ψ such that each type $\theta \in \Theta$ of the agent misses at least k linear orders in ψ.

Suppose an extended random choice ρ has a BAUP representation with state space S. As we know (Proposition 1), in this case, for each frame $f \in F$, there is a random utility representation given by $p_f \in \triangle(O)$.

DEFINITION 2. Let O be the set of all linear orders on X and ρ be an extended random choice. A family of distributions over linear orders (random utility representations) $\{p_f\}_{f \in F}$, $p_f \in \triangle(O)$ is associated with ρ if for each frame $f \in F$,

$$p_f(\{\succeq \in O \mid x \succ y \ \forall y \in A \backslash x\}) = \rho_f^A(x).$$

Theorem 2. *Assume an extended random choice ρ has a BAUP approximation (representation) with state space S, and let Ψ^k for $k = 1, 2, ...$ be defined via eq. (9), (10), (11). Then there is a family of random utility representations $\{p_f\}_{f \in F}$ associated with ρ such that for any k, for any $\psi = (\succeq_1, ..., \succeq_n) \in \Psi^k$, for any $f_1, ..., f_n \in F$*

$$\sum_{i=1}^{n} p_{f_i}(\succeq_i) \ \leq \ n - k$$

Theorem 2 says that fixation of the state space of a BAUP model places a set of implicit restrictions on the observed choice frequencies across different frames that can be induced by such

[12] The result is given in the Introduction (page 1) of Orlik and Terao (2013), and, citing them, this result "was obtained by L. Schliifli on page 39 of his great posthumous work, Theorie der vielfachen Kontinuitiit (Denkschriften der Schweizerischen naturforschenden Gesellschaft, vol. 38, 1901)".

a model. Although the number of various restrictions placed on quantities $\{p_{f\in F}\}_{f\in F}$ might be large, the statement of the Theorem is relatively weak. Indeed, the result is formulated in terms of the random utility representations rather than in terms of the observed choice frequencies, and, most importantly, it does not deliver conditions necessary for all such representations, but rather says that there *exists* a family of representation that satisfies all considered conditions. The next Corollary, however, says that for rich enough choice data, these restrictions do have a bite.

Corollary 3. *For any finite state space S, there are numbers n, m such that if $|X| \geq n$ and $|F| \geq m$, there is an extended random choice ρ that admits a random utility representation within each frame, but does not have a BAUP approximation with state space S.*

Let us now show how to find some explicit constraints on the choice frequencies induced by a BAUP model with a fixed state space. The result given in Theorem 2 is relatively weak because the same random choice ρ_f within frame f may have multiple random utility representations p_f. In a special case when $|X| = 3$, the representation is always unique—the fact that we have used to derive Proposition 2. Note that the Block-Marschak polynomial $q_f(x,A)$ gives the probability that A is an upper counter set of x according to *any* random utility representation of ρ_f. Let Γ be the set of all pairs (x,A), $x \in A$, $A \subseteq X$ such that $x \notin A$:

$$\Gamma = \left\{(x,A) \in X \times 2^X \;\middle|\; x \notin A\right\} \tag{12}$$

Thus, a pair $(x,A) \in \Gamma$ indexes a Block-Marschak polynomial $q_f(A,x)$. We are not interested in $x \in A$, since in this case, $q_f(x,A) = 0$ by definition. For $k = 1,2,\dots$, define

$$\Xi^k = \left\{\xi \subseteq \Gamma \;\middle|\; \forall \theta \in \Theta \;\; \exists (x_1,A_1),\dots,(x_k,A_k) \in \xi : \; \theta \cap \left\{ \geq \in O \;\middle|\; y \succ x_i \Leftrightarrow y \in A_i \right\} = \varnothing \; \forall i = 1,\dots,k \right\} \tag{13}$$

In words, an element $\xi \in \Xi^k$ is a collection of pairs (x,A) such that for any type $\theta = O(u)$ of the agent with some state-dependent utility function u, there are at least k distinct pairs $(x,A) \in \xi$ such that for each of these pairs, there is no linear order $\geq \in \theta$[13] for which the set of items A is a strict upper counter set of the item x.

Theorem 3. *Assume extended random choice ρ has a BAUP approximation (representation) with state space S, and let Ξ^k for $k = 1,2,\dots$ be defined via eq. (9), (10), (12), (13). Then for any k, for any*

[13] That is, for $\theta = O(u)$, an order that could be induced by an expectation of utility u_s with respect to some belief regarding the state $s \in S$.

$\xi = ((x_1, A_1), ..., (x_n, A_n)) \in \Xi^k$, for any $f_1, ..., f_n \in F$

$$\sum_{i=1}^{n} q_{f_i}(x_i, A_i) \;\le\; n - k$$

where for each $f \in F$, q_f is a Block-Marschak polynomial constructed from the random choice frequencies ρ_f according to eq. (6).

Theorem 3 provides conditions necessary for a representation (or approximation) of a given extended random choice with a BAUP model with a fixed state space S. These constraints are given in terms of the familiar Block-Marschak polynomials that are used to characterize a classical random utility model. The advantage of this formulation is that the Block-Marschak polynomials are well-defined linear functions of the observed choice data. Theorem 3, thus, provides constraints on the possible joint values of the random choice frequencies—expressed via a Block-Marschak polynomial within each frame—across various frames in the form of a simple additive statistic.

As an illustration, let us show a way to derive Proposition 2 from Theorem 3. For that, let $X = \{x, y, z\}$, and consider

$$\xi \;=\; \{\; (y, \{x\}),\; (z, \{y\}),\; (x, \{z\}) \;\}$$

Note that there is only one linear order such that $\{x\}$ is a strict upper counter set of item y, namely, $x \succ y \succ z$; and similarly, for pairs $(z, \{y\})$ and $(x, \{z\})$ the corresponding unique linear orders are $y \succ z \succ x$ and $z \succ x \succ y$. Thus, to prove that $\xi \in \Xi^1$, we only need to argue that each $\theta \in \Theta$ has empty intersection with one of the singleton sets of linear orders $\{xyz\}, \{yzx\}, \{zxy\} \subseteq O$; but we have already checked it in Section 3.1. Applying Theorem 3 for $\xi \in \Xi^1$, we get

$$q_f(y, \{x\}) + q_g(z, \{y\}) + q_h(x, \{z\}) \;\le\; 3 - 1$$

Supstituting the expressions for the corresponding Block-Marschak polynomials, we get $W^{xyz}_{fgh} \le 2$. We can get the inequality $Q^{xyz}_{fghj} \le 3$ in a similar fashion by considering $\xi' = \{(x, \{z\}), (x, \{y\}), (y, \{z\}), (z, \{x\})\} \in \Xi^1$.

It is worth to mention that, analogous to Proposition 2, both Theorem 2 and Theorem 3 also hold when the agent's updating procedure is not Bayesian. In fact, both theorems hold whenever the agents in the population are expected utility maximizers with state space S. Finally, whether Theorem 2 describes a different set of constraints on the extended random choice than Theorem 3 is an open question.

4 Appendix

4.1 Proof of Proposition 1

Let BAUP model (S,π,σ,μ) represents an extended random choice ρ. For each frame f, consider the distribution r over utility functions $v: X \to \mathbb{R}$ given by the following formula:

$$r_f(v) \;=\; \sum_u \mu(u)\cdot Pr\Big(u^{M(f)} = v\Big)$$

Then

$$\begin{aligned}
\rho_f^A(x) &= \sum_u \mu(u)\cdot Pr\Big(\Big\{u^{M(f)}(x) \geq u^{M(f)}(y)\ \forall y \in A\Big\}\Big) \\
&= \sum_u \mu(u)\cdot Pr\Big(\Big\{u^{M(f)} = v \;\Big|\; v(x) \geq v(y)\ \forall y \in A\Big\}\Big) \;=\; r_f\Big(\Big\{v \;\Big|\; v(x) \geq v(y)\ \forall y \in A\Big\}\Big)
\end{aligned}$$

Thus, for each frame $f \in F$, a random utility model r_f represents the random choice ρ_f within this frame. ■

4.2 Proof of Corollary 1

When $|S| = 1$, the state space is irrelevant, and the expected utility is given by function u_1. Hence,

$$\rho_f^A(x) \;=\; \mu\Big(\Big\{u \;\Big|\; u_1(x) \geq u_1(y)\ \forall y \in A\Big\}\Big)$$

The Corollary 1 follows immediately ■

4.3 Proof of Theorem 1

We start with the second statement. Without loss of generality, assume that any combination of frame details is a possible frame; that is, $F = 2^D$.

Let O be the set of all linear orders (complete, transitive and antisymmetric binary relations) on the set of alternatives X, and let $\mathcal{O} = O^F$ be the set of vectors of linear orders indexed by the frames $f \in F$. For $o \in \mathcal{O}$, we write $x o_f y$ when $x \in X$ is ordered higher than $y \in X$ according to order o_f under frame $f \in F$. Let $p \in \triangle\mathcal{O}$ be a probability distribution over $\mathcal{O}$. We say that a BAUP model (S,π,σ,μ) represents (or induces) probability distribution p over vectors of linear orders $o \in \mathcal{O}$ if for all $f \in F$, and $x,y \in A$,

$$p(o) \;=\; \sum_u \mu(u)\cdot Pr\Big(\Big\{u^{M(f)}(x) \geq u^{M(f)}(y) \Longleftrightarrow x o_f y \quad \forall f \in F\Big\}\Big) \tag{14}$$

where the summation is performed over all state-dependent utility functions $u : S \times X \to \mathbb{R}$ such that $\mu(u) > 0$, and $u^{M(f)}(x)$ is a function given in eq. (2) of a random signal $M(f)$ associated with frame f and distributed according to eq. (5).

Lemma 2. *There is a state space S and a framing function $\sigma : D \to \Sigma(S)$ such that for any $\epsilon > 0$, there is a prior probability $\pi \in \triangle(S)$ such that for any $o \in \mathcal{O}$, there is a state-dependent utility function $u : S \times X \to \mathbb{R}$ such that BAUP model $(S, \pi, \sigma, \delta_u)$ induces a distribution over vectors of orders $p \in \triangle(\mathcal{O})$ such that $p(o) \geq 1 - \epsilon$, where δ_u denotes a degenerate distribution that places probability 1 on u.*

Proof of Lemma 2. Consider a state space $S = \{0,1\}^D$; thus, $s = (s_1, ..., s_{|D|}) \in S$, where $s_d \in \{0,1\}$ for $d \in D$. Consider binary message spaces $\mathcal{M}_d = \{d\} \times \{0,1\}$ for frame details $d \in D$. Denote by $\mathcal{M}$ the space of joint messages (all combinations of messages/absence of messages for different frame details) and define the signal structures $\sigma(d)$ as follows:

$$\sigma(d)(s,(d,i)) \;=\; \mathbb{1}\{s_d = i\} \qquad \text{for } i \in \{0,1\}$$

In other words, signal (d,i) reveals the d-th component of the state: $s_d = i$.

Denote by ι the state $\iota \in S$ such that $\iota_d = 1$ for all $d \in D$. Define $\pi \in \triangle(S)$ as follows:

$$\pi(\iota) \;=\; 1 - \epsilon, \qquad \pi(s) \;=\; \frac{\epsilon}{2^{|D|} - 1} \quad \text{if } s \neq \iota$$

For all $f \in F$, let $V_f : X \to \mathbb{R}$ be the following utility function:

$$V_f(x) \;=\; |\{y \in X | x o_f y\}| \tag{15}$$

Next, for $s \in S$, denote by

$$L(s) = \{d \in D | s_d = 1\}$$

and define the following state-dependent utility function:

$$u_s(x) \;=\; (\pi(s))^{-1} \cdot \sum_{J:\, L(s) \subseteq J \subseteq D} (-1)^{|J| - |L(s)|} \cdot 2^{|D| - |J|} \cdot V_J(x) + v(x)$$

where function $v : X \to (0,1)$ will be specified later.

Consider an arbitrary frame $f \in F$. Note that state s is a vertex of a binary D-dimensional cube. For convenience, denote by s_f the projection of s on $f \subseteq D$. Consider signal realization

$\iota_f \equiv \{(d,1)\}_{d\in f} \in \mathcal{M}$. The corresponding expected utility of item x is

$$
\begin{aligned}
u^{\iota_f}(x) &= \frac{\sum_{s:\ s_f=\iota_f} \pi(s) u_s(x)}{\sum_{s:\ s_f=\iota_f} \pi(s)} = \\
&= \frac{1}{\pi(\{s\in S | s_f = \iota_f\})} \cdot \sum_{s:\ s_f=\iota_f} \ \sum_{J:\ L(s)\subseteq J\subseteq D} (-1)^{|J|-|L(s)|} \cdot 2^{|D|-|J|} \cdot V_J(x) + v(x) = \\
&= \frac{1}{1-\frac{2^{|D|}-2^{|f|}}{2^{|D|}-1}\cdot\epsilon} \cdot \sum_{L,J:\ f\subseteq L\subseteq J\subseteq D} (-1)^{|J|-|L|} \cdot 2^{|D|-|J|} \cdot V_J(x) + v(x) = \\
&= \frac{1}{1-\frac{2^{|D|}-2^{|f|}}{2^{|D|}-1}\cdot\epsilon} \cdot \left[2^{|D|-|f|} \cdot V_f(x) + \sum_{L,J:\ f\subseteq L\subseteq J\subseteq D,\ J\neq f} (-1)^{|J|-|L|} \cdot 2^{|D|-|J|} \cdot V_J(x) \right] + v(x) = \\
&= \frac{1}{1-\frac{2^{|D|}-2^{|f|}}{2^{|D|}-1}\cdot\epsilon} \cdot \left[2^{|D|-|f|} \cdot V_f(x) + \sum_{J:\ f\subseteq J\subseteq D,\ J\neq f} (-1)^{|J|} \cdot 2^{|D|-|J|} \cdot V_J(x) \cdot \sum_{L:\ f\subseteq L\subseteq J} (-1)^{-|L|} \right] + v(x) = \\
&= \frac{1}{1-\frac{2^{|D|}-2^{|f|}}{2^{|D|}-1}\cdot\epsilon} \cdot V_f(x) + v(x)
\end{aligned}
$$

where we used the Binomial theorem to get

$$
\sum_{L:\ f\subseteq L\subseteq J} (-1)^{-|L|} = \sum_{k=|f|}^{|J|} \ \sum_{L:\ f\subseteq L\subseteq J,\ |L|=k} (-1)^{-|L|} = (-1)^{-|f|} \sum_{k=0}^{|J|-|f|} \binom{|J|-|f|}{k} (-1)^{-k} = (-1)^{-|f|} \cdot (1-1)^{|J|-|f|} = 0
$$

for $J \neq f$.

Claim 1. *There is a function $v : X \to (0,1)$ such that for any frame $f \in F$, function $u^{\iota_f}(\cdot)$ represents linear order o_f, and for any signal realization $m \in \{0,1\}^f$, the expected utility $u^m(\cdot)$ represents some linear order $\tilde{o} \in O$.*

Proof. Note that $[x o_f y$ and $x \neq y]$ implies

$$
\frac{1}{1-\frac{2^{|D|}-2^{|f|}}{2^{|D|}-1}\cdot\epsilon} \cdot V_f(x) - \frac{1}{1-\frac{2^{|D|}-2^{|f|}}{2^{|D|}-1}\cdot\epsilon} \cdot V_f(y) > \frac{1}{1-\frac{2^{|D|}-2^{|f|}}{2^{|D|}-1}\cdot\epsilon} \geq 1
$$

Hence, $u^{\iota_f}(x) \geq u^{\iota_f}(y)$ if and only if $x o_f y$, proving that the first statement of the Claim holds for any function $v : X \to (0,1)$.

Next, note that for any signal realization $m \in \mathcal{M}$, for any $x, y \in X$,

$$
u^m(x) - u^m(y) = C_m(x) - C_m(y) + v(x) - v(y)
$$

where $\{C_m(x)\}_{m\in\mathcal{M}, x\in X}$ is a finite set of quantities that does not depend on $v(\cdot)$. Let

$$
\widetilde{C} = \min_{m\in M; x,y\in X} \left\{ |C_m(x) - C_m(y)| \text{ such that } |C_m(x) - C_m(y)| > 0 \right\}
$$

hence, $\widetilde{C} > 0$. Enumerate alternatives $x \in X$ arbitrary by $k = 1,2,3,...,|X|$ and let $k(x)$ to be the number of alternative x. Define

$$v(x) \; = \; \min\{\widetilde{C},1\} \cdot e^{-k(x)}$$

It is straightforward to see that $u^m(x) \neq u^m(y)$ for all $m \in \mathcal{M}$, $x,y \in X$, hence the second statement of the Claim holds as well. □

By Claim 1, BAUP model (S,π,σ,δ_u) induces some distribution of linear orders p. Next, note that for state $\iota = (1,...,1)$,

$$\sigma(d)(\iota,(d,1)) \; = \; 1$$

Hence, using Claim 1, we get

$$p(o) \; \geq \; Pr\Big(\Big\{M_f = \iota_f \; \forall f \in F\Big\}\Big) \; = \; \pi(\iota) \; = \; 1-\epsilon$$

proving the Lemma. ■

Lemma 3. *Let $Y = \{1,...,n,n+1\}$, and $y^j = (y^j_1,...,y^j_{n+1})$ for $j = 1,2,...$ be a sequence such that $y^j_i \in \triangle(Y)$ for all i,j, and $y^j_i \longrightarrow \delta_i$ for all $i = 1,...,n,n+1$. Then for any $p \in \triangle(Y)$ such that $p(i) > 0$ for all $i = 1,...,n,n+1$, there is j such that $p \in conv(\{y^j_1,...,y^j_{n+1}\})$.*

Proof of Lemma 3. Since $\sum_{i=1}^{n+1} y'_i = 1$ for all $y' \in \triangle(Y)$, it is enough to consider $Y_0 = \{x \in \mathbb{R}^n | x_i \geq 0 \; \forall i \in \{1,...,n\}, \; \sum_{i=1}^n x_i \leq 1\}$ and a sequence $z^j = (z^j_1,...,z^j_n,z^j_{n+1})$, $j = 1,2,..$ such that $z^j \in Y_0$, $z^j_i \longrightarrow e_i$ for $i \in \{1,...,n\}$, where $e_{ii} = 1$, e_{ik}=0, for $k \neq i$, and $z^j_{n+1} \longrightarrow e_0$, where $e_0 = (0,...,0)$. Consider now $q \in Y_0$ with $q_i = p_i$ for $i \in \{1,...,n\}$. Note that half-spaces $x_i \geq 0$, $\sum_i^n x_i \leq 1$ include $n = 1$-dimensional faces $F_i = \text{conv}\Big(\{e_0,e_1,...,e_n\} \backslash \{e_i\}\Big)$, $i = 0,1,...,n$ of the polytope S_0, and their intersection is equal to Y_0. Clearly, for large enough j, $F^j_i = \text{conv}\Big(\{y^j_0,y^j_1,...,y^j_n\} \backslash \{y^j_i\}\Big)$ for $i = 0,1,...,n$ comprise the set of $n-1$ dimensional faces of $Z^j = \text{conv}\{y^j_0,y^j_1,...,y^j_n\}$. Since $z^j_i \longrightarrow e_i$, it follows that Z^j is equal to the intersection of half-spaces H^j_i, $i \in \{0,1,...,n\}$, where $H^j_i = \{x \in \mathbb{R}^n | \sum_{k=1}^n \epsilon^j_{ik} x_k + x_i \geq \epsilon^j_{i0}\}$ for $i \in \{1,...,n\}$, $H^j_0 = \{x \in \mathbb{R}^n | \sum_{k=1}^n (1+\epsilon^j_{0k}) x_k \leq 1 + \epsilon^j_{00}\}$ for some $\epsilon^j_{ik} \longrightarrow 0$. Since $q_i > 0$, and $\sum_{i=1}^n q_i < 1$, for large enough j, $\sum_{k=1}^n \epsilon^j_{ik} q_k + q_i \geq \epsilon^j_{i0}$, and $\sum_{k=1}^n (1+\epsilon^j_{0k}) q_k \leq 1 + \epsilon^j_{00}$. Thus, $q \in H^j_i$ for all $i \in \{0,1,...,n\}$, hence $q \in Z^j$, and $p \in \text{conv}(\{y^j_1,...,y^j_{n+1}\})$. ■

Lemma 4. *There is a state space S and a framing function $\sigma : D \to \Sigma(S)$ such that if the distribution $p \in \triangle\mathcal{O}$ of vectors of linear orders satisfies $p(o) > 0$ for all $o \in \mathcal{O}$, then there is a prior $\pi \in \triangle(S)$ and a distribution over state-dependent utility functions μ such that the BAUP model (S, π, σ, μ) induces p.*

Proof of Lemma 4. Let S be a state space, σ be a framing function, and, for given $\epsilon > 0$, let $\pi \in \triangle(S)$ be a prior probability and for each $o \in \mathcal{O}$, let $u^\epsilon(o): S \times X \to \mathbb{R}$ be a state-dependent utility function such that $(S, \pi, \sigma, \delta_{u_o^\epsilon})$ induces a distribution of vectors of linear orders p_o^ϵ with the property $p_o^\epsilon(o) \geq 1 - \epsilon$. By Lemma 2, we can find such objects. Next, a space $\triangle\mathcal{O}$ of a probability distributions over vectors of linear orders is a $|X!|^{|F|}$-dimensional probability simplex, and $p_o^\epsilon \longrightarrow \delta_o$ for $\epsilon \longrightarrow 0$. Therefore, by Lemma 3, for small enough ϵ, $p \in \text{conv}\Big(\{p_o^\epsilon\}_{o \in \mathcal{O}}\Big)$, hence $p = \sum_{o \in \mathcal{O}} \alpha_o p_o^\epsilon$ for some convex weights α. Consider a distribution over state-dependent utility functions μ such that $\mu(u_o^\epsilon) = \alpha_o$, then a BAUP model (S, π, σ, μ) represents p. ∎

Now, let ρ be an extended random choice such that $q_f(A, x) > 0$ for all $f \in F$, $A \in \mathcal{X}$, $x \in A$, where quantities q are defined in eq. (6). Consider $\widetilde{\rho}: F \times \mathcal{X} \times X \to \mathbb{R}$ defined by

$$\widetilde{\rho}_f^A(x) \;=\; \frac{1}{1-\epsilon}\rho_f^A(x) - \frac{\epsilon}{1-\epsilon} \cdot \frac{\mathbb{1}\{x \in A\}}{|A|}$$

where $\epsilon > 0$ is a small parameter. Note that $\widetilde{\rho}_f^A(x) = 0$ for $x \notin A$, and $\sum_{x \in A} \widetilde{\rho}_f^A(x) = 1$. Moreover, since $\rho_f^A(x) = q_f(x, X \backslash A) + \sum_{B \subsetneq A} q_f(x, B) > 0$ for $x \in A$, it follows that for small enough ϵ, $\widetilde{\rho}_f^A(x) > 0$ for all $f \in F$, $A \in \mathcal{X}$, $x \in A$, hence $\widetilde{\rho}$ is an extended random choice. Let $\widetilde{q}$ be an associated with $\widetilde{\rho}$ set of linear polynomials according to eq. (6). Then for small enough ϵ, $\widetilde{q}_f^A(x) \geq 0$ for all $f \in F$, $A \in \mathcal{X}$, $x \in A$. For each frame f, by the results of Falmagne (1978), $\widetilde{\rho}_f$ admits a random utility representation; following Barberá and Pattanaik (1986), we consider the associated distribution over linear orders $\widetilde{\tau}_f \in \triangle(O)$. Let $\tau^{uniform}$ be a uniform distribution over linear orders given by $\tau^{uniform}(d) = 1/|O|$ for all $d \in O$. Notice that $\tau^{uniform}$ represents random choice given by $Pr(\text{choose } x \text{ from } A) = \dfrac{\mathbb{1}\{x \in A\}}{|A|}$. Since

$$\rho_f^A(x) \;=\; (1-\epsilon) \cdot \widetilde{\rho}_f^A(x) + \epsilon \cdot \frac{\mathbb{1}\{x \in A\}}{|A|}$$

it follows that ρ^f has a random utility (random order) representation given by

$$\tau_f \;=\; (1-\epsilon) \cdot \widetilde{\tau}_f + \epsilon \cdot \tau^{uniform}$$

Construct a distribution of linear orders $p \in \mathcal{O}$ given by

$$p(o) \;=\; \prod_{f \in F} \tau_f(o_f)$$

Note that, since $\tau_f(o_f) \geq \epsilon/|O|$, then for any $o \in \mathcal{O}$, $p(o) \geq \left(\frac{\epsilon}{|O|}\right)^{|F|} > 0$. By Lemma 4, p has some BAUP representation (S, π, σ, μ). Since marginal distributions of p are given by τ_f for each frame f, and they comprise random utility representations of random choice ρ_f for each f, then (S, π, σ, μ) is a BAUP representaion of the extended random choice ρ. Thus, we have proven the second statement of the Theorem.[14]

Suppose now that $q_f(x, A) \geq 0$ for all $f \in F$, $A \in \mathcal{X}$, $x \in A$, and let $\epsilon > 0$ is given. Consider an extended random choice τ given by

$$\tau_f^A(x) \;=\; (1 - \epsilon/2) \cdot \rho_f^A(x) + (\epsilon/2) \cdot \frac{\mathbb{1}\{x \in A\}}{|A|}$$

Then

$$|\rho_f^A(x) - \tau_f^A(x)| \;\leq\; (\epsilon/2) \cdot \left| \frac{\mathbb{1}\{x \in A\}}{|A|} - \rho_f^A(x) \right| \;<\; \epsilon$$

Next, let $q^{uniform}(x, A)$ for $x \in X$, $A \in \mathcal{X}$ be the Block-Marschak polynomials associated with the uniform random choice $Pr(\text{choose } x \text{ from } A) = \frac{\mathbb{1}\{x \in A\}}{|A|}$. According to Barberá and Pattanaik (1986), $q(x, A)$ is the probability that A is an upper counter set of x according to a random utility model that rationalizes the random choice. Since uniform distribution over linear orders rationalizes uniform random choice, it follows that $q^{uniform}(x, A) > 0$ for all $A \in \mathcal{X}$, $a \notin A$. Let $\tilde{q}_f(x, A)$ for $f \in F$, $A \in \mathcal{X}$, $x \in A$ be Block-Marschak polynomials constructed according to eq. 6 from the extended random choice τ, then

$$\tilde{q}_f(x, A) \;=\; (1 - \epsilon/2) \cdot q_f(x, A) + (\epsilon/2) \cdot q^{uniform}(x, A)$$

Hence, $\tilde{q}_f(x, A) > 0$ for all $f \in F$, $A \in \mathcal{X}$, $x \in A$. By the proven second statement of the Theorem, the extended random choice τ admits a BAUP representation. Since this holds for arbitrary $\epsilon > 0$, the if part of the first statement of the Theorem is proven.

Finally, assume that an extended random choice ρ has a BAUP approximation given by a sequence of BAUP models $\left\{\left(S^j, \pi^j, \sigma^j, \mu^j\right)\right\}_{j=1,2,\ldots}$. For any j, a BAUP model $\left(S^j, \pi^j, \sigma^j, \mu^j\right)$ induces a

[14] To see that our assumption $F = 2^D$ is without loss of generality, notice that we can always augment the set of frames to 2^D be creating fictitious random choice frequencies that replicate those in one of the existing frames. We then can apply the proof for the case $F = 2^D$

random utility model within each frame $f \in F$ with the induced random choice ρ^j_f. By the results of Falmagne (1978), it must be that $q^j_f(x,A) \geq 0$ for the associated Block-Marschak polynomials. Since $\rho^j_f \longrightarrow \rho_f$ for every $f \in F$, it follows that $q^j_f(x,A) \to q_f(x,A)$, implying $q_f(x,A) \geq 0$ for all $f \in F$, $A \in \mathcal{X}$, $x \in A$. The only if part of the first statement of the Theorem is proven. ■

4.4 Proof of Corollary 2

The Corollary follows immediately from Theorem 1 and the results of Falmagne (1978), Barberá and Pattanaik (1986) that the random choice ρ_f has a random utility representation if and only if $q_f(x,A) \geq 0$ for all $x \in X$, $A \in \mathcal{X}$. ■

4.5 Proof of Proposition 2

The major part of the proof is given in our analysis in Section 3.1. We complement it by calculating the set Θ. Note also that for our purposes, we do not need to calculate the set Θ precisely; it suffice to calculate a set $\Theta_0 \subseteq \Theta$ such that for each $\theta \in \Theta$, there is $\theta_0 \in \Theta_0$ such that $\theta \subseteq \theta_0$. It follows from the fact that $\psi \not\subseteq \theta_0$ for all $\theta_0 \in \Theta_0$ if and only if $\psi \not\subseteq \theta$ for all $\theta \in \Theta$.

Thus, it is without loss of generality to assume that both $u_1(\cdot)$ and $u_2(\cdot)$ do not exhibit ties. There is always a permutation of x, y, z such that $u_1(x) > u_1(y) > u_1(z)$. If $u_2(z) > u_2(y) > u_2(x)$, then either $O(u) = \theta_1 = \{xyz, xzy, zxy, zyx\}$, or $O(u) = \theta_2 = \{xyz, yxz, yzx, zyx\}$; see Figure 1. If $u_2(y) > u_2(z) > u_2(x)$, then $O(u) = \{xyz, yxz, yzx\} \subset \theta_2$. If $u_2(z) > u_2(x) > u_2(y)$, then $O(u) = \{xyz, xzy, zxy\} \subset \theta_1$. If $u_2(y) > u_2(x) > u_2(z)$, then $O(u) = \{xyz, yxz\} \subset \theta_2$. If $u_2(x) > u_2(z) > u_2(y)$, then $O(u) = \{xyz, xzy\} \subset \theta_1$. Finally, if $u_2(x) > u_2(y) > u_2(z)$, then $O(u) = \{xyz\} \subset \theta_1, \theta_2$. It follows that the set Θ_0 consists of the sets θ_1, θ_2 and the sets induced from θ_1 and θ_2 by the permutation of items x, y, z, supporting the analysis provided in Section 3.1.

Next, consider $\psi = \{xyz, yzx, zxy\}$. Note that $\psi \not\subseteq \theta_1, \theta_2$. Next, the orbit of ψ with respect to permutations of x, y, z consists of ψ and $\psi' = \{yxz, xzy, zyx\}$. We can check that $yxz \notin \theta_1$, and $xzy \notin \theta_2$, proving that ψ is not a subset of any $\theta \in \Theta_0$, and hence, ψ is not a subset of any $\theta \in \Theta$. The rest of the analysis for the necessity of $W^{xyz}_{fgh} \leq 2$ is provided in Section 3.1.

Consider now $\psi = \{xzy, yxz, yzx, zxy\}$. Since $|\psi| = 4 = |\theta_1| = |\theta_2|$, it suffice to see that ψ is not a result of applying of some permutation of x, y, z to θ_1, since the latter always features only 2 and not 3 different items on the top of the preference order, and, similarly, it is not a result of applying of some permutation of x, y, z to θ_2, since the latter always features only 2 and not 3 different items at the bottom of the preference order. The rest of the analysis for the necessity of $Q^{xyz}_{fghj} \leq 3$ is provided in Section 3.1. ■

4.6 Proof of Lemma 1

Let $|X| = n$, and assume $O(u) = O$ for a state-dependent utility function u. Consider a probability simplex $\triangle S$ as a subset of $R^{|S|-1}$ consisting of points $p \in R^{|S|-1}$ such that $p_i \geq 0$, $\sum_{i=1}^{|S|-1} p_i \leq 1$. Define (expected) utility of item $x \in X$ depending on $p \in R^{|S|-1}$ to be

$$V(p,x) \;=\; \sum_{i=1}^{|S|-1} p_i \cdot u_i(x) + \left(1 - \sum_{i=1}^{|S|-1} p_i\right) u_{|S|}(x)$$

Since all linear orders in O are induced by an expected utility for some $\pi \in \triangle S$, it follows that for each pair of items $x, y \in X$, there is an open half-space of $\mathbb{R}^{|S|-1}$ such that $V(p,x) > V(p,y)$ and an open half-space of $\mathbb{R}^{|S|-1}$ such that $V(p,y) > V(p,x)$, divided by an $|S|-2$ dimensional hyperplane consisting of points p such that $V(p,x) = V(p,y)$. Thus, $\frac{n(n-1)}{2}$ hyperplanes given by equations $V(p,x) = V(p,y)$ for various pairs $x, y \in X$ divide $R^{|S|-1}$ on multiple cells. Within each cell, the expected utility $V(p,x)$ induces the same linear order over X. The total number of linear orders induced by $V(p,\cdot)$ for various p may not exceed the maximum number of potential cells Y. By Orlik and Terao (2013) (page 1),

$$Y \;=\; \sum_{k=0}^{|S|-1} \begin{pmatrix} n(n-1)/2 \\ k \end{pmatrix}$$

where $\begin{pmatrix} m \\ l \end{pmatrix} = \dfrac{m!}{l!(m-l)!}$ is a Binomial coefficient. Since $|O| = |X|! = n!$, we have

$$\sum_{k=0}^{|S|-1} \begin{pmatrix} n(n-1)/2 \\ k \end{pmatrix} \;\geq\; n!$$

For $n > 2|S|$, we have

$$n! \;\leq\; \sum_{k=0}^{|S|-1} \begin{pmatrix} n(n-1)/2 \\ k \end{pmatrix} \;\leq\; |S| \cdot \begin{pmatrix} n(n-1)/2 \\ |S| \end{pmatrix} \;\leq\; |S| \cdot \left(n^2/2\right)^{|S|}$$

Since any fixed k, $n! > \text{const} \cdot n^k$ for large enough n, the above inequality does not hold for large enough n. Hence, for large enough $n = |X|$, $O(u) \neq O$, proving the Lemma. ■

4.7 Proof of Theorem 2

Suppose an extended random choice ρ has a BAUP representation (S,π,σ,μ). For each $u \in \text{supp}(\mu)$, consider an induced family of random utility representations $\{p_f^u\}_{f\in F}$ given by

$$p_f^u(\succeq) \;=\; Pr\Big(\Big\{x \succ y \implies U^{M(f)}(x) > U^{M(f)}(y)\ \forall x,y \in X\Big\}\Big)$$

and a family of random utility representations $\{p_f\}_{f\in F}$ given by

$$p_f(\succeq) \;=\; \sum_u \mu(u) p_f^u(\succeq_i)$$

Consider arbitrary $\psi = \{\succeq_1,\ldots,\succeq_n\} \in \Psi^k$, then for any state-dependent utility function u, there are linear orders $\succeq_{i_1},\ldots,\succeq_{i_k} \in \psi\backslash O(u)$. It follows from the definition of $O(u)$ that

$$p_f^u(\succeq_{i_j}) \;=\; 0 \qquad \forall j = 1,\ldots,k$$

Hence,

$$\sum_{i=1}^n p_f^u(\succeq_i) \;=\; \sum_{i=1}^n \mathbb{1}\{i \neq i_1,\ldots,i_k\}\cdot p_f^u(\succeq_i) \;\leq\; \sum_{i=1}^n \mathbb{1}\{i \neq i_1,\ldots,i_k\} \;=\; n-k$$

It follows that

$$\sum_{i=1}^n p_f(\succeq_i) \;=\; \sum_{i=1}^n \sum_u \mu(u) p_f^u(\succeq_i) \;=\; \sum_u \mu(u) \sum_{i=1}^n p_f^u(\succeq_i) \;\leq\; n-k$$

Finally,

$$p_f\Big(\Big\{\succeq \in O | x \succ y\ \forall y \in A\backslash x\Big\}\Big) \;=\; \sum_u \mu(u) Pr\Big(\Big\{U^{M(f)}(x) > U^{M(f)}(y)\ \forall y \in A\backslash x\Big\}\Big) \;=\; \rho_f^A(x)$$

Thus, a family of random utility models $\{p_f\}_{f\in F}$ satisfies all desired properties. Since non-strict inequalities are preserved in the limit, the statement also holds for a BAUP approximation with state space $|S|$. ■

4.8 Proof of Corollary 3

Let $|X| = n$, $|F| = m = n!$, and let $O = \{\succeq_1,\ldots,\succeq_m\}$ be the set of linear orders on X; enumerate frames by $f_1,\ldots,f_m \in F$. Consider an extended random choice ρ induced by the family of distributions of linear $\{p_f\}_{f\in F}$ such that $p_{f_i} = \delta_{\succeq_i}$. That is,

$$\rho_{f_i}^A(x) \;=\; \mathbb{1}\{x \succ_i y\ \forall y \in A\backslash x\}$$

Towards a contradiction, assume that ρ has a BAUP approximation $(S, \pi^j, \sigma^j, \mu^j)$, $j = 1, 2, \dots$ with state space S. That is, for any $\epsilon > 0$, if j is large enough, a BAUP model $(S, \pi^j, \sigma^j, \mu^j)$ induces an extended random choice τ^j such that $\left|\rho_f^A(x) - \left(\tau^j\right)_f^A(x)\right| < \epsilon$ for all x, A, f.

Claim 2. *The family $\{p_f\}_{f \in F}$ is the unique family of random utility representations associated with the extended random choice ρ that it induces.*

Proof. Assume that some other family $\{t_f\}_{f \in F}$ is associated with the same extended random choice. Then there is a frame $f_i \in F$ and a pair of items x, y such that $x >_i y$ and $t_f(\{>\in O | x > y\}) < 1$, but this contradicts τ being associated with ρ, since $\rho_f^{\{x,y\}}(x) = 1$. □

By Lemma 1, if n is large enough, $O \notin \Theta$. By Theorem 2, for any BAUP model $(S, \pi^j, \sigma^j, \mu^j)$, there is a family of random utility representations p^j (in fact, induced by the corresponding BAUP model) with the property that

$$\sum_{i=1}^{|O|} p_{f_i}^j(\geq_i) \;\leq\; |O| - 1$$

Consider $\epsilon = (2|O|)^{-2}$. Then the above inequality implies that there is $i \in \{1, \dots, |O|\}$ and a linear order $> \neq \geq_i$ such that $p_{f_i}^j(\geq) > \epsilon$. Since $\geq \neq \geq_i$, then there are items $x, y \in X$ such that $x >_i y$ and $y > x$. Hence, $\rho_{f_i}^{\{x,y\}}(x) = 1$, $(\tau^j)_{f_i}^{\{x,y\}}(x) < 1 - \epsilon$. Since such frame and pair of items could be found for any j, this contradicts our assumption that $(S, \pi^j, \sigma^j, \mu^j)$ approximates ρ. ■

4.9 Proof of Theorem 3

The proof of Theorem 3 partially repeats the proof of Theorem 2.

Assume that an extended random choice ρ has a BAUP representation (S, π, σ, μ) with state space S. For each $u \in \text{supp}(\mu)$, consider an induced family of random utility representations $\{p_f^u\}_{f \in F}$ given by

$$p_f^u(\geq) \;=\; Pr\left(\left\{x > y \implies U^{M(f)}(x) > U^{M(f)}(y) \;\forall x, y \in X\right\}\right)$$

and a family of random utility representations $\{p_f\}_{f \in F}$ given by

$$p_f(\geq) \;=\; \sum_u \mu(u) p_f^u(\geq_i)$$

Note that p_f represents ρ_f for each frame $f \in F$. Therefore,[15] for $x \notin A$,

$$q_f(x,A) \;=\; p_f\Big(\Big\{ \succeq \in O \;\Big|\; y \succ x \;\Leftrightarrow y \in A \Big\}\Big)$$

Then

$$\begin{aligned}\sum_{i=1}^{n} q_{f_i}(x_i, A_i) \;&=\; \sum_{i=1}^{n}\sum_{u} \mu(u) p^u_{f_i}\Big(\Big\{ \succeq \in O \;\Big|\; y \succ x_i \;\Leftrightarrow y \in A_i \Big\}\Big) \;= \\ &\sum_{u} \mu(u) \sum_{i=1}^{n} p^u_{f_i}\Big(\Big\{ \succeq \in O \;\Big|\; y \succ x_i \;\Leftrightarrow y \in A_i \Big\}\Big) \;\leq\; n-k\end{aligned}$$

where we used the definition of Ξ to find that for any state-dependent utility function u, there are at least k pairs (x,A) such that for any f, $p^u_f\Big(\Big\{ \succeq \in O \;\Big|\; y \succ x \;\Leftrightarrow y \in A \Big\}\Big) = 0$ and, therefore,

$$\sum_{i=1}^{n} p^u_{f_i}\Big(\Big\{ \succeq \in O \;\Big|\; y \succ x_i \;\Leftrightarrow y \in A_i \Big\}\Big) \;\leq\; n-k$$

for any state-dependent utility function u. Thus, we have proven the necessity of constraints for a BAUP representation with a fixed state space S. Since the constraints are linear polynomials in choice frequencies, they continue to hold for a BAUP approximation with state space S as well. ■

References

Ahn, D. S. and T. Sarver (2013): "Preference for flexibility and random choice," *Econometrica*, 81, 341–361.

Barberá, S. and P. K. Pattanaik (1986): "Falmagne and the rationalizability of stochastic choices in terms of random orderings," *Econometrica: Journal of the Econometric Society*, 707–715.

Block, H. D. and J. Marschak (1960): "Random orderings and stochastic theories of responses," *Contributions to probability and statistics*, 97–132.

Deb, R. and L. Renou (2021): "Dynamic Choice and Common Learning," *CEPR Discussion Paper No. DP16160*.

Dekel, E., B. L. Lipman, and A. Rustichini (2001): "Representing preferences with a unique subjective state space," *Econometrica*, 69, 891–934.

[15] See the argument in Barberá and Pattanaik (1986).

Dillenberger, D., J. S. Lleras, P. Sadowski, and N. Takeoka (2014): "A theory of subjective learning," *Journal of Economic Theory*, 153, 287–312.

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

Frick, M., R. Iijima, and T. Strzalecki (2019): "Dynamic random utility," *Econometrica*, 87, 1941–2002.

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

Huber, J., J. W. Payne, and C. Puto (1982): "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis," *Journal of consumer research*, 9, 90–98.

Karni, E. and Z. Safra (2016): "A theory of stochastic choice under uncertainty," *Journal of Mathematical Economics*, 63, 164–173.

Kreps, D. M. (1979): "A representation theorem for" preference for flexibility"," *Econometrica: Journal of the Econometric Society*, 565–577.

Lu, J. (2016): "Random choice and private information," *Econometrica*, 84, 1983–2027.

Manzini, P. and M. Mariotti (2014): "Stochastic choice and consideration sets," *Econometrica*, 82, 1153–1176.

Natenzon, P. (2019): "Random choice and learning," *Journal of Political Economy*, 127, 419–457.

Orlik, P. and H. Terao (2013): *Arrangements of hyperplanes*, vol. 300, Springer Science & Business Media.

Piermont, E. (2017): "Context dependent beliefs," *Journal of Mathematical Economics*, 71, 63–73.

Salant, Y. and A. Rubinstein (2008): "(A, f): Choice with Frames," *The Review of Economic Studies*, 75, 1287–1296.

Shmaya, E. and L. Yariv (2016): "Experiments on Decisions under Uncertainty: A Theoretical Framework," *American Economic Review*, 106, 1775–1801.

Tversky, A. and D. Kahneman (1981): "The framing of decisions and the psychology of choice," *science*, 211, 453–458.

Tversky, A. and I. Simonson (1993): "Context-dependent preferences," *Management Science*, 39, 1179–1189.