Random choice with framing effects: a Bayesian model

Evgenii Safonov

Abstract

I show that it is without loss of generality to consider frames of choice problems as Blackwell experiments defined on the space of subjective states of mind of the decision maker if analyst observes only random choice. The model has non-trivial empirical content when two conditions hold: (1) frames are observable and data contains choice probabilities for different menus within the same frame and (2) analyst makes a conjecture about the maximum size of subjective state space. I show that without these two conditions "no empirical content" result holds under strong structural assumptions with an application to the dynamic random choice where decision maker gradually learns her preferences.

Keywords: Random Choice, Framing Effects, Menu Effects, Context Dependence, Learning

1 Introduction

People may be uncertain about their preferences: probably, nobody knows herself/himself perfectly. Framing of a situation in which agent makes a decision may tell her something about herself, triggering a particular path of self-cognition. Are we brave enough, what song do we like most of all, and what sort of ice cream do we prefer this summer - answers on these questions may depend on the environment, and people could make different choices subject to seemingly irrelevant details. Expected utility framework is one of the classical models dealing with uncertainty, and it is tempting to apply it and see how "Bayesian" decision makers live with their varying tastes. In the presence of idiosyncratic uncertainty decisions become stochastic, and it is natural to consider random choice data in our study.

A good definition of a frame is given in Salant and Rubinstein (2008). In attempt to paraphrase it, let's say that frame is the way the particular choice problem is presented to the decision maker. There is an important difference between the case when frame contains some information about the objective qualities of alternatives and when it does not. I consider only the latter case in line with Salant and Rubinstein (2008). Under this definition, we may consider framing effects as a deviation from "rationality" represented by WARP in the deterministic setting or random utility model in the stochastic choice setting with the population interpretation of random choice. One might think that DM for some reasons chooses according to different orders of alternatives when she faces different frames. From this perspective, various theoretical models consider particular reasons of why this occurs: limited attention, reference points, attraction and similarity effects, etc. This paper studies situations when a rational expected utility maximizer might as well exhibit the same behavior; this happens when framing of a choice problem interferes with DM's process of learning her intrinsic valuations of alternatives.

Let's start with one of the framing effects studied in literature, contex dependence, when presence or absence of seemingly irrelevant alternatives in menu affects agent's decisions (Huber, Payne, and Puto (1982), Tversky and Simonson (1993)). Consider an example of context dependence that gives us intuition how uncertain tastes could lead to preference reversal across different menus. Imagine a clerk who comes to a tourist agency searching for a place to go during her vacation a month later. There are only two different options though: either a mountain trip (m) or a holiday at the beach (b). Both options have some benefits and shortcomings, but after thinking for a while she decides to go to a mountain trip: $c(\{m, b\}) = \{m\}$. In a counterfactual situation tourist agency also has a special offer: a skydiving tomorrow (p). Decision maker does not want to do anything tomorrow, neither she wants to jump with a parachute. However she imagines vividly the possibility of her falling down to the earth from the sky and realizes that it is somehow frightening. Thus she reconsiders her attitude toward danger of a mountain trip and decides to spent her vacation at the beach: $c(\{m, b, p\}) = \{b\}$. It may also be the case that she sticks to her initial choice because she realizes that she is not that afraid of height: $c(\{m, b, p\}) = \{m\}$. Alternatively, we could turn around the whole story: $c(\{m, b\}) = \{b\}$, but $c(\{m, b, p\}) = \{m\}$, because in the latter case DM finds herself to be braver person than she originally thought. And, of course, nothing prevents DM from choosing b from both menus.

We can argue that all four considered choice functions $(c(\{m, b\}), c(\{m, b, p\})) \in \{(m, m), (m, b), (b, m), (b, b)\}$ are rationalizable for a Bayesian agent in the deterministic choice model. For example, let $S = \{brave,\ usual\}$ be the space of states of mind, and utilities are given in the table below. Suppose firstly that the prior DM's belief

state of mind	u(*mountain*)	u(*beach*)	u(*parachute*)
brave	4	1	0
usual	-1	1	-5
0.5 brave + 0.5 usual	1.5	1	-2.5
0.1 brave + 0.9 usual	-0.5	1	-4.5

is $\pi(brave) = 0.5$ and that she does not get new information regarding her tastes while observing menu $\{m, b\}$. Then mountain trip is the best option for her. Now when she faces a menu including a parachute, she partially[1] learns her intrinsic state of mind and assigns posterior probability $p = 0.1$ to being brave (we consider this posterior as one of possible realizations of the distribution of posteriors from the ex-ante point of view). Then we observe a preference reversal: $c(\{m, s, p\}) = \{s\}$. To justify $c(\{m, s\}) = m,\ c(\{m, s, p\}) = s$, let vice versa the prior be $\pi'(brave) = 0.1$, and the *realization* of posterior be $p'(brave) = 0.5$.

Stochastic choice data generated by a Bayesian model provides us with some restrictions. Indeed, if we assume that introducing a third option gives DM more information about her own tastes, then the distribution of posteriors should be a mean preserving spread of the prior. In other words, if agent chooses a mountain trip when she is more

[1]This example works with full revelation of uncertainty as well (say, $p = 0$ corresponds to the agent who is certain that she is a usual type).

uncertain about her bravery, she should assign some positive probability to choose a mountain trip after she learns more about herself[2]. It is natural to assume that the resulting random choice should be consistent with the prior DM's beliefs, in other words that she forms correct expectations about herself being a brave or a usual person. There is no reason to consider Bayesian model without this consistency condition, otherwise we would not have any restrictions on the observed choice.

The assumption that frame does not reveal any information about the objective state is important. There are at least two stories about the objective information that are not considered in this paper. Firstly, it might be that framing of a choice problem directly tells DM some information that she has not known before. For example, an advertisement about skydiving says that, based on the statistics of deaths in mountains, jumping with a parachute is less dangerous than going to a mountain trip. We can say that agent is uncertain about danger of a mountain trip, and $S = \{dangerous, normal\}$ is an objective state space. In this case agent always gets a signal that mountains are objectively dangerous ($s = dangerous$) after looking at the skydiving advertisement. Thus it is possible that she never chooses mountain trip from menu $\{m, b, p\}$.

The second possibility is that menu could reveal information about the objective state of the world indirectly. For example, tourist agency could have different levels of expertise in handling risky events. The objective state space is $S = \{expert, amatuer\}$. DM thinks that if $s = expert$, tourist agency can offer skydiving, and thus when she observes menu $A = \{m, b, p\}$, the posterior probability that tourist agency has necessary expertise is high, and DM chooses $c(\{m, b, p\}) = m$. If $s = amatuer$, then, most probably, tourist agency is unable to offer parachute jumps; thus when DM observes $A = \{m, b\}$, she chooses $c(\{m, b\}) = b$. In this example menu itself serves as a signal realization, revealing the type of tourist agency. Piermont (2017) studies this case and gets "everything is possible" result: without further restrictions any choice behavior could be rationalized with the help of Bayesian frames with objective state space, when menus are associated with signal realizations.

The population interpretation of random choice is good for understanding the difference between models of Bayesian frames with and without revelation of objective information. When state space is objective (i.e. $S = \{dangerous, normal\}$ or $S = \{expert, amatuer\}$), all agents in the population get the same signal realization after observing menu $\{m, b, p\}$ (roughly speaking, "$s = dangerous$" or "$s = expert$" in the considered examples), and the same is true for menu $\{m, b\}$ ("use your prior" or "$s = amatuer$" respectively). When state space is subjective (i.e. $S = \{brave, usual\}$), we may assume that states of minds are realized independently for different agents: the presence of skydiving option only *triggers* the revelation of uncertainty for each agent, but does not serve as a correlation device for the states of minds of different people.

Our basic model (which we call a Bayesian agent with uncertain preferences, or BAUP model) consists of an unobserved by analyst subjective state space, state-dependent utility function, beliefs and signal structure which associates some Blackwell experiment with each frame. This model, as usual for models with state-dependent utilities, is highly flexible. It has a characterization in terms of one relation revealed via stochastic choice.

[2]DM could be indifferent between m and b when facing both menus and choose with the help of a tie-breaking rule. We rule out this possibility in the formal model.

We say that one alternative is revealed to dominate another if the first one is chosen for sure from some menu including both of them. Expected utility framework implies that this Revealed Dominance relation should be acyclic. Theorem 1 argues that this is also a sufficient condition for the representation of random choice by BAUP model. Hence Revealed Dominance Acyclicity axiom provides a characterization of BAUP model.

Theorem 1 is not a strong result though: almost every random choice is rationalizable by BAUP model. We then start to search for a natural structural assumptions providing some discipline to our model. One could try to put restrictions on each of the components of BAUP model separately or consider some complex restrictions. The four components are: state space, utility function, prior probability distribution and structure of signals. Prior probability distribution is not the best object to work with by two reasons: firstly, there is a classical issue with the separation of utilities and probabilities, and secondly, probability distribution is a good instrument to fit the other distribution (random choice), thus it is not wise to break it. Putting restrictions on utilities is more reasonable; however, it might be too rude way to discipline the model. One might argue that we could be happy with an "ad hoc" model providing restriction on preference flexibility of decision makers, and we do not need Bayesian structures to proceed further. Thus it would be better for our purposes if we derive restrictions on preference's flexibility of DM endogenously from some considerations involving the assumption that agents are expected utility maximizers.

We pay our attention to the last two components of BAUP model: state space and signal structure. The extensive treatment of them is given in the main part of the paper, and it is not necessary to go into details here. The main results are given by Theorems 1-4. We start by definition of fundamentals, then go over the benchmark Random Utility model, and then introduce our main framework: BAUP model. Then we formulate and prove Theorem 1. Next in Theorem 2 we show how to justify any non-degenerate random choice by a decision maker who has one order of alternatives in a given state and the reverse of that order in the another state. It turns out that this preference variation is enough even if DM has only two described above subjective states.

One important feature of this study is that we consider a situation when analyst is able to observe the frames (as in the Salant and Rubinstein (2008) paper). Working in this environment, we provide an extensive treatment of models with constrains on the structure of signals coming from the framings of choice problems. Results are summarized in "everything goes" Theorem 3 which says that reasonable constrains on information structure does not provide any discipline to the BAUP model. However the ability to observe frames allows us to get some necessary conditions imposed on the random choice which admits BAUP representation with state space of size 2. An example of such condition includes a probabilities of choice under 3 different frames: α, β, γ. Let $\rho^A_\beta(x)$ be the probability that DM chooses x from menu A when choice problem has frame β. Then the following condition holds:

$$\rho^{xy}_\alpha(x) - \rho^{xyz}_\alpha(x) + \rho^{yz}_\beta(y) - \rho^{xyz}_\beta(y) + \rho^{xz}_\gamma(z) - \rho^{xyz}_\gamma(z) < 2 \tag{1}$$

Observe that random utility model requires this polynomial to be ≤ 1, and without any restrictions it could be as large as 3. Therefore, BAUP model with state space of size 2 does provide us with moderate framing effects. Theorem 4 summarizes our findings about representations with restrictions on the size of state space, when analyst observes frames.

Propositions are given for reference and does not provide us with some insights outside of the scope of the model. Lemma 1 and Lemma 3 are key technical lemmas needed for the proofs of Theorem 3 and Theorem 4 respectively. We proceed with a review of the related literature.

1.1 Related Literature

The benchmark model of random choice for current paper is random utility model, which describes a choice of a population of rational agents whose choice function satisfies WARP. Falmagne (1978) provides a characterization of this model in terms of linear inequalities on random choice probabilities. McFadden (2005) considers random choice induced by a population of agents given by a general deterministic choice function. Framing effects are studied in a great number of papers. To name a few of them, Huber, Payne, and Puto (1982), Tversky and Simonson (1993) and relatively recent paper Salant and Rubinstein (2008) from which we borrowed a couple of definitions and examples. To the best of my knowledge, the first paper describing preference uncertainty via subjective state space is Kreps (1979) followed by a more deliberate consideration by Dekel, Lipman, and Rustichini (2001)

Somewhat more recent and related to the current study papers are Karni and Safra (2016), Lu (2016), Ahn and Sarver (2013), Manzini and Mariotti (2014), Dillenberger et al. (2014). Two papers, Shmaya and Yariv (2016) and Piermont (2017), contains some "everything is possible" results similar to Theorems 1,2 and 3 of the current paper.

There are two closely related papers which study the connection between DM's private information over subjective state space and the observed stochastic choice. Natenzon (2012) considers a DM who learns her state of mind gradually by thinking about a choice problem. In his setup signals are associated with alternatives (similar to Example 3 in the current paper). He analyzes DM's random choice with the help of "Bayesian Probit" model in continuous time focusing on the attraction and similarity effects, when introduction of similar object to the menu correspondingly increases and decreases probability to choose the given object. Natenson finds that if signals coming from the presence of these two objects are correlated and objects are sufficiently similar, then attraction effect appears at the beginning of learning process, when DM is relatively less informed about her preferences, and then it reverses to the similarity effect when time goes on. The current paper complements Natenson's analysis by considering a general non-parametric setup. Since we get "everything goes" result for the similar setup[3] (Theorem 3), but with less structure, it is interesting to analyze what particular assumptions could give us something meaningful, and to what extend they are without loss of generality.

Recent study (Frick, Iijima, and Strzalecki (2017)) deals with the dynamic random choice over lotteries with the presence of private information in great generality, including gradual learning. Current paper, in comparison, focuses more on the framing effects rather than history dependence of choice and work without assumption of linear structure, since here the fundamental objects of choice are deterministic alternatives rather than lotteries.

[3]Natenson works in continuous time, and our model lives in a finite world. Moreover, dynamic choice is not in the focus of this paper, thus our results could have relatively mild connection to those in his paper.

2 Basic Model

In the basic model we assume that analyst observes probabilities of choosing alternatives from different menus. Let X be a finite set of alternatives, and $\mathcal{F} = \{A_i\}_{i\in\mathcal{I}}$ be a finite set of menus (choice field), enumerated by some index $i \in \mathcal{I}$. We study situations in which agent's decisions to pick an alternative from a set may depend not only on the offered menu, but also on other seemingly irrelevant details (i.e., framing). Therefore we allow for repetitions in $\mathcal{F}$. That is, $A_i = A_j$ for $i \neq j$ means that analyst observes choice probabilities from the same menu, but, possibly, these choice problems appear to be different for the decision maker because they have different framing. If it is the same for both menus or if agents do not experience framing effects, we might expect choice probabilities to be the same whenever $A_i = A_j$.

Random choice $\{\rho^{A_i}(x)\}_{i\in\mathcal{I}}$ is a family of probability distributions on X with supports contained in A_i. That is,

$$\rho : \ \mathcal{I} \times X \to [0,1] \qquad , \qquad \sum_{x\in A_i} \rho^{A_i}(x) = 1 \ \ \forall i \in \mathcal{I} \tag{2}$$

We can define a space of choice probabilities $\mathcal{P}$ as a Cartesian product of the corresponding probability simplexes:

$$\mathcal{P} = \bigtimes_{i\in\mathcal{I}} \left\{ (\rho^{A_i}(x))_{x\in A_i} \in R_+^{|A_i|} \ \ | \ \ \sum_{x\in A_i} \rho^{A_i}(x) = 1 \right\} \tag{3}$$

and identify random choice ρ with a point in this space. We use the usual Euclidean metric in $\mathcal{P}$:

$$d(\rho, \hat{\rho}) \ = \ ||\rho - \hat{\rho}|| \ = \ \sqrt{\sum_{i\in I} \sum_{x\in A_i} (\rho^{A_i}(x) - \hat{\rho}^{A_i}(x))^2} \tag{4}$$

Benchmark: Random Utility model

Random order (random utility) model serves as a benchmark and also plays an important role throughout the paper. Following Barberá and Pattanaik (1986) and using slightly different notations, define D to be the set of all total orders[4] on X with generic element $d \in D$. Let $w : \ D \to [0,1]$ such that $\sum\limits_{d\in D} w(d) = 1$ be a probability distribution over total orders on X. We say that probability distribution w over total orders on X induces random choice $\{\rho^{A_i}(x)\}$ if

$$\forall i \in \mathcal{I} \qquad \rho^{A_i}(x) = \sum_{d\in D} w(d) \cdot \mathbb{1}\{xdy \ \forall y \in A_i\} \tag{5}$$

That is, probability of choosing some alternative x from menu A is equal to the probability that this alternative is maximum in A according to a random order. We say that random choice data admits random order representation if there exists a probability distribution w over total orders on X such that it induces random choice data. From eq. (5) we

[4]Complete, transitive and antisymmetric binary relations.

immediately get one necessary condition to have a random order representation: $\rho^{A_i}(x) = \rho^{A_j}(x)$ for all $A_i = A_j$.

Assume now that there are no repetitions of menus in $\mathcal{F}$. If $\mathcal{F}$ includes all non-empty subsets of X, Falmagne (1978)'s result provides a characterization of the set of stochastic choices that have random order representation. Barberá and Pattanaik (1986) formulate this result in terms of functions[5] $q_\rho(x, A)$, defined recursively for $A \in 2^X$:

$$q_\rho(x, \varnothing) = \rho^X(x) \;\; \forall\, x \in X \tag{6}$$

$$q_\rho(x, A) = \begin{cases} \rho^{X \setminus A}(x) - \sum\limits_{B \subsetneq A} q_\rho(x, B) & \text{if } x \notin A \\ 0 & \text{if } x \in A \end{cases}$$

Stochastic choice ρ admits a random order representation if and only if $q_\rho(x, A) \geq 0$ $\forall A \in 2^X$, $x \in X$. These conditions take a form of linear polynomials in $\rho^A(x)$, and therefore they define a polygon $RO = \{\rho \in \mathcal{P} \mid q_\rho(x, A) \geq 0 \;\forall\; x, A\}$ in the space of choice probabilities $\mathcal{P}$. If analyst observes only some menus, it is easy to conclude that random choice admits random order representation if and only if $\{\rho^{A_i}(x)\}$ belongs to the projection of polygon RO on the subspace of observable choices.

A closely related concept is random utility representation, which operates with a probability distribution defined on the appropriate sigma-algebra on the space of utility functions $u : \; X \to R$. We are going to use distributions of utility functions with finite support, therefore we only need to require singletons to be measurable. Following Gul and Pesendorfer (2006), we say that distribution of utility functions μ with finite support is regular if ties occur with zero probability:

$$\mu(\{U \mid U(x) = U(y) \;\; \text{for some } x \neq y\}) = 0 \tag{7}$$

Any regular distribution of utilities with finite support μ induces a distribution w over random orders on X:

$$w_\mu(d) = \mu\{U \mid xdy \;\; \text{iff} \;\; U(x) \geq U(y) \quad \forall x, y \in X\} \tag{8}$$

w_μ is well-defined because U has no ties on the support of μ. We say that random choice $\rho(.)$ admits a random utility representation if there is a regular μ such that the distribution of total orders w_μ induces $\rho(.)$.

2.1 Model of Bayesian agents with uncertain preferences

We start an introduction of our model by describing choices of a single decision maker and then we consider a population of them. As we will see, we can get the latter as a special case of the former in this setup. Decision maker (DM) is uncertain about a subjective state of the world, $s \in S$, where S is a finite state space, which we may interpret as a space of moods, or "selves", or states of mind, or any other source of preference uncertainty. DM has a state-dependent utility function $u : S \times X \to R$, which describes

[5] These functions are known as Block-Marschak polynomials. Block and Marschak (1960) paper shows that these polynomials should be non-negative whenever data is generated by some random utility (random order) model. This is one part of the Falmagne (1978)'s theorem.

her preferences over alternatives in X in each state of the world. We assume that DM is an expected utility maximizer, and thus when she faces a choice problem, she picks an alternative that gives her the highest utility given her beliefs (i.e. probability distribution on the state space S) about her tastes. Let $\pi(s)$ be a prior probability distribution of DM over S. Framing of a choice problem may help DM to think about her preferences and resolve some of the uncertainty. Since framing is external with respect to the subjective space of decision maker S, a correct way to model this uncertainty resolution is to associate each frame with a Blackwell experiment σ defined on S. Formally, let $\Sigma(S)$ be the set of Blackwell experiments on S with finite signal space[6] and $h: \mathcal{I} \to \Sigma(S)$ be a function that assigns experiment $h(i)$ to choice problem A_i; we call $h(.)$ a *framing* function.

Thus, we model the DM's choice as follows. Firstly, some state s of her mind is realized, but she does not think about it before she faces a decision problem[7]. Then DM observes a frame of a particular choice problem A_i and receives signal τ^i according to the conditional distribution $h(i)(s,\tau)$. She forms a posterior belief $p(s \mid \tau^i)$ and evaluates alternatives using expected utility $\mathbb{E}[u_s(x) \mid \tau^i]$. In principle, maximum of the expected utility is not necessary unique, and we should introduce a tie-breaking rule to figure out which alternative is picked eventually. However, the finite environment of the setup (finite number of alternatives and menus) allows us to require that this situation does not occur. Tie-breaking rule would give the model additional flexibility, but it is flexible enough. Putting things all together, a Bayesian agent with uncertain preferences (BAUP) model of choice consists of a tuple $\{S, u, \pi, h\}$ and a tie-breaking rule. We say that BAUP model is *regular* if ties occur with zero probability. Formally,

Definition 1. Regular BAUP model. Regular BAUP model defined for a set of choice problems $i \in \mathcal{I}$ is a tuple $\{S, u, \pi, h\}$ consisting of finite state space S, state-dependent utility function $u: S \times X \to R$, prior probability distribution π on S, and framing function $h: \mathcal{I} \to \Sigma(S)$ assigning Blackwell experiment on S with finite signal space for each choice problem, such that

$$\forall i \in \mathcal{I} \;\; \forall x, y \in X \quad \Pr\left(\mathbb{E}[u_s(x) - u_s(y) \mid \tau^i] = 0\right) \;=\; 0 \tag{9}$$

where τ^i is a random signal observed by DM according to Blackwell experiment $h(i)$.

There are two sources of stochasticity build in the model. Firstly, there is a random subjective state of the world s with distribution $\pi(s)$. Secondly, there are random signals with conditional distributions $h(i)(s,\tau)$ associated with framing of different choice problems $A_i \in \mathcal{F}$. We say that BAUP model $\{S, u, \pi, h\}$ induces random choice ρ if

[6] The restriction that signal space is finite is for the technical ease. I am going to relax it in the further work. Infinite signal space requires consideration of measurability of expected utilities appearing in our calculations. Observe that proofs involving construction of some Bayesian model (Theorem 2, Theorem 3, if part of Theorem 1) continue to be valid in the case when signal space is allowed to be infinite because we can use the same Blackwell experiments with finite signal space. The only if part of Theorem 1 requires some consideration, and it is discussed in the Appendix. Other results rely on the fact that expected utility is a convex combination of state-dependent utilities weighted by a posterior distribution and are not sensitive to the finiteness of a signal space.

[7] Alternatively, she could get some signal regarding her tastes. This case is discussed in the population interpretation of the random choice.

$$\forall x \in X \;\; \forall i \in \mathcal{I} \qquad \rho^{A_i}(x) = \Pr\left\{ \, \mathbb{E}[u_s(x) \mid \tau^i] \geq \mathbb{E}[u_s(y) \mid \tau^i] \;\; \forall y \in A_i \, \right\} \tag{10}$$

where τ^i is a random signal observed by DM according to Blackwell experiment $h(i)$, and

$$\mathbb{E}\left[u_s(x) \mid \tau^i\right] \; = \; \sum_{s \in S} p(s \mid \tau^i) \cdot u_s(x) \; = \; \sum_{s \in S} \left(\frac{\pi(s) \cdot h(i)(s, \tau^i)}{\sum_{s' \in S} \pi(s') \cdot h(i)(s', \tau^i)} \cdot u_s(x) \right)$$

We say that random choice ρ admits a regular BAUP representation if there exists a subjective state space S, state-dependent utility function $u_s(x)$, prior probability distribution $\pi(s)$, and framing function $h : \mathcal{I} \to \Sigma(S)$ mapping indexes of choice problems to the Blackwell experiments on S with finite signal space, such that BAUP model $\{S, u, \pi, h\}$ is regular and it induces random choice ρ.

2.2 A characterization theorem for BAUP model

The goal of this section is to characterize random choice that admits regular BAUP representation analogously to the Falmagne (1978)'s characterization of random choice consistent with some random utility model. For this purpose we firstly define a binary relation $\hat{\succeq}_\rho$ on X induced by random choice.

Definition 2. Revealed Dominance relation. Given random choice $\{\rho^{A_i}(x)\}_{i \in \mathcal{I}}$, we define a revealed by this choice dominance binary relation $\hat{\succeq}_\rho$ by

$$x \; \hat{\succeq}_\rho \; y \qquad \text{iff} \qquad x \neq y \text{ and } [\exists i \in \mathcal{I} : \; x, y \in A_i \text{ and } \rho^{A_i}(x) = 1] \tag{11}$$

Choice behavior, consistent with BAUP model, is characterized by the following axiom.

Revealed Dominance Acyclicity Axiom. Revealed Dominance relation $\hat{\succeq}_\rho$ is acyclic[8].

Theorem 1. Random choice $\{\rho^{A_i}(x)\}_{i \in \mathcal{I}}$ has a regular BAUP representation if and only if it satisfies Revealed Dominance Acyclicity Axiom.

Proof.
Only if part. Suppose ρ has a BAUP representation. Consider some pair of alternatives $x \neq y$ such that $x \; \hat{\succeq}_\rho \; y$. Then there is $i \in \mathcal{I}$ such that $\rho^{A_i}(x) = 1$ and $y \in A_i$. We thus have

$$\Pr\{\tau^i \mid \mathbb{E}[u_s(x) \mid \tau^i] > \mathbb{E}[u_s(y) \mid \tau^i]\} = 1 \tag{12}$$

therefore, using the Law of Total Expectations,

$$\mathbb{E}[u_s(x) - u_s(y)] = \mathbb{E}[\; \mathbb{E}[u_s(x) - u_s(y) \mid \tau^i] \;] > 0 \tag{13}$$

[8]Binary relation $\succeq$ is acyclic if $x_1 \succeq ... \succeq x_n$ implies $x_n \not\succeq x_1$.

Hence,

$$x \mathrel{\hat{\succsim}_\rho} y \quad \Rightarrow \quad \mathbb{E}[u_s(x)] > \mathbb{E}[u_s(y)] \tag{14}$$

If $\hat{\succsim}_\rho$ has a cycle $x_1 \mathrel{\hat{\succsim}_\rho} ... \mathrel{\hat{\succsim}_\rho} x_n \mathrel{\hat{\succsim}_\rho} x_1$, then $\mathbb{E}[u_s(x_1)] > ... > \mathbb{E}[u_s(x_n)] > \mathbb{E}[u_s(x_1)]$, contradiction. We conclude that $\{\rho^{A_i}(x)\}_{i\in\mathcal{I}}$ satisfies Revealed Dominance Acyclicity Axiom.

If part. Consider $\{\rho^{A_i}(x)\}_{i\in\mathcal{I}}$ such that $\hat{\succsim}_\rho$ has no cycles. We proceed by constructing a BAUP model $\{S, u, \pi, h\}$ that induces ρ. Observe that $\hat{\succsim}_\rho$ is a subset of some total order[9] $\succeq^*$. WLOG enumerate elements of X by $x^1, x^2, ..., x^n$ such that $x^n \succ^* x^{n-1} \succ^* ... \succ^* x^2 \succ^* x^1$. Denote

$$\mathcal{I}_d = \{i \in \mathcal{I} \mid \exists x : \rho^{A_i}(x) = 1\} \qquad ; \qquad \mathcal{I}_{nd} = \mathcal{I} \backslash \mathcal{I}_d \tag{15}$$

and WLOG assume that menus are enumerated such that $\mathcal{I}_{nd} = \{1, ..., K\}$, $\mathcal{I}_d = \{K + 1, ..., \mid \mathcal{I} \mid\}$. Consider state space

$$S = \{s_{m_1...m_K}\}_{m_i...m_K} \tag{16}$$

where $K = \mid \mathcal{I}_{nd} \mid$ indexes are such that $m_j \in \{m \mid \rho^j(x^m) > 0\}$. That is, $(m_1, ..., m_K)$ corresponds to x^{m_1} being chosen from A_1, x^{m_2} being chosen from A_2, etc. Note that j goes through all choice problems for which at least two alternatives are chosen with positive probabilities and m_j goes through alternatives that are chosen from A_j with positive probabilities. Let $m_j(s)$ be the j-th index of state s:

$$s \equiv s_{m_1(s)...m_j(s)...m_K(s)} \tag{17}$$

Denote

$$v_i^m(x) = \begin{cases} 0 & \text{if } \rho^{A_i}(x) = 0 \\ 2 \mid X \mid \cdot \left[\dfrac{1}{\rho^{A_i}(x)}\mathbb{1}\{x = x^m\} - \dfrac{1}{1-\rho^{A_i}(x)}\mathbb{1}\{x \neq x^m\}\right] & \text{if } \rho^{A_i}(x) > 0 \end{cases} \tag{18}$$

and consider the following state-dependent utility function:

$$u_s(x^q) = q + \sum_{i=1}^{K} v_i^{m_i(s)}(x^q) \tag{19}$$

Define prior probability distribution

$$\pi(s_{m_1...m_K}) = \prod_{i=1}^{K} \rho^{A_i}(x^{m_i}) \tag{20}$$

[9] We can show this by induction in the number of alternatives. Firstly, set $x \succeq^* x$ for all $x \in X$. Induction base: $X = \{x_1, x_2\}$. In case if $x \mathrel{\hat{\succsim}_\rho} y$ for some $x, y \in X$, define $x \succeq^* y$. No cycles condition for $\hat{\succsim}_\rho$ guarantees that $y \not\succeq^* x$. In the other case, define $x_2 \succeq^* x_1$. Induction base: find an alternative z such that $x \mathrel{\hat{\not\succsim}_\rho} z$ for all $x \in X\backslash\{z\}$ This alternative exists, since otherwise each alternative is revealed to be dominated by another, and we have a cycle in $\hat{\succsim}_\rho$ (recall that X is finite). Define $z \succeq^* x \; \forall x \in X$ and apply induction hypothesis to $X\backslash\{z\}$. It is easy to see that $\succeq^*$ is a total order, which includes $\hat{\succsim}_\rho$.

and consider the following messages:

$$\tau^j_m \quad = \quad \text{“ } c(A_j) = x^m \text{ ”} \tag{21}$$

and a “default” message τ_0. Let $\boldsymbol{\tau} = (\tau^1, ..., \tau^{\mathcal{I}})$ be a vector of messages corresponding to framing of choice problems $A_1, ..., A_{\mathcal{I}}$. Choose framing function h such that

$$(h(1), ..., h(\mathcal{I}))(s_{m_1 \dots m_K}, \boldsymbol{\tau}) \;\equiv \tag{22}$$

$$\equiv \; \Pr(\boldsymbol{\tau} \mid s_{m_1 \dots m_K}) = \begin{cases} 1 & \text{if } \tau^j = \tau^j_{m_j} \;\; \forall j \in \mathcal{I}_{nd} \;\text{ and }\; \tau^j = \tau_0 \;\; \forall j \in \mathcal{I}_d \\ 0 & \text{otherwise} \end{cases}$$

In other words, when DM faces choice problem A_j with $j \in \mathcal{I}_{nd}$ at state $s_{\dots m_j \dots}$, she receives signal “$c(A_j) = x^{m_j}$” with probability one. And if she faces A_j with $j \in \mathcal{I}_d$ at any state of the world, she receives signal τ_0 with probability one.

Let's show that $\{S, u, \pi, h\}$ represents ρ. Suppose that subjective state is $s = s_{r_1 \dots r_K}$. When DM faces menu A_j with $j \in \mathcal{I}_{nd}$, she receives signal $\tau^j = \tau^j_{r_j}$ with probability one and forms a posterior belief

$$p(s_{m_1 \dots m_K} \mid \tau^j = \tau^j_{r_j}) \quad = \quad \mathbb{1}\{m_j = r_j\} \cdot \prod_{i \in \mathcal{I}_{nd} \setminus \{j\}} \rho^{A_i}(x^{m_i}) \tag{23}$$

If she faces menu A_j with $j \in \mathcal{I}_d$, her posterior belief remains equal to her prior $\pi(s)$. Suppose $j \in \mathcal{I}_{nd}$. Then

$$\mathbb{E}[u_s \mid \tau^j = \tau^j_{r_j}] \; = \; q + \sum_{i \in \mathcal{I}_{nd}} \mathbb{E}[v_i^{m_i(s)}(x^q) \mid \tau^j = \tau^j_{r_j}] \tag{24}$$

Consider $i \neq j$ and $\rho^{A_i}(x^q) > 0$, then

$$\mathbb{E}[v_i^{m_i(s)}(x^q) \mid \tau^j = \tau^j_{r_j}] \quad = \quad \sum_{s:\; m_j(s) = r_j} \left(\prod_{k \in \mathcal{I}_{nd} \setminus \{j\}} \rho^{A_k}(x^{m_k(s)}) \;\; \times \right. \tag{25}$$

$$\left. \times \;\; 2 \mid X \mid \cdot \left[\frac{1}{\rho^{A_i}(x^q)} \mathbb{1}\{q = m_i(s)\} - \frac{1}{1 - \rho^{A_i}(x^q)} \mathbb{1}\{q \neq m_i(s)\} \right] \right) \quad =$$

$$= \;\; 2 \mid X \mid \sum_{s:\; m_j(s) = r_j \;\; , \;\; m_i(s) = w} \left(\prod_{k \in \mathcal{I}_{nd} \setminus \{j, i\}} \rho^{A_k}(x^{m_k(s)}) \;\; \times \right.$$

$$\left. \times \;\; \underbrace{\left[\sum_{w:\; \rho^{A_i}(x^w) > 0} \rho^{A_i}(x^w) \left[\frac{1}{\rho^{A_i}(x^q)} \mathbb{1}\{q = w\} - \frac{1}{1 - \rho^{A_i}(x^q)} \mathbb{1}\{q \neq w\} \right] \right]}_{=\,0} \right) \quad = \quad 0$$

Analogous calculation shows that if $\rho^{A_j}(x^q) > 0$, then

$$\mathbb{E}[v_j^{m_j(s)}(x^q) \mid \tau^j = \tau^j_{r_j}] = \begin{cases} -\dfrac{2 \mid X \mid}{1-\rho^{A_j}(x^q)} & \text{if } q \neq r_j \\ \\ \dfrac{2 \mid X \mid}{\rho^{A_j}(x^q)} & \text{if } q = r_j \end{cases} \tag{26}$$

Therefore, for $j \in \mathcal{I}_{nd}$ we have

$$\mathbb{E}[u_s(x^q) \mid \tau^j = \tau^j_{r_j}] = q + \begin{cases} 0 & \text{if } \rho^{A_j}(x^q) = 0 \\ \\ -\dfrac{2 \mid X \mid}{1-\rho^{A_j}(x^q)} & \text{if } \rho^{A_j}(x^q) > 0 \text{ and } q \neq m \\ \\ \dfrac{2 \mid X \mid}{\rho^{A_j}(x^q)} & \text{if } q = r_j \end{cases} \tag{27}$$

Note that $\dfrac{2 \mid X \mid}{\rho^{A_j}(x^q)} > q'$, $\forall q' \in \{1, ..., \mid X \mid\}$. Thus

$$\underset{x \in A_j}{\text{Argmax}}\{\mathbb{E}[u_s(x) \mid \tau^j = \tau^j_{r_j}]\} = \{x^{r_j}\} \tag{28}$$

Using these results, we find that

$$\Pr\left\{\ \mathbb{E}[u_s(x^q) \mid \tau^j] \geq \mathbb{E}[u_s(y) \mid \tau^j] \ \ \forall y \in A_j\ \right\} \ \ = \tag{29}$$

$$= \ \ \Pr\left\{\tau^j = \tau^j_q\right\} \ \ = \ \ \pi\left(\{s \in S \mid m_j(s) = q\}\right) \ \ = \ \ \rho^{A_j}(x^q)$$

which means that $\{S, u, \pi, h\}$ induces random choice $\rho^{A_j}(x)$ for menus A_j with $j \in \mathcal{I}_{nd}$.

Now consider $j \in \mathcal{I}_d$. When DM faces A_j, she receives uninformative signal τ^0 and therefore

$$\mathbb{E}[u(x^q) \mid \tau^j = \tau_0] = \mathbb{E}[u(x^q)] = q \tag{30}$$

Thus

$$\underset{x \in A_j}{\text{Argmax}}\{\mathbb{E}[u_s(x) \mid \tau^j = \tau_0]\} = \{x^{M(j)}\} \qquad , \qquad M(j) = \underset{x^m \in A_j}{\text{argmax}}(m) \tag{31}$$

Hence,

$$\Pr\left\{\ \mathbb{E}[u_s(x^q) \mid \tau^j] \geq \mathbb{E}[u_s(y) \mid \tau^j] \ \ \forall y \in A_j\ \right\} \ \ = \tag{32}$$

$$= \ \ \Pr\left\{\tau^j = \tau_0\right\} \cdot \mathbb{1}\{q = M(j)\} \ \ = \ \ \mathbb{1}\{q = M(j)\}$$

Note that A_j is such that $\rho^{A_j}(x^w) = 1$ for some $x^w \in A_j$. Assume that $w \neq M(j)$. Then

$$x^w \ \hat{\succeq}_\rho \ x^{M(j)} \quad \Rightarrow \quad x^w \ \succeq^* \ x^{M(j)} \quad \Rightarrow \quad w > M(j) \tag{33}$$

contradicting the definition of $M(j)$. Therefore, $\rho^{A_j}(x^{M(j)}) = 1$, and thus $\{S, u, \pi, h\}$ induces random choice $\rho^{A_j}(x)$ for menus A_j with $j \in \mathcal{I}_d$ as well. The if part is proven. Theorem 1 is proven.

Each model of choice M has a corresponding subspace $\mathcal{P}(M) \subseteq \mathcal{P}$ of choice probabilities that could be induced by this model:

$$\mathcal{P}(M) = \{\rho \in \mathcal{P} \mid \rho \text{ could be induced by model M}\} \tag{34}$$

If $M = M_{RUM}$ is a random utility model, then $\mathcal{P}(M_{RUM}) = RO$ is a polygon in the space of choice probabilities $\mathcal{P}$ such that at each point of this polygon all Block-Marschak polynomials are non-negative. Clearly, random utility is a special case of BAUP model such that the subjective state is fully revealed to DM when she makes a decision. This is summarized in Proposition 1.

Proposition 1. Let BAUP model M_{perfect} be such that $h(i) = \sigma_{\text{perfect}}$ $\forall i \in \mathcal{I}$, where $\sigma_{\text{perfect}}(s, s') = \mathbb{1}\{s = s'\}$ $\forall s, s' \in S$; in other words, DM learns her subjective state of mind perfectly. Then M_{perfect} could induce the same set of random choices as random utility model:

$$\mathcal{P}(M_{\text{perfect}}) \quad = \quad RO \tag{35}$$

The proof is simple and omitted.

If $M = M_{BAUP}$ is a BAUP model, then, according to Theorem 1, $\mathcal{P}(M_{BAUP})$ is a set of choice probabilities such that $\hat{\succeq}_\rho$ is acyclic. One might argue that it is a very light requirement. Indeed, BAUP model is incompatible with random choice if there are $x_1, ..., x_n \in X$ and $i_1, ..., i_n \in \mathcal{I}$ such that $x_k, x_{k+1 \text{ mod}(n)} \in A_k$ and $\rho^{A_k}(x_k) = 1$. Suppose we have a finite data set of choice decisions. Then any test of hypothesis $H_0: \ \rho^A(x) < 1$ versus $H_1: \ \rho^A(x) = 1$ has a trivial power. This is a general observation: since probabilities are estimated with errors, there is no good way to reject one model in favor of another if the corresponding sets of choice probabilities induced by two models have the same closures. Let's formalize this concept.

Definition 3. Almost indistinguishable models. We say that two models of random choice, M_1 and M_2, are almost indistinguishable if $cl(\mathcal{P}(M_1)) = cl(\mathcal{P}(M_2))$.[10]

So far we considered a situation in which analyst had no exogenously given information about frames of choice problems. It is natural to say that model $M = M_{arbitrary}$ exhibits arbitrary framing effects in this setup if $\mathcal{P}(M_{arbitrary}) = \mathcal{P}$, that is, we can see any random choice. We may conclude from Theorem 1 that BAUP model is almost indistinguishable from a model of arbitrary framing effects. It is a negative ("everything goes") result from the point of view of falsification of the model. The rest of the paper is devoted to finding ways to discipline BAUP model by putting restrictions on its components (state space, utilities and framing function) and considering richer choice data such that BAUP model becomes distinguishable from a model of arbitrary framing effects. Before we come to the examples of BAUP models that give moderate framing effects, we consider a couple of "everything goes" theorems for the cases when we put restrictions on the size of state space and on the framing function correspondingly; the latter model in-

[10]Where closure is taken with respect to topology on $\mathcal{P}$ induced by Euclidean metric (eq. 4).

cludes dynamic random choice as a special case. Before that we show that it is enough to have subjective state space with only two states ($|S| = 2$) for the BAUP model to justify any non-degenerate random choice (and thus this model is indistinguishable with model of arbitrary framing effects), provided that analyst has no information about frames. But firstly we discuss how to apply BAUP model to a population of decision makers who are uncertain about their own preferences.

2.3 Population interpretation of BAUP model

Population interpretation of random choice is important for the BAUP model. Indeed, suppose that analyst observes sequential choices made by the same person from menus A_1, A_2, A_3. When DM chooses an alternative from A_1, she observes signal realization τ^1 associated with framing $h(1)$. When she faces A_2, she observes signal realization τ^2 associated with framing $h(2)$, but she also knows τ^1 at the time when she makes her choice from A_2. Similarly the analyst sees choice from menu A_3 under the condition that DM has information given by realizations of all 3 signals τ^1, τ^2, τ^3. The problem is that analyst does not see choice from menu A_2 with framing given by $h(2)$ and choice from A_3 with framing given by $h(3)$. Therefore only the very first choice of a person is made under the actual framing of the BAUP model, and framings associated with the other choice problems are contaminated by the framings of choice problems that DM faced before.

A good way to avoid this complication might be to have a large group of decision makers and give each of them one choice problem. This requires us to consider random choice exhibited by a population of decision makers. The simplest way to model this is to say that we have some distribution $\nu(\theta)$ defined on the space of types $\theta \in \Theta$, where we assume Θ to be finite. The observed random choice is then a convex combination of choice frequencies for each type θ in the population:

$$\rho^{A_i}(x) \quad = \quad \sum_{\theta \in \Theta} \nu(\theta) \cdot \rho_\theta^{A_i}(x) \tag{36}$$

where $\rho_\theta^{A_i}(x)$ is random choice that could be exhibited by individual decision maker of type θ. If $\mathcal{P}(M)$ is the set of choice probabilities that could be induced by model M describing choices of an individual agent, then

$$\mathcal{P}^{\text{population}}(M) \quad = \quad \operatorname{conv}\left(\mathcal{P}(M)\right) \tag{37}$$

is the set of choice probabilities that could be induced by choices of population of decision makers acting according to model M, and it is equal to a convex closure of $\mathcal{P}(M)$. We can see that $\mathcal{P}(M_{BAUP})$ is a convex set ($\hat{\succeq}_\rho$ being acyclic is preserved under convex combinations), therefore Theorem 1 works well for the population interpretation of random choice as well.

BAUP model provides a good way to incorporate population interpretation in itself. Indeed, consider a BAUP model in which DM receives some signal τ_0 regarding her tastes according to the conditional distribution $\sigma_0(s, \tau)$ before she faces any choice problem. We can say that there are types $\theta \in \Theta$ of decision makers who receive signals "your type is θ". Different types thus have different information before facing choice problems, and decision makers of the same type have identical information. There is one caveat though: signal τ_0 may provide some hidden information that may be used in combination with a signal received after observing framing of a choice problem[11].

2.4 Representation with two subjective states

Suppose DM is uncertain about some aspect of her personality: whether she is brave or not, is she depressed or in the good mood, does she like specific sea food or not, what is her attitude toward some policy. Depending on the circumstances she may look at this aspect from various angles and to a different extend. Thus her decision to undertake a risky trip, choose a way to spend a weekend, go to a specific restaurant or vote for a politician may depend on the seemingly irrelevant details framing the corresponding choice problem. We can use a BAUP model with two states to model her choice: $S = \{brave,\ usual\}$, $S = \{depressed,\ good\}$, $S = \{like\ octopus,\ don't\ like\ octopus\}$, $S = \{support\ universal\ income,\ reject\ universal\ income\}$. Unfortunately, this restriction of the model does not allow by itself to narrow the set of possible random choice frequencies, which is proven in Theorem 2.

Theorem 2. Any non-degenerate random choice ρ (that is, $\rho^{A_i}(x) > 0\ \forall x \in A_i\ \forall i \in \mathcal{I}$) has a regular BAUP representation $\{S, u, \pi, h\}$ with state space consisted of two states $S = \{s_1, s_2\}$. Moreover, this representation exists for any non-degenerate prior $\overline{\nu} \equiv \pi(s_2) \in (0, 1)$.

Corollary. Regular BAUP model with a state space of size two (call it M^2) is almost indistinguishable from arbitrary framing effects model.

Proof. Since choices from singletons are trivial ($\rho^{\{x\}}(x) = 1$) and they are induced by any BAUP model (any singleton set $\{x\}$ has maximizer x), we may assume without loss of generality that each menu consists of at least two elements ($\mid A_i \mid \geq 2\ \forall i \in \mathcal{I}$).

Step 1. Enumerate arbitrary elements of X by $x^1, ..., x^n$. Denote $U_\mu(x) \equiv E_\mu[u_s(x)] = (1-\mu) \cdot u_{s_1}(x) + \mu \cdot u_{s_2}(x)$. Then for any $0 < \mu_1 < ... < \mu_{n-1} < 1$ the following state-dependent utility function

$$u_1(x^k) = -\sum_{i=1}^{k-1} \left(\frac{\mu_i}{1-\mu_i} \right) \qquad u_2(x^k) = k \tag{38}$$

is such that:

1. $\mu \in [0, \mu_1) \quad \Rightarrow \quad U_\mu(x^k) > U_\mu(x^l) \quad \forall l > k$
2. $\mu \in (\mu_{k-1}, \mu_k) \quad \Rightarrow \quad \{x^k\} = \underset{x \in X}{\text{Argmax}}\ U_\mu(x)$
3. $\mu \in (\mu_{n-1}, 1] \quad \Rightarrow \quad U_\mu(x^k) < U_\mu(x^l) \quad \forall l > k$

[11] Nevertheless, any model with a prior signal coming from Blackwell experiment σ_0 and framing function $h(i)$ is outcome equivalent to the model where prior signal is uninformative and framing function maps to a compound Blackwell experiment $\overline{h}(i) = h(i) \cup \sigma_0$. For the model with a single decision maker this means that the assumption that DM does not receive any signals regarding her state of mind before facing a choice problem is without loss of generality.

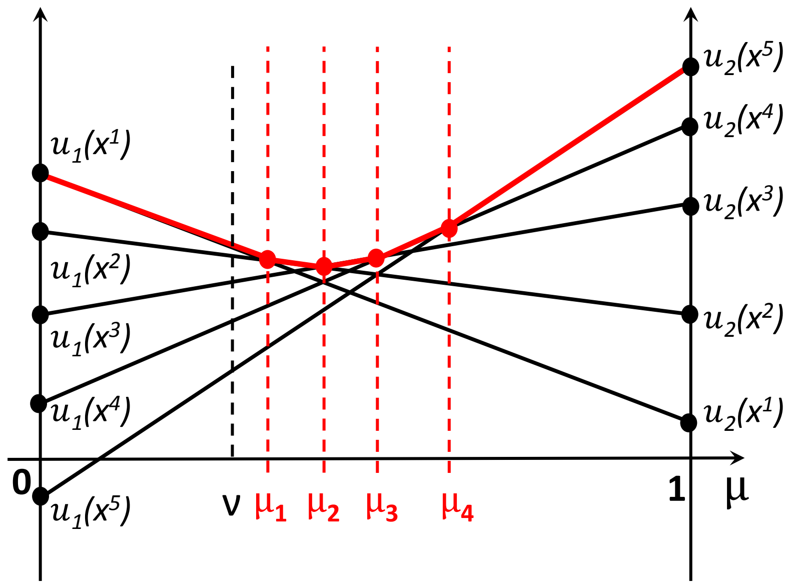

Figure 1: State-dependent utility function used in the proof of Theorem 2 (example with five alternatives). Horizontal Axis is the probability of state s_2. Vertical Axis is the expected utility of an alternative.

The corresponding calculation is given in the Appendix.

Step 2. Let $X = \{x^1, ..., x^n\}$ and

$$\epsilon = \min \left\{ \ \frac{1}{n} \cdot \min_{i \in \mathcal{I}} \min_{x \in A_i} \rho^{A_i}(x) \quad , \quad \frac{1}{\overline{\nu}} - 1 \ \right\} \tag{39}$$

Note that $\epsilon > 0$ under the conditions of Theorem 2. For each $i \in \mathcal{I}$ denote

$$\mathcal{J}_i = \{j \in \{1, ..., n\} \mid x^j \in A_i\} \qquad m_i = \min(\mathcal{J}_i) \qquad M_i = \max(\mathcal{J}_i) \tag{40}$$

and

$$\mu_k = \overline{\nu} \cdot (1 + \frac{k}{2n}\epsilon) \qquad \in \quad (\overline{\nu}, 1) \tag{41}$$

for $k \in \{1, ..., n-1\}$. Next, for each $i \in \mathcal{I}$ and $k \in \mathcal{J}_i$ define

$$\nu_i^k = \begin{cases} \overline{\nu} \cdot (1 + \frac{k - \frac{1}{2}}{2n}\epsilon) & \text{if} \ \ k \in \mathcal{J}_i \backslash \{m_i\} \\ \overline{\nu} \cdot \left(1 - \frac{\epsilon}{\rho^{A_i}(x^{m_i})} \quad \cdot \sum_{k \in \mathcal{J}_i \backslash \{m_i\}} \frac{k - \frac{1}{2}}{2n} \cdot \rho^{A_i}(x^k) \right) & \text{if} \ \ k = m_i \end{cases} \tag{42}$$

Then Blackwell experiments σ_i, having the following conditional distributions, are well-defined:

$$\sigma_i(s_1, \tau_k^i) = \rho^{A_i}(x^k) \frac{1 - \nu_i^k}{1 - \overline{\nu}} \qquad \sigma_i(s_2, \tau_k^i) = \rho^{A_i}(x^k) \frac{\nu_i^k}{\overline{\nu}} \qquad \text{for } k \in \mathcal{J}_i \tag{43}$$

The proof is given in the Appendix.

Step 3. Let's show that BAUP model $\{\{s_1, s_2\}, u_s(x), \overline{\nu}, h\}$, where $u_s(x)$ is a state-dependent utility function defined at step 1 for μ_k given by eq. (41) and h is a framing function such that $h(i) = \sigma_i$, induces random choice $\{\rho^{A_i}(x)\}_{i \in \mathcal{I}}$. We have:

$$p(s_2 \mid \tau_k^i) \;=\; \frac{\sigma_i(s_2, \tau_k^i)\overline{\nu}}{\sigma_i(s_2, \tau_k^i)\overline{\nu} + \sigma_i(s_1, \tau_k^i)(1 - \overline{\nu})} \;=\; \nu_i^k \tag{44}$$

Consider $k \in \mathcal{J}_i \backslash \{m_i, n\}$. Therefore, $1 < k < n$ and $\mu_{k-1} < \nu_i^k < \mu_k$. Statement 2 of step 1 then gives us

$$\operatorname*{Argmax}_{x \in A_i} \mathbb{E}[u_s(x) \mid \tau^i = \tau_k^i] \;=\; \operatorname*{Argmax}_{x \in A_i} U_{\nu_k^i}(x) \;=\; \{x^k\} \tag{45}$$

If $k = m_i$, then[12] $0 < \nu_i^{m_i} < \overline{\nu} < \mu_1$, and statement 1 of step 1 gives us

$$\operatorname*{Argmax}_{x \in A_i} \mathbb{E}[u_s(x) \mid \tau^i = \tau_{m_i}^i] \;=\; \operatorname*{Argmax}_{x \in A_i} U_{\nu_{m_i}^i}(x) \;=\; \{x^{m_i}\}$$

If $k = n \in \mathcal{J}_i$, then $\mu_{n-1} < \nu_i^n < 1$, and statement 3 of step 1 gives us

$$\operatorname*{Argmax}_{x \in A_i} \mathbb{E}[u_s(x) \mid \tau^i = \tau_n^i] \;=\; \operatorname*{Argmax}_{x \in A_i} U_{\nu_n^i}(x) \;=\; \{x^{M_i}\} = \{x^n\}$$

Finally,

$$\Pr(\tau^i = \tau_k^i) \;=\; \sigma_i(s_1, \tau_k^i)(1 - \overline{\nu}) + \sigma_i(s_2, \tau_k^i)\overline{\nu} \;=\; \rho^{A_i}(x^k) \tag{46}$$

Putting things all together, we get

$$\Pr\left\{\; \mathbb{E}[u_s(x^k) \mid \tau^i] \geq \mathbb{E}[u_s(y) \mid \tau^i] \;\; \forall y \in A_i \;\right\} \;=\; \Pr\left(\tau^i = \tau_k^i\right) \;=\; \rho^{A_i}(x^k) \tag{47}$$

Note also that ties in $U_\nu(x)$ do not occur for posteriors $\nu = \nu_i^k$ appearing after signal realizations τ_k^i. Therefore, BAUP model $\{\{s_1, s_2\}, u, \overline{\nu}, h\}$ is regular and it induces random choice ρ. Theorem 2 is proven.

The Corollary follows from the fact that the set of non-degenerate random choices $\mathcal{P}_{nd}$ is equal to the interior of the space of choice probabilities $\mathcal{P}$. Theorem 2 says that $\mathcal{P}_{nd} \subseteq \mathcal{P}(M^2)$. Therefore

$$\mathcal{P} \;=\; cl(\mathcal{P}_{nd}) \;\subseteq\; cl(\mathcal{P}(M^2)) \;\subseteq\; \mathcal{P} \qquad \Rightarrow \qquad cl(\mathcal{P}(M^2)) = \mathcal{P} \tag{48}$$

We conclude that M^2 is almost indistinguishable from a model of arbitrary framing effects. Corollary is proven.

3 Random choice with data on frames

In the basic model we assumed that analyst had no exogenous information about the framing of different choice problems. In a way, we determined frames endogenously

[12] See proof of step 2 in the Appendix.

as a part of model in the BAUP representations, which we considered in the basic setup. Since we did not impose any restrictions on these frames, we used this freedom to set up idiosyncratic frame for each choice problem. It was one of the drivers of "everything is possible" results that we have got so far. In this section we consider a situation in which analyst observes not only the distribution of choices from different menus, but also framings of the corresponding choice problems. We borrow some examples of how analyst can define a frame from Salant and Rubinstein (2008)[13]:

(i) Frame is a default alternative specified for a choice problem.
(ii) Frame is an order in which alternatives are presented to the decision maker.
(iii) Frame is an advertisement[14].
(iv) Frame is a maximum amount of time that DM can spend on a choice problem.

In principle, nothing prevents DM from exhibiting framing effects that are not under the control of analyst; this would return us to the basic scenario. In this section we *assume* that analyst has full information about frames. Therefore we expect that choice frequencies from menus associated with the same frame should be rationalizable by some random utility model. For the ease of analysis we also assume that analyst observes choices from all possible menus in each frame[15].

To formalize these concepts, let's follow Salant and Rubinstein (2008) and define an *extended choice problem* (A, β) to be a pair consisting of menu A and frame $\beta \in \mathcal{B}$, where $\mathcal{B}$ is a finite set of frames. Next, define *extended random choice*[16] $\{\rho_\beta^A(x)\}_{\beta\in\mathcal{B},\ \varnothing\neq A\subseteq X}$ to be a family of probability distributions on X indexed by frame $\beta \in \mathcal{B}$ and menu $A \in 2^X\backslash\{\varnothing\}$ with supports contained in A:

$$\rho: \quad \left(2^X\backslash\{\varnothing\}\right) \times \mathcal{B} \times X \to [0,1] \quad , \quad \sum_{x\in A}\rho_\beta^A(x) = 1 \ \ \forall\beta \in \mathcal{B},\ \varnothing \neq A \subseteq X \tag{49}$$

Frame Consistency Axiom. We say that an extended random choice $\{\rho_\beta^A(x)\}_{\beta\in\mathcal{B},\ \varnothing\neq A\subseteq X}$ is frame consistent if for any fixed $\beta \in \mathcal{B}$ random choice $\{\rho_\beta^A(x)\}_{\varnothing\neq A\subseteq X}$ admits a random utility representation.

As we have argued before, we consider extended random choice that satisfies Frame Consistency Axiom in this section. Define the space of choice probabilities $\mathcal{P}$ as a Cartesian product of the corresponding probability simplexes analogously to the one in the basic setup:

$$\mathcal{P} \;=\; \mathop{\times}_{\beta\in B}\left(\mathop{\times}_{A:\ \varnothing\neq A\subseteq X}\left\{(\rho_\beta^A(x))_{x\in A} \in R_+^{|A|} \;\middle|\; \sum_{x\in}\rho_\beta^A(x) = 1\right\}\right) \;=\; \mathop{\times}_{\beta\in B}\mathcal{P}_\beta \tag{50}$$

[13]Status Quo, List, Advertisement and Deadline examples in Salant and Rubinstein (2008).

[14]We should be careful here because advertisement can provide some objective information as well, which is not covered by our analysis in this paper.

[15]This is in line with standard assumption that choice field is $2^X\backslash\{\varnothing\}$ for finite set of alternatives X. We can get results for less rich data containing only some menus $A \subseteq X$ by considering the projections of corresponding sets of choice probabilities on the subspace of observable random choices.

[16]Similar to the extended choice function $c(A, f)$ in Salant and Rubinstein (2008).

where

$$\mathcal{P}_\beta = \underset{A:\ \varnothing\neq A\subseteq X}{\times}\left\{(\rho^A_\beta(x))_{x\in A} \in R^{|A|}_+ \ \mid \ \sum_{x\in} \rho^A_\beta(x) = 1\right\} \tag{51}$$

Extended random choice that satisfies Frame Consistency axiom is a subset of this space $\mathcal{P}^{\text{consistent}} \subset \mathcal{P}$:

$$\mathcal{P}^{\text{consistent}} \quad = \quad \underset{\beta\in\mathcal{B}}{\times} RO_\beta \quad = \tag{52}$$

$$= \quad \{\rho \in \mathcal{P} \mid \{\rho^A_\beta(x)\}_{\varnothing\neq A\subseteq X} \text{ admits random utility representation } \forall\beta\in\mathcal{B}\}$$

where $RO_\beta \subset \mathcal{P}_\beta$ is a polygon in the space of choice probabilities for menus in frame β consisting of points for which Falmagne's conditions are satisfied. We say that choice model $M = M_{arbitrary}$ admits arbitrary consistent framing effects if $\mathcal{P}(M) = \mathcal{P}^{\text{consistent}}$.

We may consider extended random choice $\{\rho^A_\beta(x)\}_{\beta\in\mathcal{B},\ \varnothing\neq A\subseteq X}$ as a random choice $\{\rho^{A_i}(x)\}_{i\in\mathcal{I}}$ defined in the basic model with index set

$$\mathcal{I} \quad = \quad \mathcal{I}_\mathcal{B} \quad \equiv \quad \{i = (A', \beta) \mid A' \in 2^X \backslash\{\varnothing\}\ ,\ \beta \in \mathcal{B}\} \tag{53}$$

Thus we have

$$\rho^{A'}_\beta(x) \quad = \quad \rho^{A_{(A',\beta)}}(x) \tag{54}$$

connecting definitions of extended random choice in the setup with observable frames and random choice in the basic setup. This allows us to use the definition of BAUP model from the basic setup for the extended random choice data.

The idea of BAUP representation is that framing of a choice problem affects the way DM learns her preferences. Thus we may argue that if framings of two choice problems are the same, this should be reflected in the framing function of the BAUP representation; that is, these two choice problems should be associated with the same Blackwell experiment. We should be more precise here though. Indeed, if we think about Blackwell experiment as a conditional distribution $\Pr\{\tau^i = \tau \mid s\} = h(i)(s,\tau)$, then two menus with the same frame β could give different signal realizations notwithstanding the fact that conditional distributions of these signals are the same. This does not affect results if we observe only a family of marginal probabilities distributions (random choice, extended random choice) but do not consider their joint distribution. The latter can matter for the dynamic random choice if analyst observes sequential decisions of each person in a population. The other interpretation is that every choice problem in the frame is associated with the same random signal defined as a measurable function on the underlying probability space Ω ($\tau^\beta : \ \Omega \to T$ with T being a finite space of possible signal realizations). We assume that firstly $\omega \in \Omega$ realizes and then the analyst observes random choice; therefore signal realizations are the same across menus with the same frame. We will mostly use this modeling approach; we formulate it in Definition 4. For the sake of completeness, we also formulate the first approach, where choice problems with the same frame could give rise to different signal realizations.

Definition 4. Strong frame consistency of BAUP model. Given a set of extended choice problems indexed by $i \in \mathcal{I}_\mathcal{B}$, we say that BAUP model $\{S, u, \pi, h\}$ is strongly consistent with frames if $h(A,\beta) = h(A',\beta) = h_\beta$ for all $A, A' \in 2^X\backslash\{\varnothing\}$ in the sense of random signal. In other words, all choice problems with the same frame β are associated with the same random signal τ^{β}, which has conditional distribution $\Pr\{\tau^{\beta} = \tau \mid s\} = h_{\beta}(s, \tau)$, and DM observes the same realization of τ^{β} for each choice problem in frame β.

Definition 4'. Weak frame consistency of BAUP model. Given a set of extended choice problems indexed by $i \in \mathcal{I}_{\mathcal{B}}$, we say that BAUP model $\{S, u, \pi, h\}$ is weakly consistent with frames if $h(A, \beta) = h(A', \beta) = h_{\beta}$ for all $A, A' \in 2^X \backslash \{\varnothing\}$ in the sense of conditional distributions. In other words, all choice problems with the same frame β are associated with the same conditional distribution $\Pr\{\tau^{A,\beta} = \tau \mid s\} = h_{\beta}(s, \tau)$, but the corresponding signals could have arbitrary joint distribution subject to this condition.

Observe that Definition 4 is stronger: it (obviously) implies Definition 4', but not the other way around. The difference between strong and weak versions of frame consistency of BAUP model matters only if analyst observes some joint distribution of choice probabilities. This is summarized in Proposition 2 below:

Proposition 2. For any weakly frame consistent regular BAUP model $\{s, u, \pi, h\}$ there exists a strongly frame consistent regular BAUP model $\{s, u, \pi, \tilde{h}\}$ such that: (i) $\tilde{h}(A, \beta)$ has the same conditional distribution as $h(A, \beta)$ for all $A \in 2^X \backslash \{\varnothing\}$, $\beta \in \mathcal{B}$; (ii) it induces the same marginal conditional distributions of expected utility functions as the initial model; (iii) it induces the same extended random choice as the initial model.

The proof is simple and omitted in this version of the paper.

The notions of frame consistency for random choice and BAUP model are closely connected. In fact, the latter implies the former, as is proven in Proposition 3.

Proposition 3. Given a set of extended choice problems indexed by $i \in \mathcal{I}_{\mathcal{B}}$, any regular weakly consistent with frames BAUP model $\{S, u, \pi, h\}$ induces frame consistent extended random choice ρ.

Proof. Consider arbitrary frame $\beta \in \mathcal{B}$. Weak frame consistency of BAUP model implies that the following distribution μ over utility functions (μ has a finite support) is the same for all $A \in 2^X \backslash \{\varnothing\}$:

$$\mu_{\beta}(U) \quad = \quad \Pr\{\ \tau^{A,\beta} \ \mid \ \mathbb{E}[u_s(x) \mid \tau^{A,\beta}] = U(x)\ \forall x \in X\ \} \tag{55}$$

Observe that random utility μ_{β} is regular (that is, probability of a tie is zero), because the original BAUP model is regular. Therefore μ_{β} induces some random choice $\{\hat{\rho}^A_{\beta}\}_{A:\ \varnothing \neq A \subset X}$. It is easy to see that $\hat{\rho}$ is induced by the original BAUP model. Indeed,

$$\forall \beta \in \mathcal{B},\ A \in 2^X \backslash \{\varnothing\} \qquad \Pr\left\{\ \mathbb{E}[u_s(x) \mid \tau^{A,\beta}] \geq \mathbb{E}[u_s(y) \mid \tau^{A,\beta}]\ \ \forall y \in A\ \right\} \quad = \tag{56}$$

$$= \quad \mu_{\beta}\{U:\ U(x) \geq U(y)\ \forall y \in A\} \quad = \quad \hat{\rho}^A_{\beta}(x)$$

Thus extended random choice $\hat{\rho}$ induced by the considered BAUP model admits random utility representation for any fixed β. Proposition 3 is proven.

Corollary. Proposition 3 is true for the strongly frame consistent regular BAUP model.

The rest of the paper is devoted to the studies of conditions under which BAUP models with consistent frames are almost indistinguishable from arbitrary consistent framing effects model, and under which, in contrast, there are some constraints on the extended random choice induced by BAUP models. We will consider only extended random choices that satisfy Frame Consistency axiom and BAUP models consistent with frames, therefore we will omit words "frame consistency" for the ease of notation whenever it will not lead to ambiguity in interpretations.

3.1 BAUP model with information structure

BAUP model is an information model in the first place. Therefore it is interesting how we can apply it in the situations where it is natural to make additional assumptions on the content of information that DM receives facing this or that choice problem. Recall that we used one such assumption in the definition of BAUP model consistent with frames. But what if we have some additional structure on top of this? We motivate this section by three examples.

Example 1. Dynamic random choice with frames[17]. Suppose that analyst observes a sequence of choices of decisions makers. For example, one person goes to a shop on Monday and faces extended choice problem (A_1, β_1). Then she goes to a shop on Thursday and faces (A_4, β_4), and finally on Friday she decides what to choose from A_5 given framing β_5. In the BAUP framework we should say that she deals with "efficient" frame[18] $\beta^*_{14} = \beta_1 \cup \beta_4$ when she makes her choice on Thursday, and similarly she receives a signal associated with efficient frame $\beta^*_{145} = \beta_1 \cup \beta_4 \cup \beta_5$ on Friday because she has learned some information in the previous days. Suppose that analyst is able to keep track of sequential decisions of each individual, then in this example analyst observes the probability of different vectors of choice functions $\rho^{(A_1,\beta_1),(A_4,\beta_4),(A_5,\beta_5)}(x_1, x_4, x_5)$ and similar probabilities for individuals who go to the shop every day, or only on Wednesday, or on Tuesday and Thursday, etc. Of course, menus and frames may also vary within each day.

Example 2. Some frames are more informative than others. Suppose that observation of some particular alternatives in menu triggers DM's thought process, helping to resolve some of her uncertainty about preferences. We may argue then that larger menus should provide DM with more information than smaller ones; it is also possible that DM spends more time on making choice from larger menus. The opposite case is a *limited attention* consideration, where we think that DM pays some cognitive cost to deal with menu, and thus choice problem with larger menu provides DM with less information regarding her tastes (i.e. she becomes distracted because of the framing). The other example is *deadline*, when frame corresponds to the maximum amount of time that DM is able to spend on choice problem. Intuitively, the more time DM has, the more precise she becomes

[17]Dynamic random choice without knowledge of frames is a special case when analyst observes choice frequencies from no more than one menu in each frame.

[18]That is, she receives both signals corresponding to frames β_1 and β_4.

about her preferences[19]. BAUP model allows us to formalize this concept by assuming that analyst knows an "informativeness" relation on the set of frames

$$\succeq_{\text{info}} \ \subseteq \ \mathcal{B} \times \mathcal{B} \tag{57}$$

and BAUP model should reflect this by framing function which maps more informative frame to more informative with respect to Blackwell order of experiments:

$$\beta \ \succeq_{\text{info}} \ \beta' \qquad \Rightarrow \qquad h(A, \beta) \ \succeq_{\text{Blackwell}} \ h(A', \beta') \tag{58}$$

We require $\succeq_{\text{info}}$ to be transitive because of its interpretation.

Example 3. Analyst knows the explicit structure of news. Suppose frames are given by a set of observed conditions ("advertisements") under which DM makes her choice. Thus analyst is able to observe DM's choice in the presence of condition a, in the presence of condition b and in the presence of conditions a and b together, etc. Another example is when we associate each alternative in menu with a component of framing, which provides some information about the subjective state to DM. In these examples we can say that there are some "basic blocks" or "news", and each frame consists of some set of basic news. We can define these blocks to be "basic" Blackwell experiments in BAUP formalism in the sense that framing of a choice problem corresponds to the corresponding compound Blackwell experiment.

It turns out that example 3 covers examples 1 and 2 as special cases (see Theorem 3). Therefore we focus on developing a formalism to model the explicit structure of news, and on proving important technical statements for this setup. Then we show how to apply these results to dynamic random choice and "informativeness relation" setups. We start with a number of definitions.

An *alphabet* is a finite set $\mathcal{K} = \{1, ..., K\}$. A *letter* is a number $k \in \mathcal{K}$. A *word* $\mathcal{L}$ is a subset of an alphabet $\mathcal{L} \subseteq \mathcal{K}$. A *transcription function* is a mapping $t : \ \mathcal{B} \to 2^{\mathcal{K}}$ from the set of frames to the set of words. Given a finite family of Blackwell experiments $\{\sigma_q\}_{q \in Q}$ on S with finite signal spaces and corresponding random signals τ_q, for any non-empty subset $Q' \subseteq Q$ we define $\sigma_{Q'} \equiv \bigcup\limits_{q \in Q'} \sigma_q$ to be a compound Blackwell experiment with the corresponding vector of random signals $(\tau_q)_{q \in Q'}$; in other words, agent receives all signals τ_q with $q \in Q'$ from the compound experiment $\sigma_{Q'}$. We define $\sigma_{\{\varnothing\}}$ to be a trivial Blackwell experiment[20]. Note also that $\sigma_{\{q\}} = \sigma_q$. Now we are ready to define the notion of information structure.

Definition. BAUP model with information structure. Given a set of frames $\mathcal{B}$, we say that strongly frame consistent BAUP model $\{S, u, \pi, h\}$ exhibits information structure $(\mathcal{K}, t)$ given by alphabet $\mathcal{K} = \{1, ..., K\}$ and transcription function $t : \ \mathcal{B} \to 2^{\mathcal{K}}$ if there exists a family $\{\sigma_k\}_{k \in \mathcal{K}}$ of Blackwell experiments on S with finite signal space (let's call them *basic* experiments), indexed by letters of the alphabet, such that $h_\beta = \sigma_{t(\beta)}$. In other words, Blackwell experiment associated with frame β should be a compound

[19] Natenzon (2012) considers a variant of this model under some parametric assumptions.

[20] That is, a Blackwell experiment with signal space of size 1, providing uninformative signals.

experiment consisting of a subset of basic experiments σ_k prescribed by a transcription function t.

Note that BAUP model can exhibit multiple information structures for different families of Blackwell experiments σ_k.

Definition. Fine information structure. Given a set of frames $\mathcal{B}$, we say that information structure $(\mathcal{K}, t)$ is fine if transcription function t is a one-to-one mapping from $\mathcal{B}$ to $2^{\mathcal{K}} \backslash \{\varnothing\}$. Note that this requires $2^K - 1 \geq \mid \mathcal{B} \mid$ and $\varnothing \notin \mathrm{Im}_t(\mathcal{B})$[21].

Let D be the set of total orders on X. Given a set of frames $\mathcal{B}$, we define a vector of total orders $\mathbf{d} \in \mathbf{D} \equiv D^{|\mathcal{B}|}$ to be an element of the Cartesian product of $\mid \mathcal{B} \mid$ sets of total orders on X indexed by $\beta \in \mathcal{B}$. d_β is then the total order corresponding to β-coordinate of $\mathbf{d}$ in $\mathbf{D}$. Given a set of frames $\mathcal{B}$ and strongly frame consistent BAUP model $\{S, u, \pi, h\}$, we define a vector of signals $\boldsymbol{\tau} = (\tau^1, ..., \tau^\beta, ..., \tau^{|\mathcal{B}|})$ to be a random variable equal to the vector with components being random signals τ^β associated with Blackwell experiments h_β.

Let $\mu(\mathbf{d})$ be a distribution of vectors of total orders on X. We say that $\mu(\mathbf{d})$ is induced by a strongly frame consistent regular BAUP model $\{S, u, \pi, h\}$ if

$$\mu(\mathbf{d}) \quad = \quad \Pr\{\ \boldsymbol{\tau} \ \mid \ x d_\beta y \ \Leftrightarrow \ \mathbb{E}[u_s(x) \mid \tau_\beta] \geq \mathbb{E}[u_s(y) \mid \tau_\beta] \quad \forall x, y \in X, \ \beta \in \mathcal{B}\ \} \tag{59}$$

The following lemma is a key result, which lies in the heart of "everything goes" conclusion about the representation of random choice with a BAUP model with particular information structure.

Lemma 1. Let the finite set of frames $\mathcal{B}$ be given. Then for any fine information structure $(\mathcal{K}, t)$, for any distribution $\mu(\mathbf{d})$ of vectors of total orders on X with full support there exists a strongly frame consistent regular BAUP model $\{S, u, \pi, h\}$ that exhibits information structure $(\mathcal{K}, t)$ and induces distribution of vectors of random orders $\mu(\mathbf{d})$. Moreover, $\mu(\mathbf{d})$ could be induced by a population of the described above BAUP decision makers such that basic Blackwell experiments associated with letters of alphabet $k \in \mathcal{K}$ have two possible signal realizations $\tau^k \in \{0, 1\}$.

Proof. It is enough to show that lemma is true when $\mid \mathcal{B} \mid = 2^K - 1$. If $\mid \mathcal{B} \mid < 2^K - 1$, we can augment $\mathcal{B}$ with fake frames and enlarge vector of total orders preserving full support condition in order to use the result from the previous case. Therefore WLOG $\mid \mathcal{B} \mid = 2^K - 1$, and transcription function t is a bijection from $\mathcal{B}$ to $2^K \backslash \{\varnothing\}$, since information structure $(\mathcal{K}, t)$ is fine.

We proceed in three steps. Firstly, we show how to get any vector of expected utility

[21] This means that each frame should correspond to some informative Blackwell experiment. If we allow for the trivial Blackwell experiment, Theorem 3 continuous to work in the population interpretation. We can say that non-trivial Blackwell experiments allow us to provide DM information about her "population type".

functions $(V_{\mathcal{L}})_{\mathcal{L}\in 2^X\backslash\{\varnothing\}}$ at some (subjective) state of the world. Secondly, we show that this allows us to construct a strongly frame consistent BAUP model exhibiting information structure $(\mathcal{K}, t)$ that induces some distribution of vectors of total orders $\tilde{\mu}(\mathbf{d})$ with full support. Thirdly, we rescale probabilities and utilities such that $\tilde{\mu}(\mathbf{d})$ becomes $\mu(\mathbf{d})$.

Step 1. Consider state space

$$S \quad = \quad \{s^{\theta}_{m}\}^{\theta\in\Theta}_{m\in\{0,1\}^K} \quad \equiv \quad \{s^{\theta}_{m_1\ldots m_K}\}^{\theta\in\Theta}_{(m_1,\ldots,m_K)\in\{0,1\}^K} \tag{60}$$

indexed by the "news realizations" $m_k \in \{0,1\},\ k = 1, ..., K$ and "type" of DM $\theta \in \Theta$, where Θ is a finite set with cardinality $\mid \Theta \mid = \mid D \mid^{2^K-1}$. As usual, define $\theta(s)$ to be coordinate of s corresponding to type θ, and $m_k(s)$ to be coordinate of s corresponding to the k-th "news realization":

$$s \quad \equiv \quad s^{\theta(s)}_{m(s)} \quad \equiv \quad s^{\theta(s)}_{m_1(s)\ldots m_K(s)} \tag{61}$$

Let prior distribution be uniform: $\pi = \pi_{\text{uniform}}$ with $\pi_{\text{uniform}}(s) = \mid S \mid^{-1} \ \ \forall s$. Consider Blackwell experiments $\overline{\sigma}_0, \overline{\sigma}_1, ..., \overline{\sigma}_K$ such that Θ is a space of signal realizations for $\overline{\sigma}_0$, and $\{0,1\}$ is a space of signal realizations for each $\overline{\sigma}_k$ with $k = 1, .., K$, and

$$\overline{\sigma}_0(s, \overline{\tau}^0 = \theta') \quad = \quad \mathbb{1}\{\theta' = \theta(s)\} \quad \forall \theta' \in \Theta \tag{62}$$

$$\overline{\sigma}_k(s, \overline{\tau}^k = i) \quad = \quad \mathbb{1}\{i = m_k(s)\} \quad \forall i \in \{0,1\} \qquad \text{for } k = 1, ..., K$$

Observe that signal realization is completely determined by state s, thus we may write $\overline{\tau}^0(s)$ and $\overline{\tau}^k(s)$. Therefore these Blackwell experiments induce partitional information structures on S, and we can recover the joint distribution of the corresponding signals from the given above marginal distributions. The partitions of state space S induced by experiments $\overline{\sigma}_0$ and $\overline{\sigma}_k$ are given below:

$$P_{\overline{\sigma}_0}(s) \quad \equiv \quad \{s' \in S \mid \overline{\tau}^0(s') = \overline{\tau}^0(s)\} \quad = \quad \{s' \in S \mid \theta(s') = \theta(s)\} \tag{63}$$

$$P_{\overline{\sigma}_k}(s) \quad \equiv \quad \{s' \in S \mid \overline{\tau}^k(s') = \overline{\tau}^k(s)\} \quad = \quad \{s' \in S \mid m_k(s') = m_k(s)\}$$

Thus the compound experiments constructed from $\overline{\sigma}_0, \overline{\sigma}_1, ..., \overline{\sigma}_K$ are well-defined. Consider compound experiments

$$\sigma_k \ = \ \overline{\sigma}_0 \cup \overline{\sigma}_k \tag{64}$$

Their conditional distributions are given by

$$\sigma_k(s, (\theta', i)) \ = \ \mathbb{1}\{\ \theta' = \theta(s)\ ,\ i = m_k(s)\ \} \tag{65}$$

and the corresponding partitions are given by

$$P_{\sigma_k}(s) \quad \equiv \quad \{s' \in S \mid (\overline{\tau}^0, \overline{\tau}^k)(s') = (\overline{\tau}^0, \overline{\tau}^k)(s)\} \quad = \tag{66}$$

$$= \quad \{s' \in S \mid \theta(s') = \theta(s)\ ,\ m_k(s) = m_k(s')\}$$

Finally, for any non-empty word $\mathcal{L} \in 2^{\mathcal{K}}\backslash\{\varnothing\}$ the corresponding compound Blackwell experiment

$$\sigma_{\mathcal{L}} \quad \equiv \quad \bigcup_{k\in\mathcal{L}} \sigma_k \tag{67}$$

has the following conditional distribution and associated partition of the state space:

$$\sigma_{\mathcal{L}}(s,(\theta',\mathbf{i})) \;=\; \mathbb{1}\{\;\theta'=\theta(s)\;,\;i_k=m_k(s)\;\forall k\in\mathcal{L}\} \tag{68}$$

$$P_{\mathcal{L}}(s) \;=\; \{s'\in S\mid \theta(s')=\theta(s)\;,\;m_k(s')=m_k(s)\;\forall k\in\mathcal{L}\}$$

We define framing function h according to the information structure $(\mathcal{K},t)$ applied to the basic experiments σ_k, $k\in\mathcal{K}$. That is, $h_\beta=\sigma_{t(\beta)}$ and we require strong frame consistency for $h(i)$ as a framing function in the BAUP model.

Let $\{V^\theta_{\mathcal{L}}(x)\}_{\theta\in\Theta,\;\mathcal{L}\in 2^{\mathcal{K}}\setminus\{\varnothing\}}$ be a family of utility functions. Consider states s^θ_ι where $\iota=(1,1,...,1)$ is a vector of ones. Denote also $\mathcal{L}(m)=\{k\in\mathcal{K}\mid m_k=1\}$.
Claim: the following state-dependent utility function

$$\hat{u}_s(x) \;=\equiv \sum_{\mathcal{J}:\;\mathcal{L}(m(s))\subseteq\mathcal{J}\subseteq\mathcal{K}} (-1)^{|\mathcal{J}|-|\mathcal{L}(m(s))|}\cdot 2^{K-|\mathcal{J}|}\cdot V^\theta_{\mathcal{J}}(x) \tag{69}$$

is such that

$$\mathbb{E}_{\pi_{\text{uniform}}}[\hat{u}_{s'}(x)\mid s'\in P_{\mathcal{L}}(s^\theta_\iota)] \;=\; V^\theta_{\mathcal{L}}(x) \qquad \forall\;\mathcal{L}\in 2^{\mathcal{K}}\setminus\{\varnothing\},\;\theta\in\Theta \tag{70}$$

Indeed,

$$\mathbb{E}_{\pi_{\text{uniform}}}[\hat{u}_{s'}(x)\mid s'\in P_{\mathcal{L}}(s^\theta_\iota)] \;=\; \frac{\sum\limits_{\theta'}\sum\limits_{m}\mathbb{1}\{\theta'=\theta\}\prod\limits_{k\in\mathcal{L}}\mathbb{1}\{m_k=1\}\cdot u_{s^{\theta'}_m}(x)}{\sum\limits_{\theta'}\sum\limits_{m}\mathbb{1}\{\theta'=\theta\}\prod\limits_{k\in\mathcal{L}}\mathbb{1}\{m_k=1\}} = \tag{71}$$

$$=\;\frac{1}{2^{K-|\mathcal{L}|}}\cdot\sum_{m:\;\mathcal{L}\subseteq\mathcal{L}(m)\subseteq\mathcal{K}} u_{s^\theta_m}(x) \;=$$

$$=\;\frac{1}{2^{K-|\mathcal{L}|}}\cdot\sum_{m:\;\mathcal{L}\subseteq\mathcal{L}(m)\subseteq\mathcal{K}}\left(\sum_{\mathcal{J}:\;\mathcal{L}(m)\subseteq\mathcal{J}\subseteq\mathcal{K}}(-1)^{|\mathcal{J}|-|\mathcal{L}(m)|}\cdot 2^{K-|\mathcal{J}|}\cdot V^\theta_{\mathcal{J}}(x)\right) \;=$$

$$=V^\theta_{\mathcal{L}}(x)+\frac{1}{2^{K-|\mathcal{L}|}}\cdot\sum_{\mathcal{J}:\;\mathcal{L}\subsetneq\mathcal{J}\subset\mathcal{K}}(-1)^{|\mathcal{J}|}\cdot 2^{K-|\mathcal{J}|}\cdot V^\theta_{\mathcal{J}}(x)\cdot\underbrace{\left(\sum_{m:\;\mathcal{L}\subseteq\mathcal{L}(m)\subseteq\mathcal{J}}(-1)^{-|\mathcal{L}(m)|}\right)}_{=0\text{ for }\mathcal{J}\neq\mathcal{L}} = V^\theta_{\mathcal{L}}(x)$$

Step 2. By virtue of our choice of Θ we have $\mid\Theta\mid=\mid D\mid^{2^K-1}=\mid\mathbf{D}\mid$. Therefore there is a bijection between the space of vectors of total orders $\mathbf{D}$ and Θ. We can identify each $\theta\in\Theta$ with some vector of total orders $\mathbf{d}\in\mathbf{D}$ and vice versa; abusing notations, let's say that $\theta=\mathbf{d}$. Consider the following family of utility functions

$$V^\theta_{\mathcal{L}}(x)=\#\{y\in X\mid x\theta_{t^{-1}(\mathcal{L})}y\} \qquad \theta\in\Theta,\;\mathcal{L}\in 2^{\mathcal{K}}\setminus\{\varnothing\} \tag{72}$$

where $\theta_{t^{-1}(\mathcal{L})}$ is total order equal to the coordinate of vector $\mathbf{d}$ corresponding to the frame β such that $t(\beta)=\mathcal{L}$ (recall that transcription function t is a bjiection under our assumptions and thus the inverse transcription function t^{-1} is well-defined for all $\mathcal{L}\in 2^{\mathcal{K}}\setminus\{\varnothing\}$).

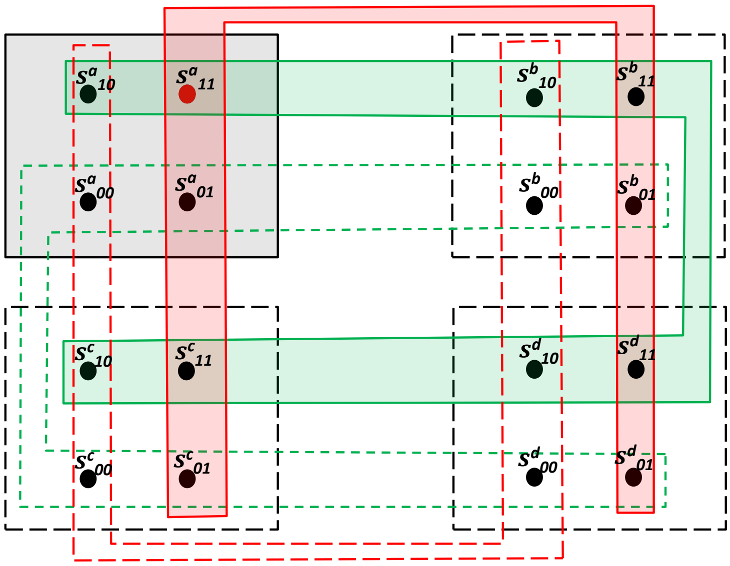

Figure 2: State space used in the proof of Lemma 1, an example with $\Theta = \{a, b, c, d\}$ and $K = 2$. Black lines corresponds to the partition induced by $\overline{\sigma}_0$ (i.e. "population type"), green lines - to the partition induced by $\overline{\sigma}_1$, and red lines - by $\overline{\sigma}_2$. Realized state is s^a_{11}, and solid lines with filling correspond to the realized elements of partitions $P_{\overline{\sigma}_0}(s^a_{11})$, $P_{\overline{\sigma}_1}(s^a_{11})$ and $P_{\overline{\sigma}_2}(s^a_{11})$.

Observe that utility function $V^\theta_{\mathcal{L}}$ represents total order $\theta_{t^{-1}(\mathcal{L})}$.

Now use step 1 to come up with a state-dependent utility function $\hat{u}_s(x)$ such that expected utility of DM at state s^θ_ι under frame β is $V^\theta_{t(\beta)}$. Pick some transcendental number $\epsilon \in (0, 1)$ (say, $\epsilon = e^{-1}$), enumerate arbitrary $x \in X$ by $x^1, ..., x^n$ and consider the following state-dependent utility function:

$$\tilde{u}_s(x^q) \;=\; \hat{u}_s(x^q) \;+\; \epsilon^q \tag{73}$$

where ϵ^q is just ϵ to the power q. Because of the partitional informational structure, expected utility of DM at frame β is completely determined by the state s. Observe that $\mathbb{E}[\tilde{u}_{s'}(x^q) \mid s \in P_{t(\beta)}(s)] = a(s, \beta, x^q) + \epsilon^q$ where $a(s, \beta, x^q)$ is some rational number. Therefore ties occur with zero probability, since $a + \epsilon^q = a' + \epsilon^{q'}$ for some rational numbers a, a' and natural numbers $q \neq q'$ contradicts ϵ being transcendental. Hence BAUP model $\{S, \tilde{u}, \pi_{\text{uniform}}, h\}$ is regular. It is also strongly frame consistent by construction. Next, we have:

$$E_{\pi_{\text{uniform}}}[\tilde{u}_{s'}(x^q) \mid s' \in P_{t(\beta)}(s^\theta_\iota)] \;=\; \#\{y \in X \mid x^q \theta_\beta y\} + \epsilon^q \tag{74}$$

Therefore, for $x^q \neq x^r$ we have

$$x^q \theta_\beta x^r \;\Rightarrow\; E_{\pi_{\text{uniform}}}[\tilde{u}_{s'}(x^q) \mid s' \in P_{t(\beta)}(s^\theta_\iota)] \;=\; \#\{y \in X \mid x^q \theta_\beta y\} + \epsilon^q \;\geq \tag{75}$$

$$\geq\; \#\{y \in X \mid x^r \theta_\beta y\} + 1 + \epsilon^q \;>\; \#\{y \in X \mid x^r \theta_\beta y\} + \epsilon^r \;=$$

$$=\; E_{\pi_{\text{uniform}}}[\tilde{u}_{s'}(x^r) \mid s' \in P_{t(\beta)}(s^\theta_\iota)]$$

which means that at state s_ι^θ DM's expected utility at frame β represents total order θ_β. We conclude that

$$\tilde{\mu}(\mathbf{d}) \;\equiv\; \Pr\{\, \boldsymbol{\tau} \;\mid\; x d_\beta y \;\Leftrightarrow\; \mathbb{E}[u_s(x) \mid \tau_\beta] \geq \mathbb{E}[u_s(y) \mid \tau_\beta] \quad \forall x,y \in X,\; \beta \in \mathcal{B} \,\} \;\geq \tag{76}$$

$$\geq \;\; \Pr\{s_\iota^{\mathbf{d}}\} \;=\, \mid S \mid^{-1} \,>\; 0 \qquad\quad \forall \mathbf{d} \in \mathbf{D}$$

and hence strongly frame consistent regular BAUP model $\{S, \tilde{u}, \pi_{\text{uniform}}, h\}$ represents distribution over vectors of total orders $\tilde{\mu}(\mathbf{d})$ with full support.

Step 3. The last step is the easiest one to prove. We use flexibility of BAUP model allowing to rescale utilities and probabilities, preserving the values of expected utilities. Define

$$\mathbf{d}(s): \qquad \mathbb{E}[\tilde{u}_{s'}(x) \mid s' \in P_{t(\beta)}(s)] \geq \mathbb{E}[\tilde{u}_{s'}(y) \mid s' \in P_{t(\beta)}(s)] \quad \Leftrightarrow \quad x d_\beta(s) y \;\; \forall \beta \in \mathcal{B} \tag{77}$$

Next, $\forall \mathbf{d} \in \mathbf{D}$ define

$$S(\mathbf{d}) \;\;=\;\; \{s \in S \mid \mathbf{d}(s) = \mathbf{d}\} \;\; \overset{\text{by step 2}}{\neq} \;\; \varnothing \tag{78}$$

Consider prior probability distribution

$$\pi(s) \;\;=\;\; \frac{\mu(\mathbf{d})}{\mid S(\mathbf{d}) \mid} \;\;>\;\; 0 \tag{79}$$

since $\mu(\mathbf{d})$ has full support and $S(\mathbf{d}) \neq \varnothing$ by step 2. Consider state-dependent utility function

$$u_s(x) \;\;=\;\; \frac{\pi_{\text{uniform}}(s)}{\pi(s)} \tilde{u}_s(x) \tag{80}$$

Then it is easy to check that BAUP model $\{S, u, \pi, h\}$ is regular, strongly frame consistent and it induces distribution of vectors of total orders $\mathbf{d}$. The key observation is that since $\pi_{\text{utinform}}(s)\tilde{u}_s(x) = \pi_u(s) u_s(x)$, we have

$$\mathbb{E}_\pi[u_{s'}(x) \mid s' \in P_{t(\beta)}(s)] \;=\; \mathbb{E}_{\pi_{\text{uniform}}}[\tilde{u}_{s'}(x) \mid s' \in P_{t(\beta)}(s)] \quad \forall \beta \in \mathcal{B}\;, x \in X \tag{81}$$

and therefore $\mathbf{d}(s)$ and $S(\mathbf{d})$ remains the same in the BAUP model $\{S, u, \pi, h\}$ as in the model $\{S, \tilde{u}, \pi_{\text{uniform}}, h\}$. Next, for the distribution of vectors of total orders $\mu_{\{S,u,\pi,f\}}(\mathbf{d})$, induced by model $\{S, u, \pi, h\}$, we have:

$$\mu_{\{S,u,\pi,f\}}(\mathbf{d}) = \pi(S(\mathbf{d})) = \sum_{s \in S(\mathbf{d})} \frac{\mu(\mathbf{d})}{\mid S(\mathbf{d}) \mid} = \mu(\mathbf{d}) \tag{82}$$

This concludes calculations of step 3.

Finally, we get the desired population interpretation of random choice with basic experiments $\overline{\sigma}_k$ $k = 1, .., K$ having binary signal realizations, and population types θ. Lemma 1 is proven.

With the help of lemma 1 we are able to formulate and prove the main "everything goes" result of this section. It turns out that information structure $(\mathcal{K}, t)$ imposed on the top of BAUP model does not provide any discipline on the induced random choice. One of the consequences of this fact is that there is no hope to infer any knowledge about the information structure of frames from the point of view of BAUP decision maker by observing the (extended) random choice. Formal statement is given in Theorem 3 below.

Theorem 3. Weakly frame consistent regular BAUP model with any fine information structure $(\mathcal{K}, t)$ is almost indistinguishable with a model of arbitrary consistent framing effects and, therefore, is almost indistinguishable with any other weakly frame consistent regular BAUP model with different fine information structure $(\mathcal{K}', t')$. Moreover, if BAUP model is also strongly frame consistent, this result continues to hold for the case when analyst observes joint distribution of choice probabilities[22].

- **Corollary 1.** Regular BAUP model is almost indistinguishable with a model of arbitrary framing effects in the dynamic random choice setup, when analyst observes sequence of choices of individual decision makers. Strongly frame consistent regular BAUP model is almost indistinguishable with a model of arbitrary consistent framing effects in case of the extended dynamic random choice.
- **Corollary 2.** Weakly frame consistent regular BAUP model is almost indistinguishable with a model of arbitrary framing effects even if analyst additionally knows a transitive informativeness relation defined on the set of frames.

Proof. We'll show here the second statement (about joint distributions). The statement about marginal distributions could be proven using the same logic.

The theorem follows almost immediately from Lemma 1. Let's work in the space of joint probability distributions. We are going to show that the set of random choices induced by the described above BAUP model contains an interior of the space of joint probability distributions corresponding to arbitrary consistent framing effects (let's call it $\mathcal{R}$). If we prove that, then the closure of the former set is equal to the latter (recall that by Proposition 3 the latter set contains the former), and we have the desired result. Pick some point in the interior of the space of joint probability distributions corresponding to arbitrary consistent framing effects ρ. Then it has some representation via random of vector orders $\mu(\mathbf{d})$. Now let $\mu_{uniform}(.)$ be a uniform distribution on the space $\mathbf{D}$ and $\rho_{uniform}$ be the corresponding random choice. Consider ρ' given by

$$\rho' = \frac{1}{1-\epsilon}\rho - \frac{\epsilon}{1-\epsilon}\rho_{uniform} \tag{83}$$

Since ρ is an interior point of $\mathcal{R}$, then we can choose ϵ small enough such that $\rho' \in \mathcal{R}$. Then it has some representation by distribution $\mu'(.)$ over $\mathbf{D}$. Finally,

$$\rho = (1-\epsilon)\rho + \epsilon\rho_{uniform} \tag{84}$$

[22] With the notion of almost indistinguishability extended naturally by the help of Euclidean metrics.

Therefore, since ρ is linear in the probability measure over vectors of total orders that induces ρ, it has a representation

$$\mu''(\mathbf{d}) = (1-\epsilon)\mu'(\mathbf{d}) + \epsilon\mu_{uniform}(\mathbf{d}) \tag{85}$$

Observe that μ'' has full support and therefore we can apply Lemma 1 to conclude the proof.

Proof of Corollary 1.
Denote time by $\tau = 1, 2, 3, ..T$. We assume time to have finitely menu periods in order to avoid technical complications. Divide DMs by types $\theta \in \Theta$ such that each type has a unique ordered sequence of choice problems indexed by frame β and time τ. Let

$$\gamma_\tau^\theta = \{\beta \mid \text{DM}^\theta \text{ faced frame } \beta \text{ at time } \tau' \leq \tau\} \tag{86}$$

Note that it is possible to have $\gamma_\tau^\theta = \gamma_{\tau'}^{\theta'}$; note also that we have finitely many γ_τ^θ for different τ and θ. Therefore, we can construct a new set of frames

$$\Gamma = \{\gamma \mid \gamma = \gamma_\tau^\theta \text{ for some } \tau, \theta\} \tag{87}$$

Now, abusing notations, pick $\mathcal{K} = \mathcal{B}$ and define transcription function $t: \Gamma \to 2^{\mathcal{K}} \backslash \{\varnothing\}$ such that

$$t(\gamma) = \gamma \tag{88}$$

Then $(\mathcal{K}, t)$ is a fine information structure and we can apply Theorem 3 to get Corollary 1.

Proof of Corollary 2.
We have a transitive relation $\succeq_{\text{info}} \subseteq \mathcal{B} \times \mathcal{B}$. Let's construct the following information structure $(\mathcal{K}, t)$. Firstly, pick $K = \mid \mathcal{B} \mid$, enumerate $\beta \in \mathcal{B}$ by $\beta_1, ..., \beta_K$ and consider bijection $j: \mathcal{B} \to \mathcal{K}$ given by $j(\beta_i) = i$. Next, denote

$$W(\beta) = \{\beta' \in \mathcal{B} \mid \beta \succeq_{\text{info}} \beta'\} \tag{89}$$

Let the transcription function be

$$t(\beta) \;=\; j(W(\beta)) \;\equiv\; \{i \in \mathcal{K} \mid \exists \beta' \in W(\beta):\ i = j(\beta')\} \tag{90}$$

Then, using transitivity of $\succeq_{\text{info}}$, we get

$$\beta \succeq_{\text{info}} \alpha \;\Rightarrow\; W(\alpha) \subseteq W(\beta) \;\Rightarrow\; j(W(\alpha)) \subseteq j(W(\beta)) \;\Rightarrow \tag{91}$$

$$\Rightarrow\; h_\beta \;=\; \sigma_{W(\beta)} \succeq_{\text{Blackwell}} \sigma_{W(\alpha)} \;=\; h_\alpha$$

Application of Theorem 3 then concludes the proof.

3.2 Representation with state space of limited size

Theorem 2 says that if analyst does not observe frames, then BAUP model allows for (almost) arbitrary framing effects even if DM's uncertainty about her tastes is encoded by the state space of size 2. It is no longer true when analyst observes the extended random choice. Moreover, frame consistent[23] BAUP model with limited state space of any size is falsifiable in finite data samples (i.e. it is not almost indistinguishable with the model of arbitrary consistent framing effects) if the number of alternatives $\mid X \mid$ is big enough. This section contains some explicit examples of necessary conditions for the extended random choice to be representable by a frame-consistent BAUP model with limited state space. We also give one general result: a low bound on the size of subjective state space of DM necessary for the frame consistent BAUP model to induce arbitrary consistent framing effects.

Importantly, in this section we consider an extended random choice of a population of decision makers, rather then that of an individual DM. Previously we had no constrains on the size of state space, and this allowed us to entwine implicitly a "population" type of DM via some signal that told DM in what part of state space she is. Therefore we always could interpret models that we considered before as models of population random choice as well and vice versa. When we put constrains on the size of the state space, we should choose one of the interpretations explicitly, and we consider population random choice here, using eq. (36) for the definition of representation.

Example 4. Suppose that analyst observes an extended random choice $\rho^A_\beta(.)$; we write $\rho^{xy}_\beta(.)$ instead of $\rho^{\{x,y\}}_\beta(.)$ and similarly for other ρ-s for the ease of notations. Let M^2 be a frame consistent BAUP model with binary state space $S = \{s_1, s_2\}$. Below is a comparison of some of the necessary conditions on the extended random choice given by different models. Denote

$$W_{\alpha\beta\gamma}(x,y,z) \equiv \rho^{xy}_\alpha(x) - \rho^{xyz}_\alpha(x) + \rho^{yz}_\beta(y) - \rho^{xyz}_\beta(y) + \rho^{xz}_\gamma(z) - \rho^{xyz}_\gamma(z) \tag{92}$$

Then $\forall \alpha, \beta, \gamma \in \mathcal{B}$ and distinct $x, y, z \in X$ ranges of $W_{\alpha\beta\gamma}(x,y,z)$ that could be induced by the corresponding models are given below:

$$\begin{array}{lll} \text{(i)} & \text{Random utility model :} & 0 \leq W_{\alpha\beta\gamma}(x,y,z) \leq 1 \\ \text{(ii)} & \text{Frame consistent BAUP model } M^2 : & 0 \leq W_{\alpha\beta\gamma}(x,y,z) < 2 \\ \text{(iii)} & \text{Frame consistent BAUP model :} & 0 \leq W_{\alpha\beta\gamma}(x,y,z) < 3 \\ \text{(iv)} & \text{Arbitrary consistent framing effects :} & 0 \leq W_{\alpha\beta\gamma}(x,y,z) \leq 3 \\ \text{(v)} & \text{BAUP model :} & -3 < W_{\alpha\beta\gamma}(x,y,z) < 3 \\ \text{(vi)} & \text{Arbitrary framing effects :} & -3 \leq W_{\alpha\beta\gamma}(x,y,z) \leq 3 \end{array} \tag{93}$$

To see this, note that arbitrary framing effects model does not provide any restrictions on the choice probabilities, justifying (vi). Similarly general BAUP model (not necessarily consistent with frames) allows, according to Theorem 1, for everything except of $W_{\alpha\beta\gamma}(x,y,z) = -3$ and $W_{\alpha\beta\gamma}(x,y,z) = 3$, because these extreme cases violate Revealed Dominance Acyclicity Axiom. Next, when extended random choice is consistent with

[23] Since we do not consider joint distribution of random choices here, strong and weak versions of frame consistency are equivalent for our purposes and we do not distinguish them.

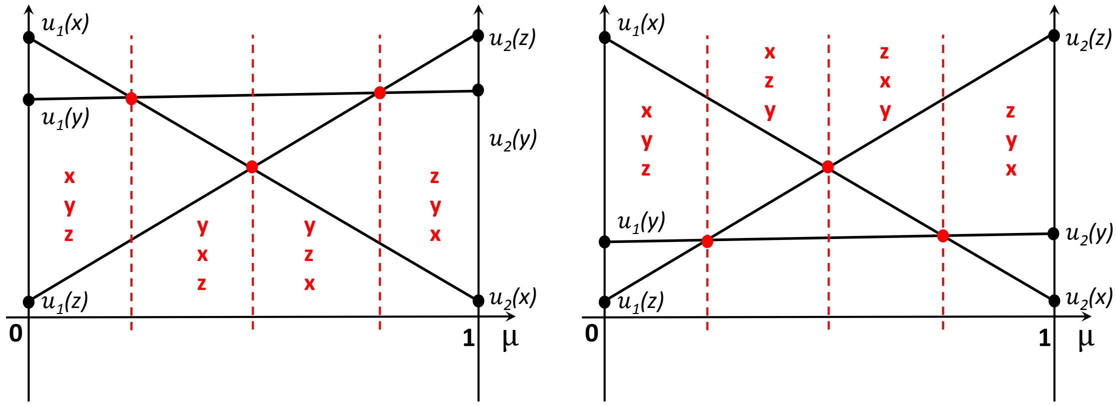

Figure 3: Expected utilities for decision makers of types $\psi_1 = \{xyz, yxz, yzx, zyx\}$ (left plot) and $\psi_2 = \{xyz, xzy, zxy, zyx\}$ (right plot) when state space is binary $S = \{s_1, s_2\}$. Horizontal Axis corresponds to the probability of state s_2 and vertical Axis corresponds to the expected utility of alternative.

frames, then

$$\begin{array}{lcll} p_\alpha(zxy) & \equiv & \rho_\alpha^{xy}(x) - \rho_\alpha^{xyz}(x) & \geq 0 \\ p_\beta(xyz) & \equiv & \rho_\beta^{yz}(y) - \rho_\beta^{xyz}(y) & \geq 0 \\ p_\gamma(yzx) & \equiv & \rho_\gamma^{xz}(z) - \rho_\gamma^{xyz}(z) & \geq 0 \end{array} \tag{94}$$

where $p_\alpha(zxy)$ stands for the probability of total order d such that $zdxdy$ in the random order representation of choices in frame α, and similar for other ρ-s. Thus $p_\alpha(zxy) \geq 0$ and similar inequalities stand for the Monotonicity property of random choice consistent with random utility model for frame α. Using this observation we get (i) and (iv) for free and (iii) with the help of Theorems 3 and Theorem 1. Now let's show how we can get (ii) and other similar inequalities for the BAUP models with limited state space.

Consider Fig. (3). The left plot shows the type of DM that was used in the proof of Theorem 2. Let μ be a probability that DM assigns to state s_2. At state s_1 ($\mu = 0$) and when μ is small DM orders alternatives accroding to order d_0: xd_0yd_0z; we make use of short notation $d_0 = xyz$. When μ increases, at some intermediate values of μ DM has orders yxz and yzx. Finally, when μ is close to one, DM orders alternative according to $d_1 = zyx$. The core feature that we used in the proof of Theorem 2 is that, depending on μ, each alternative could be the best among all, and also that at state 2 alternatives are ordered in reverse order in comparison to state 1. Observation of random choice from each menu in the frame allows analyst to recover not only the distribution of the best alternatives in a given menu, but also the distribution of total orders on the triple of alternatives x, y, z. Therefore decision maker of the type considered at the left plot of fig. 3 cannot justify arbitrary distribution over orders, she can only justify distributions with support equal to $\psi_1 = \{xyz, yxz, yzx, zyx\}$.

Similar considerations apply to DM whose expected utility is given at the right plot of Fig. (3); let's call her type $\psi_4 = \{xyz, xzy, zxy, zyx\}$. Using permutations of x, y, z we get types ψ_2, ψ_3 from ψ_1 and ψ_5, ψ_6 from ψ_2. Next, if $d_0 = xyz$ and $d_1 = xzy$, we get type $\psi_7 = \{xyz, xzy, zxy\}$ and ψ_8, ψ_9 as permutations of ψ_{10}. Starting from $d_0 = xyz$ and $d_1 = yxz$, we get $\psi_{10} = \{xyz, yxz\}$ and similarly ψ_{11}, ψ_{12}; starting from $d_0 = xyz$

and $d_1 = xzy$, we get $\psi_{13} = \{xyz, xzy\}$ and ψ_{14}, ψ_{15}; finally, $d_0 = d_1 = \{xyz\}$ gives us $\psi_{16} = \{xyz\}$ and similarly for $\psi_{17} - \psi_{21}$. This classification of population types $\psi \in \Theta$ is summarized below[24].

$$
\begin{array}{lll}
\psi_1 = \{xyz, yxz, yzx, zyx\} & \psi_2 = \{yzx, zyx, zxy, xzy\} & \psi_3 = \{zxy, xzy, xyz, yxz\} \\
\psi_4 = \{xyz, xyz, zyx, zyx\} & \psi_5 = \{yzx, yzx, xzy, xzy\} & \psi_6 = \{zxy, zxy, yxz, yxz\} \\
\psi_7 = \{xyz, xzy, zxy\} & \psi_8 = \{yzx, yxz, xyz\} & \psi_9 = \{zxy, zyx, yzx\} \\
\psi_{10} = \{xyz, yxz\} & \psi_{11} = \{yzx, zyx\} & \psi_{12} = \{zxy, xzy\} \\
\psi_{13} = \{xyz, xzy\} & \psi_{14} = \{yzx, yxz\} & \psi_{15} = \{zxy, zyx\} \\
\psi_{16} = \{xyz\} & \psi_{17} = \{yzx\} & \psi_{18} = \{zxy\} \\
\psi_{19} = \{xzy\} & \psi_{20} = \{yxz\} & \psi_{21} = \{zyx\}
\end{array}
\tag{95}
$$

Considerations analogous to those used in Theorem 2 allow us to argue that DM of each type could exhibit arbitrary non-degenerate distribution of orders prescribed by her type at each frame. Denote $p^{\psi}_{\beta}(d)$ to be the probability that type ψ exhibits total order d at frame β, and $W^{\psi}_{\alpha\beta\gamma}(x, y, z)$ to be a function W of extended random choice induced by type ψ. Observe that each type of DM has at most two of the three total orders xyz, yzx, zxy. Therefore we have

$$
p^{\psi}_{\alpha}(zxy) + p^{\psi}_{\beta}(xyz) + p^{\psi}_{\gamma}(yzx) \;\; < \;\; 2 \;\;\; \forall \;\; \psi \in \Theta \;\;\; \Rightarrow \tag{96}
$$

$$
W^{\psi}_{\alpha\beta\gamma}(x, y, z) \;\; < \;\; 2 \quad \forall \psi \in \Theta \quad \Rightarrow \quad W_{\alpha\beta\gamma}(x, y, z) \;\; = \;\; \sum_{\psi \in \Theta} \nu(\psi) \cdot W^{\psi}_{\alpha\beta\gamma}(x, y, z) \;\; < \;\; 2
$$

where we used linearity of p (and, therefore, W) in random choice probability ρ. This proves eq.(93) in our example.

Observe that each type ψ misses also one of the orders xzy, yxz, yzx, zxy, which gives us new necessary condition, independent of $W_{\alpha\beta\gamma}(x, y, z) < 2$. We can get a few other conditions by permutations of these two. Other sets of four total orders, which are partially missed by each type ψ, do not give us independent inequalities. We also can consider sets of 5 and 6 orders. Obviously, there is no type of DM who has any of these sets, therefore we can write the corresponding conditions as well; but the corresponding inequalities are not independent. Summarizing, M^2 representation requires the following necessary[25] conditions for any triple of alternatives x, y, z:

$$
\begin{array}{lll}
0 \;\le\; p_{\alpha}(zxy) + p_{\beta}(xyz) + p_{\gamma}(yzx) \;<\; 2 & & (\times 2) \\
0 \;\le\; p_{\alpha}(xzy) + p_{\beta}(yxz) + p_{\gamma}(yzx) + p_{\delta}(zxy) \;<\; 3 & & (\times 3)
\end{array}
\tag{97}
$$

where p are expressed via extended random choice according to eq.(94), and $(\times m)$ stands for the m conditions emerging as the result of permutations among x, y, z.

Conditions given by eq.(97) hold for any triple of alternatives x, y, z. Do we have more independent conditions for larger number of alternatives? The answer is: probably,

[24]Note that we do not care which state to call s_1 and which s_2, thus type $\{xyz, yxz, yzx, zyx\}$ is equivalent to type $\{zyx, yzx, yxz, xyz\}$

[25]If $X = \{x, y, z\}$, these conditions should be sufficient as well; we may try to show this with the help of Axiom of Revealed Stochastic Preference from McFadden (2005)

but before trying to find these conditions, let's investigate what kind of information about the distribution of orders for each frame analyst could possible get from the extended random choice data. In case of 3 alternatives analyst can recover the full distribution of orders in the random utility model describing choices inside a frame. It is no longer the case if the number of alternatives are larger. Falmagne (1978) shows that the best that analyst can do is to recover the distribution of the order for each alternative (i.e. probability that x is the best, probability that x is second best, etc., probability that x is the worst). Therefore $p_\beta(xyzw)$ has no longer meaning of the probability that order $xyzw$ appears in the random order representation of random choice, since there are many representations. However, we can define it to be the *minimum*[26] of this probability among all random order representations:

$$p(d) = \min\{\mu(d) \mid \mu(.) \text{ represents random choice } \rho\} \tag{98}$$

One can try to find an explicit formula for $p_\beta(d)$ and, most probably, this formula would be a linear polynomial in ρ; we leave it for the further work. For our purposes it is important that $p_\beta(d)$ could be close to one for some non-degenerate random choice probabilities. To see this, denote $\rho(d)$ to be a (degenerate) random choice induced by total order d. For example, $d = xyz$, then $\rho^{xyz}(x) = \rho^{xy}(x) = \rho^{xz}(x) = \rho^{yz}(y) = 1$ and the rest ρ-s are zeros.

Lemma 2. Suppose ρ admits random utility representation (i.e. $\rho \in RO$). Then

$$p(d) \quad \geq \quad 1- \mid D \mid \cdot ||\rho - \rho(d)|| \quad \forall d \in D \tag{99}$$

where $\mid D \mid = \mid X \mid !$ is the number of distinct total orders on X.

Proof. Consider any random order representation of ρ given by distribution $\mu(.)$. Any total order $d' \neq d$ is such that $\exists x, y \in X$: $x \neq y$, xdy but $yd'x$. Next, $\rho^{xy}(y) \leq ||\rho - \rho(d)||$ implies that $\mu(d') \leq ||\rho - \rho(d)||$. Summing this over all $d' \neq d$, we get

$$\mu(d) \quad = \quad 1 - \sum_{d' \neq d} \mu(d') \quad \geq \quad 1- \mid D \mid \cdot ||\rho - \rho(d)|| \tag{100}$$

Having definition of $p(d)$, we can find the following examples of necessary conditions for the M^2 representation, which include 4 alternatives[27]:

$$\begin{aligned} p_\alpha(xyzw) + p_\beta(wzxy) + p_\gamma(wyxz) &\leq 2 \\ p_\alpha(xyzw) + p_\beta(wzxy) + p_\gamma(zwyx) &\leq 2 \end{aligned} \tag{101}$$

The corresponding calculations are bulky and omitted in this paper.

Note that it is enough to have 3 frames to falsify BAUP model M^2. Now let's turn our attention to the general case. Denote BAUP model with size of state space $\mid S \mid = L$

[26] Clearly, this minimization problem has a solution.

[27] Note that we cannot get these inequalities straightforwardly by dropping one of the alternatives from all orders and considering the case with 3 alternatives. This gives us hope that they are independent of conditions (97).

by M^L and consider it. We might expect to have similar picture with types ψ_i of DM corresponding to the sets of total orders emerging as expected utilities with some specific state-dependent utility function. Formally, given state-dependent utility function $u_s(x)$ with $s \in S = \{1, ..., L\}$ and $x \in X$, let

$$\psi(u, L) = \{ \ d \in D \ \mid \ \exists \pi : \quad E_\pi[u_s(x)] \geq E_\pi[u_s(y)] \quad \Leftrightarrow \quad xdy \ \} \tag{102}$$

where π is a probability measure on S. Next, denote $\Psi(L)$ to be the set of all possible $\psi(u, L)$:

$$\Psi(L) = \{\psi \in 2^D \backslash \{\varnothing\} \ \mid \ \exists u : \ \psi = \psi(u, L)\} \tag{103}$$

The following lemma is an important technical result.

Lemma 3. Let the set of frames $\mathcal{B}$ be given. Consider frame-consistent BAUP model M^L with the size of state space $\mid S \mid = L$. If $D \notin \Psi(L)$ and $\mid \mathcal{B} \mid \geq \max\{\mid \psi \mid \ \mid \psi \in \Psi(L)\} + 1$, then M^L is distinguishable (meaning that it is not almost indistinguishable) with the model of arbitrary consistent framing effects.

Proof. Note that each DM of population type $\theta \in \Theta$ belongs to some type ψ; denote this type by $\psi(\theta)$ and define $\mu(\psi) = \mu(\theta : \ \psi(\theta) = \psi)$. Take any $C \subseteq D$ with $\mid C \mid = \max\{\mid \psi \mid \ \mid \psi \in \Psi(L)\} + 1$ (such C exists, since $D \notin \Psi(L)$). Note that $C \nsubseteq \psi$ $\forall \psi \in \Psi$. Let $\{\beta_d\}_{d \in C} \subseteq \mathcal{B}$ be a set of $\mid C \mid$ pairwise different frames (such set exists because $\mathcal{B}$ is large enough). Let $\epsilon > 0$ be small enough for our purposes and pick some $\rho \in int\,(\mathcal{P}_{\text{consistent}})$ such that $||\rho_{\beta_d} - \rho(d)|| < \epsilon \ \ \forall d \in C$. Suppose that M^L induces ρ. Consider some DM of type ψ. Since $C \nsubseteq \psi$, there is $d \in C$ such that $d \notin \psi$. Consider

$$R(\psi) = \text{conv}\left(\{\rho(d)\}_{d \in \psi}\right) \tag{104}$$

The set of random choices which could be induced by decision maker of type ψ at some frame is contained in $R(\psi)$. From the other hand, $d \notin R(\psi)$, since d is known to be one of the vertices of polygon RO of random choices induced by the random utility model, and $R(\psi)$ is a polygon equal to a convex closure of a subset of vertices of RO which does not include d. Thus,

$$\text{distance}\left(R(\psi), d\right) \equiv q_\psi > 0 \tag{105}$$

since $R(\psi)$ is closed (distance is a usual Euclidean distance in the space of choice probabilities for fixed frame). If $\epsilon < q_\psi$ then we should have $\mu(\psi) \cdot q_\psi \leq \epsilon$, because ρ_β is in the ϵ-neighborhood[28] of d, d is a vertex of convex polygon RO, and ρ_β is a convex combination of $\rho_\beta^{\psi'}$, $\psi' \in \Psi$ with $\rho_\beta^{\psi'} \in RO$. Pick

$$\epsilon = \frac{1}{2}\left(\sum_{\psi \in \Psi} \frac{1}{q_\psi}\right)^{-1} \tag{106}$$

[28]Observe that $||\rho - \rho(d)|| < \epsilon$ in the space of extended random choices implies $||\rho_\beta - \rho_\beta(d)||_\beta < \epsilon$ of $\rho(d)$ for the distance in the projection on any frame β

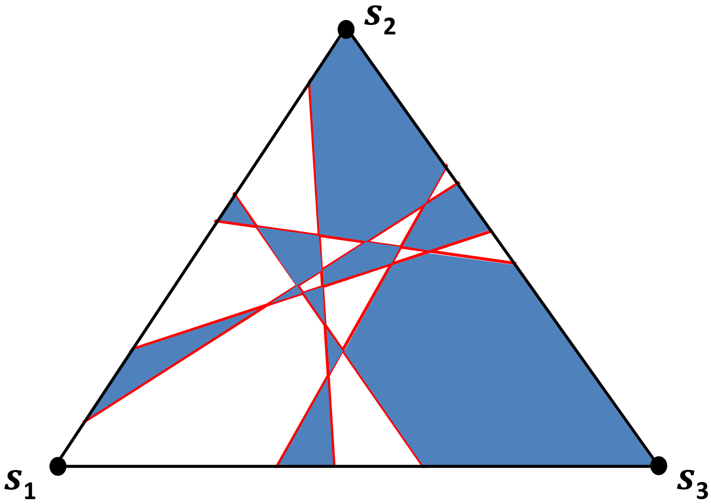

Figure 4: Maximal assessment of probability simplex by 6 lines (example for $S = \{s_1, s_2, s_3\}$ and $n = 4$). There are 22 different regions.

Then $\epsilon < q_\psi \ \forall \psi \in \Psi$ and

$$\sum_{\theta\in\Theta} \mu(\theta) \;=\; \sum_{\psi\in\Psi} \mu(\psi) \;\leq\; \sum_{\psi\in\Psi} \frac{\epsilon}{q_\psi} \;=\; \frac{1}{2}\left(\sum_{\psi\in\Psi}\frac{1}{q_\psi}\right)^{-1}\sum_{\psi}\frac{1}{q_\psi} \;=\; \frac{1}{2} \tag{107}$$

contradicting μ being a probability. Therefore, M^L does not induce $\rho \in int\,(\mathcal{P}_{\text{consistent}})$ and hence $cl(\mathcal{P}(M^L)) \neq \mathcal{P}^{\text{consistent}}$. Lemma 3 is proven.

Lemma 3 says us that it is enough to argue that $D \notin \Psi(L)$ and have enough frames in order to distinguish a frame consistent BAUP model M^L from the model of arbitrary consistent framing effects. The following geometrical result helps us to find situations in which $D \notin \Psi(L)$:

Fact A. Hyperplane arrangement. (Orlik and Terao (2013)) The maximal number of pieces by which r hyperplane of dimension $m-1$ divide R^m is given by the following formula (provided that $r \geq m$) :

$$Y^r_m \;=\; \sum_{k=0}^{m}\begin{pmatrix} r \\ k \end{pmatrix} \tag{108}$$

Theorem 4. Let $n =\mid X \mid$. For some set of frames $\mathcal{B}$ with $\mid \mathcal{B} \mid \geq 3$, consider frame consistent regular BAUP models M^l with state space of size $l = 1, 2, ...$ Then:
(i) Family of models M^l induces a family of nested sets of extended random choices such that

$$RO = \mathcal{P}(M^1) \subseteq cl(\mathcal{P}(M^2)) \subseteq ... \subseteq cl(\mathcal{P}(M^{\overline{L}})) = \mathcal{P}^{\text{consistent}} \tag{109}$$

for some $\overline{L}$ which we assume to be minimal subject to $M^{\overline{L}}$ being almost indistinguishable with a model of arbitrary consistent framing effects. M^1 is equivalent to the random utility model, and M^2 is strictly in between of random utility model and arbitrary consistent framing effects model.

(ii) $\overline{L} \quad \leq \quad | \mathcal{B} | \cdot n!^{|\mathcal{B}|}$

(iii) If $| \mathcal{B} | \geq \ n!$, then

$$\overline{L} \quad > \quad \max\{ \ L \ \mid \ Y_{L-1}^{\frac{n(n-1)}{2}} < n! \ \} \tag{110}$$

and asymptotically as $n \to \infty$

$$\overline{L} \geq \frac{n}{2} \tag{111}$$

Proof. (i) follows from our results for example 4 and simple considerations. (ii) follows from Theorem 3 and Lemma 1. Let's show how we can get (iii).

Consider a couple of alternatives $x, y \in X$. WLOG either $u_s(x) \geq u_s(y)$ for any state and expected utility of x is strictly greater than that of y, or $u_s(x) \geq u_s(y)$ for some states and vice versa for other states. In the latter case the probability simplex is divided by two parts on the region where expected utility of x is greater than that of y, and the region where we have the opposite. Note that these regions are divided by a hyperplane at which $\mathbb{E}_\mu[u_s(x)] = \mathbb{E}_\mu[u_s(y)]$. Similar considerations apply to each of $\frac{n(n-1)}{2}$ distinct pairs of alternatives. The corresponding $r \leq \frac{n(n-1)}{2}$ hyperplanes divide probability simplex on no more than $Y_{|S|-1}^{\frac{n(n-1)}{2}}$ pieces. Continuity considerations imply that expected utility constructed with the help of probability from a certain piece should induce the same total order for all probabilities in the interior of a piece. Therefore the total number of distinct orders of the corresponding type[29] ψ should be no more than $Y_{|S|-1}^{\frac{n(n-1)}{2}}$. We may apply then Lemma 3 to argue that M^L is distinguishable with arbitrary consistent framing effects model whenever $Y_{L-1}^{\frac{n(n-1)}{2}} < n!$, implying the first formula of (iii). For the asymptotic behavior[30], consider

$$\ln(Y_{L-1}^{\frac{n(n-1)}{2}}) \quad \sim \quad 2L \cdot \ln(n) \quad < \quad n \cdot \ln(n) \quad \sim \quad \ln(n!) \tag{112}$$

Theorem 4 is proven.

Theorem 4 says that if we have enough frames and alternatives, BAUP models M^L with state space of sizes L are falsifiable. Moreover, the number of alternatives which we need to distinguish M^L from arbitrary consistent framing effects model is not large (approximately, we need twice as many alternatives as states). However the proof of the theorem requires us to use a lot of frames. It is not clear whether it is necessary or there are better ways to falsify models M^L. Below is some statistics on the number of alterna-

[29]For the definition of types $\psi \in \Psi$ see eq. (102),(103).

[30]One might check that asymptotic formula works for $L << \frac{n(n-1)}{2}$ which is true for $L = n/2$ and large n

tives and frames which have shown to be sufficient for M^L to exhibit "moderate" framing effects.

#alternatives	falsifiable size of S	#frames used	for reference	
n	$\overline{L} > ...$	$\mid \mathcal{B} \mid$	$Y_{L-1}^{\frac{n(n-1)}{2}}$	$n!$
3	2	3	4	6
4	3	23	22	24
5	3	23	26	120
6	4	587	586	720
7	4	587	1561	5040
8	5	24159	24158	40320
...	...	...	...	...
$n >> 1$	$n/2$	$Y+1$	Y	$n!$

As we can see, current theoretical constructions show us how to deal, realistically, with state space of size 2 and, somewhat ambitious, with state space of size 3.

4 Conclusion

This paper considers preference uncertainty as a source of framing effects observed in the random choice. The central framework is a *BAUP* model which assumes that stochastic choice is a result of decisions of a population of agents. Each agent has a subjective state space, which we interpret as space of "states of mind", and state-dependent utility function encoding variation of her preferences. The information part consists of a prior probability distribution and Blackwell experiments associated with framings of choice problems. Agents are expected utility maximizers. It is assumed that frames allow agents to learn only their own tastes and nothing else. There are two types of data: usual random choice where framing environment could be arbitrary, and *extended random choice* where analyst observes frames.

The main results are as follows. Firstly, random choice admits BAUP representation if and only if it satisfies *Revealed Dominance Acyclicity Axiom*. We say that x is revealed to dominate y if there is a choice problem with menu including both x and y, such that x is chosen with probability one. Secondly, any non-degenerate random choice is rationalizable by BAUP decision maker with a binary state space. Thirdly, knowledge of extended random choice by itself does not allow to discipline the model. In particular, *almost any* extended random choice could be rationalized by a BAUP model with any given information structure which satisfies mild conditions of being *consistent with frames* and *fine*. This result implies that almost any dynamic random choice is rationalizable by BAUP model. Finally, paper argues that BAUP model with limited size of state space produces moderate framing effects if analyst observes extended random choice. The size of the state space could be as large as a half of size of the space of alternatives and still BAUP model would provide testable predictions. However test which has been found so far requires choice data with a lot of different frames. Thus the only sizes of state spaces which seemed to be useful are 2 and 3.

This paper contributes to the literature which considers how a (possibly hidden) resolution of uncertainty interferes with a process of decision making. Beside the results stated above, the paper argues that it is impossible to perform a good empirical test of a Bayesian model which makes predictions only about non-degeneracy of some choice probabilities; therefore we should try to discipline Bayesian models by imposing stronger structures on them. The paper provides two technical contributions: a way to construct subjective state space rationalizing DM's choices[31], and a hyperplanes' arrangement argument allowing to connect the size of the state space with the variation of preferences exhibited by the decision maker[32].

5 Appendix

5.0.1 Only if part of Theorem 1 for $h(i)$ with infinite signal space.

This is a direction for further work.

5.0.2 Proofs for Theorem 2

Proof for step 1. Consider $u_2(x^k) = k$, $u_1(x^k) = -\sum_{i=1}^{k-1}\left(\frac{\mu_i}{1-\mu_i}\right)$. Then

$$U_\mu(x^k) = \mu \cdot k - (1-\mu) \cdot \sum_{i=1}^{k-1}\left(\frac{\mu_i}{1-\mu_i}\right) \tag{113}$$

Let's start with the second statement. Consider $\mu \in (\mu_{k-1}, \mu_k)$. Assume $l > k$. Then $\mu < \mu_k < ... < \mu_l$ and we get

$$U_\mu(x^k) - U_\mu(x^l) = \mu \cdot (l-k)\left[\frac{1}{l-k}\sum_{i=k}^{l-1}\frac{\mu_i}{\mu}\cdot\frac{1-\mu}{1-\mu_i} - 1\right] > 0 \tag{114}$$

where we use

$$0 < a < b < 1 \quad \Rightarrow \quad \frac{b}{a}\cdot\frac{1-a}{1-b} = \frac{a^{-1}-1}{b^{-1}-1} > 1 \tag{115}$$

Analogously let $l < k$. Then $\mu_l < ... < \mu_{k-1} < \mu$ and we have

$$U_\mu(x^k) - U_\mu(x^l) = \mu \cdot (k-l)\left[1 - \frac{1}{l-k}\sum_{i=l}^{k-1}\frac{\mu_i}{\mu}\cdot\frac{1-\mu}{1-\mu_i}\right] > 0 \tag{116}$$

Therefore,

$$\mu \in (\mu_{k-1}, \mu_k) \quad \Rightarrow \quad \{x^k\} = \operatorname*{Argmax}_{x\in X} U_\mu(x) \tag{117}$$

Now assume $\mu \in [0, \mu_1)$. Consider $l > k$. Then again $\frac{\mu_i}{\mu}\cdot\frac{1-\mu}{1-\mu_i} > 1$ for $i = 1, ..., n-1$, thus eq. (114) gives us $U_\mu(x^k) > U_\mu(x^l)$. Analogously if $\mu \in (\mu_{n-1}, 1]$, then $\frac{\mu_i}{\mu}\cdot\frac{1-\mu}{1-\mu_i} < 1$

[31] See Theorem 1 and Lemma 1.

[32] See Theorem 4.

and eq. (116) gives us $U_\mu(x^k) < U_\mu(x^l)$ for any $l > k$. Step 1 is proven.

Proof of step 2. We want to show that σ_i given in step 2 of the proof of theorem 2 are well-defined Blackwell experiments.
(i) Firstly note that since $\sum\limits_{k\in\mathcal{J}_i} \rho^{A_i}(x^k) = 1$, we can check that

$$\sum_{k\in\mathcal{J}_i} \sigma_i(s_1, \tau_k^i) = 1 \qquad\qquad \sum_{k\in\mathcal{J}_i} \sigma_i(s_2, \tau_k^i) = 1 \tag{118}$$

(ii) Secondly, we have $\sigma_i(s_1, \tau_k) > 0$ and $\sigma_i(s_2, \tau_k) > 0$ for $k \in \mathcal{J}\backslash\{m_i\}$ by construction.
(iii) Finally,

$$\begin{aligned} 0 \quad &< \quad \frac{\epsilon}{\rho^{A_i}(x^{m_i})} \cdot \sum_{k\in\mathcal{J}_i\backslash\{m_i\}} \frac{k-\frac{1}{2}}{2n} \cdot \rho^{A_i}(x^k) \quad \leq \\ &\leq \quad \frac{\epsilon}{\rho^{A_i}(x^{m_i})} \cdot \frac{n-\frac{1}{2}}{2} \cdot (1-\rho^{A_i}(x^{m_i})) < 1 \end{aligned} \tag{119}$$

Thus $\nu_i^{m_i} \in (0, \overline{\nu})$ and we get $\sigma_i(s_1, \tau_{m_i}) > 0$, $\sigma_i(s_1, \tau_{m_i}) > 0$.

References

Ahn, David S and Todd Sarver (2013). "Preference for flexibility and random choice". In: *Econometrica* 81(1), pp. 341–361.

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

Block, Henry David, Jacob Marschak, et al. (1960). "Random orderings and stochastic theories of responses". In: *Contributions to probability and statistics* 2, pp. 97–132.

Dekel, Eddie, Barton L Lipman, and Aldo Rustichini (2001). "Representing preferences with a unique subjective state space". In: *Econometrica* 69(4), pp. 891–934.

Dillenberger, David et al. (2014). "A theory of subjective learning". In: *Journal of Economic Theory* 153, pp. 287–312.

Falmagne, Jean-Claude (1978). "A representation theorem for finite random scale systems". In: *Journal of Mathematical Psychology* 18(1), pp. 52–72.

Frick, Mira, Ryota Iijima, and Tomasz Strzalecki (2017). "Dynamic Random Utility". In: *http://cowles.yale.edu /sites /default /files /files /pub/d20/d2092.pdf*.

Gul, Faruk and Wolfgang Pesendorfer (2006). "Random expected utility". In: *Econometrica* 74(1), pp. 121–146.

Huber, Joel, John W Payne, and Christopher Puto (1982). "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis". In: *Journal of consumer research* 9(1), pp. 90–98.

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

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

Lu, Jay (2016). "Random choice and private information". In: *Econometrica* 84(6), pp. 1983–2027.

Manzini, Paola and Marco Mariotti (2014). "Stochastic choice and consideration sets". In: *Econometrica* 82(3), pp. 1153–1176.

McFadden, Daniel L (2005). "Revealed stochastic preference: a synthesis". In: *Economic Theory* 26(2), pp. 245–264.

Natenzon, Paulo (2012). "Random choice and learning". In: *34º Meeting of the Brazilian Econometric Society*.

Orlik, Peter and Hiroaki Terao (2013). *Arrangements of hyperplanes*. Vol. 300. Springer Science & Business Media.

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

Salant, Yuval and Ariel Rubinstein (2008). "(A, f): Choice with Frames". In: *The Review of Economic Studies* 75(4), pp. 1287–1296.

Shmaya, Eran and Leeat Yariv (2016). "Experiments on Decisions under Uncertainty: A Theoretical Framework". In: *American Economic Review* 106(7), pp. 1775–1801.

Tversky, Amos and Itamar Simonson (1993). "Context-dependent preferences". In: *Management Science* 39(10), pp. 1179–1189.