Plans of Action*

Evan Piermont[†] and Roee Teper[‡]

May 17, 2016

Abstract

We introduce a decision theoretic foundation for a class of learning models in which the decision maker's beliefs over the present uncertainty is dictated by the outcomes of her past actions. This type of learning underlies models of strategic experimentation. We construct a framework in which an alternative is a recursive function contingent at any stage on the outcomes of previous actions, and provide axiomatizations for subjective discounted expected utility maximization, both for independent actions and correlated actions. We point out that models of strategic experimentation have inherent limited observability, which in turn leads to partial identification of the subjective belief structure. We show that a class of processes we refer to as *strongly exchangeable* are the full characterization of Bayesianism in such environments.

Key words: Responsive learning; bandit problems; correlated arms; strong exchangeability.

*The authors wish to thank Eddie Dekel, Ehud Lehrer, Kyoungwon Seo, Teddy Seidenfeld and Eran Shmaya for their comments and remarks. Special acknowledgment to Kfir Eliaz for discussions which motivated us to think about this project.

[†]University of Pittsburgh, Department of Economics, `ehp5@pitt.edu`

[‡]University of Pittsburgh, Department of Economics, `rteper@pitt.edu`

1 Introduction

An information structure is *responsive* if observing the effect of taking actions is informative and shapes the decision maker's (DM's) belief regarding the uncertainty present in the environment. Responsive information structures have been extensively studied in the statistics literature (as bandit problems), and widely incorporated in economic models (as search problems, stopping problems, research and development, experimentation, portfolio design, etc).[1] These models (in both statistics and economics), with rare exception, assume the information structure is known and are concerned with characterizing the optimal or equilibrium behavior. There has been little work, however, on understanding the behavioral implications of these models in a general context. That is, if we had access to the data generated by the preferences of economic actors (as opposed to the underlying model of uncertainty in which they were making these choices), could we conclude they had in mind some responsive information structure, and if so, could we identify it?

In answering these questions, this paper achieves three goals. First, we introduce a new dynamic and recursive decision theoretic framework capturing the exploration-exploitation tradeoffs faced by a DM who needs to decide whether to try and learn about elements in the environment or take a known path. Second, in this framework we consider the behavioral (axiomatic) restrictions corresponding to responsive learning (generally), as well as the classic processes adhering to Bayesianism. Third, we discuss the limits in identifying the DM's subjective model of uncertainty, and how partial identification rises from the natural structure of a bandit problem.

In our framework, a DM is tasked with ranking sequential and contingent choice objects: the action the agent takes at any stage depends on the outcomes of previous actions. Formally, our primitive is a preference over *plans of action* (*PoAs*). Each action, a, is associated with a set of consumption prizes the action might yield, S_a. Then, a PoA is recursively defined as a *lottery over pairs* (a, f), where a is an action and f is a mapping that specifies the continuation PoA for each possible outcome in S_a. Theorem 1 shows that the construction of PoAs is well defined. So, a PoA specifies an action to be taken each period which can depend on the outcome of all previously taken actions. See Figures 1 and 2, where $f(x), f(y), f(z)$ are themselves PoAs. Each node in a PoA can be identified by a *history* of action-outcome realizations preceding it.

The set of actions in our model is in direct analogy to the set of arms in a bandit problem, or the set of actions in a repeated game. PoAs correspond to the set of all strategies in these environments. Note, however, the DM's perception of which outcome in S_a will result form taking action a is not specified. This is subjective and should be identified from the DM's preferences over PoAs.

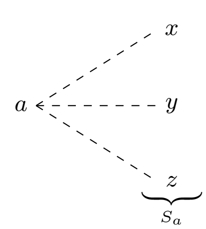

Figure 1: An action, a, and its support, S_a.

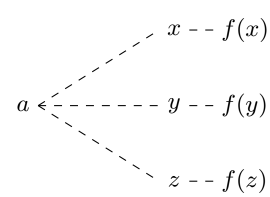

Figure 2: A degenerate PoA, (a, f).

[1] Bandit problems were introduced by Robbins (1952). See Berry and Fristedt (1985), for an overview of classic results within the statistics literature. For a survey on the applications to economics, see Bergemann and Välimäki (2008).

Theorem 2 axiomatizes preferences over PoAs of a DM who at each history entertains a belief regarding the outcome of future actions. That is, at each history h and for every action a, the DM entertains a belief $\mu_{h,a}$ over the possible outcomes S_a. In other words, $\mu_{h,a}(x)$ is the DM's subjective probability that action a will yield outcome x, contingent on having observed the history h. Given this family of beliefs, the DM acts as a subjective discounted expected utility maximizer, valuing a PoA p according to a *Subjective Expected Experimentation* (*SEE*) representation:

$$U_h(p) = \mathbb{E}_p\left[\mathbb{E}_{\mu_{h,a}}\left[u(x) + \delta U_{h'}(f(x))\right]\right], \tag{SEE}$$

where h' is the updated history (following h) when action a is taken and x is realized. All the parameters of the model –the consumption utility over outcomes, u, the discount factor, δ, and the history dependent subjective beliefs, $\{\mu_{h,a}\}_{h,a}$– are identified uniquely. The framework of PoAs we construct is a requisite to identifying responsive learning. Considering simpler environments (for example, sequences of actions), will not allow us to condition on the history of action-outcome realizations and modeling beliefs dependence on actions and their consequences will not be possible.

Our setup requires a formulation of a novel axiom termed *proportionality* (PRP): at any given history, the manner in which the DM evaluates continuation problems must be proportional to the manner in which she evaluates the consumption utility. Indeed, in order to ensure that the DM is acting consistently with a family of beliefs it must be that she assesses the value of each action according to the expectation of the consumption utility and discounted continuation utility it induces. Furthermore, it is necessary that the probabilistic weight she places on a given consumption utilities is the same as the weight she places on the corresponding continuation value.

To see how this is accomplished, consider the following simple example. There are three actions a, a' and b, all of which have support $\{x, y\} = X$. Assume that the continuation value for PoAs has been identified, so that continuation mappings are simply assignments into $\mathbb{R}$. In this environment, $f : X \to \mathbb{R}$, and (a, f) is the PoA such that action a is taken and the utility will be $u(x)+\delta f(x)$ if it yields x, and will be $u(y)+\delta f(y)$ if it yields y:

$$U(a, f) = \mu_a(x)[u(x) + \delta f(x)] + \mu_a(y)[u(y) + \delta f(y)]$$

Clearly, $(a, f) \succcurlyeq (a, g)$ if and only if $\mathbb{E}_{\mu_a}[f(\cdot)] \geqslant \mathbb{E}_{\mu_a}[g(\cdot)]$. Now imagine there is some $\alpha \in (0, 1)$ such that for any $f, g : X \to \mathbb{R}$,

$$\alpha(a, f) + (1 - \alpha)(a', f) \succcurlyeq \alpha(a, g) + (1 - \alpha)(a', g) \iff (b, f) \succcurlyeq (b, g). \tag{1}$$

By the above observation, this means $\mathbb{E}_{\alpha\mu_a+(1-\alpha)\mu_{a'}}$ and $\mathbb{E}_{\mu_b}$ induce the same ordering over $\mathbb{R}^X$. It can be concluded that $\alpha\mu_a + (1 - \alpha)\mu_{a'} = \mu_b$, and so, by the representation, it must be that

$$\alpha(a, f) + (1 - \alpha)(a', f) \sim (b, f), \tag{2}$$

for any $f : X \to \mathbb{R}$. PRP is the generalization of this observation. It states, whenever two PoAs stand in the relation as given by (1) they must also stand in the relation as given by (2). This guarantees that the DM's valuation is consistent with a single subjective distribution.

Theorem 2 shows that PRP, along with (some of the) standard behavioral conditions for discounted expected utility, is necessary and sufficient for an SEE representation. While the axiomatization does not point to the optimal strategy in general strategic experimentation problems, which is known to be a hard problem to solve when actions are correlated, it provides (like most axiomatization theorems) a unifying guidance as to what might or might not be ruled out.

This result provides the groundwork for the analysis of subjective responsive belief updating, but does not in any way restrict the evolution of the beliefs. Nonetheless, the representation is still *responsive* in the sense that beliefs at any point in time are a function of the prior action-outcome realizations. Therefore, this is the most general level at which the exploration-exploitation tradeoff can be captured in a decision problem. From here we consider commonly used responsive learning models, and ask whether beliefs evolving according to such models could be identified in a meaningful way.

For every history, h, the modeler can identify a family of action-indexed beliefs $\{\mu_{h,a}\}_a$. Note, this is not necessarily a full characterization of the (subjective) stochastic process the DM considers. The modeler does not have direct access to the DM's joint belief over the outcomes of the different actions, or, in other words, the probability that action a yields outcome x contingent on (the counterfactual) that action b would have yielded outcome y *in the same period*. We only have access to the marginals of this joint distribution. At first glance, this might seem like an irrelevant piece of data to consider, since only one action can be taken each period. However, the canonical specification of bandit problems is with respect to a *joint* distribution over all arms. If we wish to fully understand the updating process, characterize the DM's world view, or extrapolate beyond the observations, the marginal distributions may prove inadequate. To be clear, this is not a limitation of the current setup any more than of bandit problems in general: observing a single pull of an arm each period simply does not provide sufficient data to uniquely identify the underlying stochastic process.

Bearing in mind this observability concern, we first examine the relatively simple environment in which there is no informational spill-over from one action to another. That is, observing the outcome of action b does not affect the DM's beliefs regarding the outcome of action a. This corresponds to the widely studied class of bandit problems with independent arms. Axiomatically, this is ensured by the restriction that histories sharing the same outcomes for action a induce the same beliefs for action a, regardless of the outcome of other actions. With such restrictions in place, the evolution of beliefs of each arm can be looked at in isolation –this is dual to the notion of index rules (see Gittins and Jones (1974)), where the optimal strategy can be formed by analyzing each arm independently. In line with an expanding literature on Bayesianism, adding a within-arm symmetry axiom suited for the current framework provides a Bayesian (in particular, exchangeable) structure to the family of beliefs. Theorem 3 formally presents this axiomatization result.

Because of the observability issues discussed above, matters are more complicated when considering correlated actions. Nevertheless, we show that exchangeable processes can still be characterized, but only partial identification is possible. Our results indicate that, while it is a hard question finding the optimal strategy in a general correlated bandit problem, it is possible to point to a unified underlying behavior for all such problems.

We elaborate. Consider a DM who believes there is an underlying distribution governing the joint realization of actions. She does not know exactly what this distribution is, but entertains a prior probability over what it might be. In addition, she believes that conditional on the true distribution, the actions are *inter-temporally* i.i.d. Due to de Finetti (1931); Hewitt and Savage (1955), these classical Bayesian updating processes are referred to as exchangeable. In our framework, as in the classic bandit literature, histories do not specify the joint realization of all actions; contrast this to the standard symmetry axiom (see, for example, Klibanoff et al. (2013)) that assumes that a history fully specifies the evolution of the state. Nevertheless, Proposition 4 states that a preference-based symmetry axiom, Sym, adapted to our framework is still necessary and sufficient for the beliefs of an SEE representation to follow an exchangeable processes. The inherent observability constraint, however, bears a cost; the exchangeable process with which these marginals are consistent is not unique.

In light of this generic non-uniqueness, we introduce what we term *strongly exchangeable* processes. This a subclass of the widely studied exchangeable processes, where the periodic state-space takes a product structure, and conditional on the true distribution the marginals are independent. Thus, a strongly exchangeable process is one in which, conditional on the distributional parameter, outcomes are both inter-temporally and *contemporaneously* independent.

Theorem 5 characterizes strongly exchangeable processes via a strengthening of the standard (probabilistic) symmetry property. Theorem 6 states, however, that this characterization provides no additional restrictions when projected to a domain similar to the one studied here, in which only marginals are observable. Finally, Theorem 7 synthesizes these results, showing that if an SEE DM satisfies Sym, then there is a unique consistent strongly exchangeable process; every exchangeable process corresponds to a unique strongly exchangeable process which induces the exact same preferences over PoAs. Put differently, there is no observable choice data over PoAs (i.e., no behavior in bandit type problems) that separates exchangeability and strong exchangeability. We conclude, strong exchangeability is the full characterization of Bayesianism in bandit type problems.

The paper is structured as follows. The subsequent section formally introduces the environment and the construction of plans of action. Section 3 presents the primitive and the axioms for an SEE representation. Section 4 introduces the notion of an SEE representation being consistent with a stochastic process. Then, Section 5 presents exchangeable processes and the consistency of an SEE representation with such processes when actions are independent (Section 5.1) and correlated (Section 5.2). Section 6 introduces the notion of strong exchangeability and presents our (non) uniqueness result. Section 7.1 discusses the related literature. An informal discussion regarding how a decision theoretic model would incorporate exogenous information in addition to responsive learning appears in Section 7.2. Lastly, Section 7.3 discusses the point of disagreement among Bayesians in environments of experimentation. All the proofs are in the Appendices.

2 Technical Framework

The purpose of the current section is to formally construct the different choice objects, termed *plans of action* (*PoAs*). The primitive of our model, as presented in the subsequent section, is a preference relation over all PoAs.

Let X be a finite set of outcomes, endowed with a metric d_X. Outcomes are consumption prizes. For any metric space, M, let $\mathcal{K}(M)$ denote the set non-empty compact subsets of M, endowed with the Hausdorff metric. Likewise, for any metric space M, denote $\Delta^{\mathcal{B}}(M)$ as the set of Borel probability distributions over M, endowed with the weak*-topology, and $\Delta(M)$ the subset of distributions with denumerable support.

Let $\mathcal{A}$ be a compact and metrizable set of actions. Each action, a, is associated with a set of outcomes, $S_a \in \mathcal{K}(X)$, which is called the support of the action. We assume the map $a \mapsto S_a$ is continuous and surjective. For any metric space M, let $\mathcal{A} \otimes M = \{(a, f) | a \in \mathcal{A}, f \colon S_a \to M\} = \{(a, \{(x_i, m_i)\}_{i \in I}) \in \mathcal{A} \times \mathcal{K}(X \times M) | \bigcup_{i \in I}\{x_i\} = S_a \text{ and } x_i \neq x_j, \forall i \neq j \in I)\}$, endowed with the subspace topology inherited from the product topology. By the continuity of $a \mapsto S_a$ we know that the relevant subspace is closed and hence the topology on $\mathcal{A} \otimes M$ is compact whenever M is. We can think of f as the assignment into M for each outcome in th-e support of action a. For any $f : X \to M$ we will abuse notation and write (a, f) rather than $(a, f|_{S_a})$.

With these definitions we can define PoAs. A PoA is a tree of actions, such that each period the DM receives a lottery (with denumerable support) over actions conditional on the outcomes for each of the previous actions. The formal construction of PoAs is both conceptually and technically involved, and hence, relegated Appendix C.

Informally, however, set $P_0 = \Delta(\mathcal{A})$; a 0-period plan is a lottery (with denumerable support) over actions. Given an action, an element of its support is realized and the plan is over. Then a 1-period plan, p_1 is a lottery over actions, and continuation mappings into 0-period plans: $p_1 \in \Delta(\mathcal{A} \otimes P_0)$. Given the realization of an action-continuation pair, (a, f), in the support of p_1, and the realized element of the support, $x \in S_a$, the DM receives a 0-period plan, as given by $f(x)$. Continuing in this fashion, we can define recursively,

$$\hat{P}_n = \mathcal{A} \otimes P_{n-1}$$
$$P_n = \Delta(\hat{P}_n).$$

Define $P^* = \prod_{n \geqslant 0} P_n$. P^* is the set of all PoAs (including inconsistent ones). We restrict ourselves to the set of *consistent* elements of P^*: those elements such that, the $(n-1)$-period plan implied by the n-period plan is the same as the $(n-1)$-period plan. Let $P \subset P^*$ denote the set of consistent PoAs.[2] Finally, set $\hat{P} = \mathcal{A} \otimes P$. We can associate $P \cong \Delta(\mathcal{A} \otimes P) = \Delta(\hat{P})$. Of course, one might be worried that, for a given p, although each p_n is a denumerable lottery, the associated element might live in $\Delta^B(\mathcal{A} \otimes P)$ rather than $\Delta(\mathcal{A} \otimes P)$. Indeed, we need also to restrict our attention to the set of plans that have countable support not just for each finite level, but also "in the limit," and whose implied continuation plans are also well behaved in such a manner. Fortunately, this can be done:[3]

Theorem 1. *There exists a homeomorphism between P and $\Delta(\mathcal{A} \otimes P)$.*

Proof. In Appendix C. ■

As a final notational comment, denote by $\Sigma \subset P$ the set of *objective* PoAs. That is the plans of action which specify only actions with deterministic outcomes (i.e., $\{a | S_a = \{x\}, x \in X\}$). We can identify $\Sigma \cong \Delta(X \times \Sigma)$.

[2]See Appendix C for details.

[3]One can also consider measurable lotteries (instead of lotteries with countable support). In fact, the construction of the homeomorphism in Appendix C considers measurable lotteries. In the paper we focus on discrete support for notational cleanliness (see footnote 4) and tractability (to avoid measurability issues in proofs). We justify our focus by noting that $\Delta(\hat{P})$ is dense in $\Delta^{\mathcal{B}}(\hat{P})$ and so, given continuity (Axiom vNM), preferences over the more general objects are recoverable.

Notation	Meaning	Notes
$x, y, z \in X$	Single period consumption prizes	
$a \in A$	actions	associated with the support S_a
$(a, f) \in \hat{P}$	action-continuation pair	$a \in A$ and $f : S_a \to P$
$p, q \in P$	PoA	Lottery over action-continuation pairs
$h \in \mathcal{H}(p)$	a history	corresponds to a unique node of p.

Figure 3: List of notational conventions.

2.1 Histories

PoAs are infinite trees; each node, therefore, is itself a PoA, p –a distribution over action-continuation pairs. Each action-continuation, (a, f), in the support of a node contains branches to new nodes (PoAs). The branches emanating from an action coincide with the outcomes in the support of that action, $x \in S_a$. The node that follows x is the PoA specified by $f(x)$. Each node, therefore, is reached after a unique history: the history specifies the realization of the distribution of each pervious node, and outcome of the action realized. Thus, for a given PoA, p, each history of length n is an element of $\prod_{t=1}^{n} P \times \hat{P} \times X$ such that $p^1 = p$ and

$$\begin{aligned}(a^t, f^t) &\in supp(p^t)\\ x^t &\in S_{a^t}\\ p^{t+1} &= f^t(x^t)\end{aligned}$$

Define the set of all histories of length n for p as $\mathcal{H}(p, n)$ and the set of all finite histories as $\mathcal{H}(p)$. Let $\mathcal{H}(n) = \bigcup_{p \in P} \mathcal{H}(p, n)$ and, $\mathcal{H} = \bigcup_{n \in \mathbb{N}} \mathcal{H}(n)$. For each $h \in \mathcal{H}(p, n)$, h corresponds to the node (PoA) defined by $f^n(x^n)$. Lastly, for any $p, q \in P$ and $h \in \mathcal{H}(p)$ define $p_{-h}q$ as the (unique!) element of P that coincides with p everywhere except after h in which case $f^n(x^n)$ is replaced by q. Note that the n period plan implied p and $p_{-h}q$ are the same. For any $p, q \in P$ and $n \in \mathbb{N}$, let $p_{-n}q \equiv \bigcup_{h \in \mathcal{H}(p,n)} p_{-h}q$.

Finally, for any h, h' in $\mathcal{H}(n)$ we say that h and h' are $\mathcal{A}$-*equivalent*, denoted by $h \overset{\mathcal{A}}{\sim} h'$ if $\text{proj}_{\prod_{t=1}^{n} \mathcal{A} \times X}(h) = \text{proj}_{\prod_{t=1}^{n} \mathcal{A} \times X}(h')$. That is, two histories of length n are $\mathcal{A}$-equivalent, whenever they correspond to the same sequence of action-realization pairs, ignoring the objective randomization stage of each period and the continuation assignment to outcomes that did not occur. It will turn out, we are only interested in the $\mathcal{A}$-equivalence classes of histories. Technically, this is the consequence of the linearity of preference and indifference to the resolution of uncertainty (as shown in Lemma 2); conceptually, this is because all uncertainty in the model regards the realization of actions, and so, observing objective lotteries has no informational benefit.

3 Subjective Expected Experimentation

The primitive in our model is a preference relation $\succcurlyeq \subseteq P \times P$ over all PoAs. When specific PoA and history are fixed, the preferences induce history dependent preferences as follows: for any $p \in P$, and $h \in \mathcal{H}(p)$ define $\succcurlyeq_h \subseteq P \times P$ by

$$q \succcurlyeq_h r \iff p_{-h}q \succcurlyeq p_{-h}r.$$

The following axioms will be employed over all history induced preferences.[4] A history is *null* if $\succcurlyeq_h$ is a trivial relation. This first four axioms are variants on the standard fare for discounted expected utility. They guarantee the expected utility structure, non-triviality, stationarity and separability (regarding objects over which learning cannot take place), respectively.

A1. (vNM). *The binary relation, $\succcurlyeq_h$ satisfies the expected utility axioms. That is: weak order, continuity (defined over the relevant topology, see Appendix C) and independence.*

We require a stronger non-triviality condition that is standard, because of the subjective nature of the dynamic problem. We need to ensure the DM believes *some* outcome will obtain. Therefore, not all histories following a given action can be null.

A2. (NT). *For any non-null h, and any (a, f), not all $h' \in h \times \mathcal{H}((a, f), n)$ are null.*

Of course, the nature of the problem at hand precludes stationarity and separability in full generality. Since the objective is to let the DM's beliefs depend on prior outcomes explicitly, her preferences will as well. However, the DM's beliefs do not influence her assessment of objective plans (i.e., elements of Σ), and so it is over this domain that stationarity and separability are retained. This means, the DM's preferences *in utility terms* are stationary and separable, but we still allow the conversion between actions and utils to depend on her beliefs which change responsively.

A3. (SST). *For all non-null $h \in \mathcal{H}$, and $\sigma, \sigma' \in \Sigma$,*

$$\sigma \succcurlyeq \sigma' \iff \sigma \succcurlyeq_h \sigma'.$$

A4. (SEP). *For all $x, x' \in X, \rho, \rho' \in \Sigma$ and $h \in \mathcal{H}$,*

$$\big(\frac{1}{2}(x, \rho) + \frac{1}{2}(x', \rho')\big) \sim_h \big(\frac{1}{2}(x, \rho') + \frac{1}{2}(x', \rho)\big).$$

Because of the two-stage nature of the resolution of uncertainty each period (first, the resolution of lottery over $\hat{P}$, and then the resolution of the action over X), we need an additional separability constraint. From the point of view of period n, and when considering the continuation problem beginning in period $n+1$, the DM should not care if uncertainty is resolved in period n (when the action-continuation pair is realized), or in period $n + 1$. That is, we also assume the DM is indifferent to the timing of objective lotteries *given a fixed action.*

A5. (IT). *For all $a \in \mathcal{A}$, $h \in \mathcal{H}$, $\alpha \in (0, 1)$, and $(a, f), (a, g) \in \hat{P}$,*

$$\alpha(a, f) + (1 - \alpha)(a, g) \sim_h (a, \alpha f + (1 - \alpha)g),$$

where mixtures of f and g are taken point-wise.

[4] It is via the use of this construction that our appeal to denumerably supported lotteries provides tractability. If we were to employ lotteries with uncountable support, then histories as constructed in Section 2.1 would, in general, be zero probability events; under the expected utility hypothesis, $\succcurlyeq_h$ would be null for all $h \in \mathcal{H}$. This could be remedied by appealing to histories as *events* in $\mathcal{H}$, measurable with respect to the filtration induced by previous resolutions of lottery-action-outcome tuples. We believe that this imposes a unnecessary notational burden.

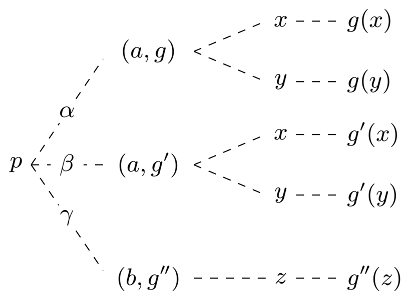

Figure 4: A PoA, p, defined by $p(a, g) = \alpha$, $p(a, g') = \beta$ and $p(b, g'') = \gamma = 1 - \alpha - \beta$.

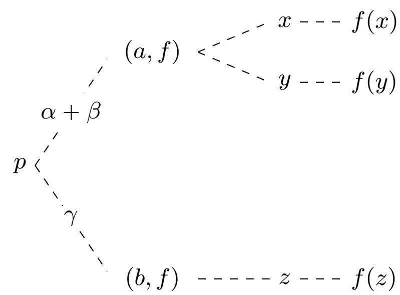

Figure 5: The PoA, $p.f$ where $f : X \to P$, and p is defined in figure 4. Notice, $p.f(a, f) = p(a, g) + p(a, g') = \alpha + \beta$.

Thus far the axioms introduced are somewhat standard. However, in our particular framework these assumptions do not guarantee that the value of the action is tied together with its realization of consumption alternatives. To do this requires two steps. First, construct a subjective distribution over each action. This will be done by looking at the ranking of continuation mappings for each action (i.e., (a, f) compared to (a, g)). These continuation mappings function exactly as "acts" in the standard subjective expected utility paradigm –and so, standard techniques allow for the identification of such a subjective distribution. Second, we need to ensure that the value assigned to arbitrary PoAs is the expectation according to these beliefs. Towards this, the following notation is introduced.

Definition. *For any function $f : X \to \mathcal{P}$, define $p.f \in P$ as $p.f[(a, g)] = p[\{(b, h)|b = a\}]$ if $g = f$, and $p.f[(a, g)] = 0$ if $g \neq f$.*

Take note, because we are dealing with simple probability distributions, we have no measurability concerns. The plan of action $p.f$ has the same distribution over actions in the first period, but the continuation plan is unambiguously assigned by f as shown in Figures 4 and 5. If the original plan is in $\hat{P}$, then the dot operation is simply a switch of the continuation mapping: $(a, g).f = (a, f)$. This operation is introduced because it allows us to isolate the subjective distribution of the first period action.

Definition. *We say that $p, q \in P$ are **h-proportional** if for all $f, g : X \to \Sigma$.*

$$p.f \succcurlyeq_h p.g \iff q.f \succcurlyeq_h q.g$$

Intuitively, this is just the Savagian (1954) notion of more or less likely than. Since the images of f and g are in Σ, there is no informational effect from observing the outcome of p. Hence, f and g can be thought of as objective assignments into continuation utilities, and so, $p.f$ is preferred to $p.g$ if f is more likely than g to yield good continuation utilities, given the distribution induced by the actions in the support of p.

A6. (PRP). *For all $p, q \in P$, and $f : X \to \Sigma$ if p and q are h-proportional then $p.f \sim_h q.f$.*

If p and q are h-proportional, then f is more likely than g to provide good continuation utilities under p whenever it is also the case under q. Obviously, p and q must induce the same distributions over consumption outcomes today. But then, the preference between p and q must be the consequence of only the continuation values. If we replace the continuation problems with objectively equivalent plans, the DM should be indifferent between p and q.

The following is our general axiomatization result. It states that the properties above characterize a DM who, when facing a PoA, calculates the subjective expected utility according to a collection of history dependent beliefs over action-outcome pairs, and among different PoAs contemplates the benefits of consumption versus learning.

Theorem 2 (Subjective Expected Experimentation Representation). *If $\succcurlyeq_h$ satisfies vNM, NT, SST, SEP, IT and PRP then there exists a utility index $u : X \to \mathbb{R}$, a discount factor $\delta \in (0,1)$, and a family of beliefs $\{\mu_{h,a} \in \Delta(S_A)\}_{h \in \mathcal{H}, a \in \mathcal{A}}$ such that*

$$U_h(p) = \mathbb{E}_p \left[\mathbb{E}_{\mu_{h,a}} \left[u(x) + \delta U_{h'(a,x)}(f(x)) \right] \right], \tag{SEE}$$

jointly represents $\{\succcurlyeq_h\}_{h \in \mathcal{H}}$, where $h'(a,x) = (h, p, (a, f), x)$. Moreover, u is cardinally unique, δ is unique, the family of beliefs is unique, and $\mu_{h,a} = \mu_{h',a}$ whenever $h \overset{\mathcal{A}}{\sim} h'$.

Proof. In Appendix B. ∎

SEE representation does not put any restrictions on the dynamics of the beliefs. The behavioral restrictions imposed by classical Bayesian models is the focus of subsequent sections.

4 Consistent Stochastic Processes

In order for a modeler to understand the DM's updating process (in particular, when actions are not independent), we need to construct her beliefs regarding not only each action individually but also her joint beliefs regarding the correlation between actions. Because only a single action can be taken each period, the beliefs identified in Theorem 2 correspond only to the marginals (on each action) of the process governing all actions jointly. As we will see, in the generic case we have insufficient data to uniquely identify the (subjective) joint distribution. We will still, however, be able to identify a representative with unique properties.

Towards this goal, we need to define the general class of processes and the notion of consistency. Let $\mathcal{S}_\mathcal{A} \equiv \prod_{a \in \mathcal{A}} S_a$, and $\mathcal{S} \equiv \prod_{n \geqslant 0} \mathcal{S}_\mathcal{A}$. Let $\mathcal{S}_n$ be the projection onto the n^{th} component. For any $\mathcal{A}' \subseteq \mathcal{A}$, let $\mathcal{S}_{\mathcal{A}'} \equiv \prod_{a' \in \mathcal{A}'} S_{a'}$ (so, $\mathcal{S}_a \cong S_a$), and $\mathcal{S}_{n,\mathcal{A}'}$ the projection of $\mathcal{S}_n$ onto $\mathcal{S}_{\mathcal{A}'}$. Equip $\mathcal{S}$ with the standard product σ-algebra.

$\mathcal{S}$ represents the grand state-space; a state, s, determines the realization of each action in each period –an entity unobservable to the modeler. What we can observe are the DM's beliefs regarding the set of events that are generated by histories. For any $n \in \mathbb{N}$ and $h = (a^1, x^1 \ldots a^n, x^n) \in \mathcal{H}(n)/\overset{\mathcal{A}}{\sim}$, identify h with the cylinder it generates in $\mathcal{S}$, that is $\{s \in \mathcal{S} | s_{m,a^m} = x^m, m \leqslant n\}$.

Definition. *An SEE belief structure, $\{\mu_{h,a}\}_{h \in \mathcal{H}, a \in \mathcal{A}}$, is **consistent** with $\zeta \in \Delta^\mathcal{B}(\mathcal{S})$ if $\mu_{h,a}(x) = \frac{\zeta(h,a,x)}{\zeta(h)}$ for all $h \in \mathcal{H}$ and $a \in \mathcal{A}$.*

That is, $\mu_{h,a}$ is consistent with ζ if following every history h and for every action a, the conditional ζ-distribution over S_a generated by action a coincides with $\mu_{h,a}$. Given an SEE belief structure, the consistent processes are those that the modeler cannot rule out as the DM's subjective world-view. Because, in general, this set of consistent processes will be well populated, it is prudent to make precise exactly the limits of observability.

Definition. *We say the processes $\zeta, \xi \in \Delta^{\mathcal{B}}(\mathcal{S})$ are **SEE-equivalent** if there is an SEE representation consistent with both processes.*

Two SEE decision makers holding beliefs over $\mathcal{S}$ given by ζ and ξ, respectively, will have the same evolution of SEE beliefs, and therefore, assuming their utilities indices coincide, the same preferences over PoAs. Thus, given the data encoded in $\succcurlyeq$, we can only hope to recover beliefs upto SEE-equivalence. The remainder of this paper characterizes the relation between behavioral axioms and the resulting set of SEE-equivalent consistent processes.

5 Exchangeable Processes and Consistency

Because it forms the basis for the statistical literature on bandit problems, we will pay particular attention to the class of *exchangeable* processes.

Definition. *Let $\hat{\Omega} = \prod_{n \geqslant 0} \Omega$ be a probability space. The process $\zeta \in \Delta^{\mathcal{B}}(\hat{\Omega})$ is **exchangeable**, if there exists probability measure θ over $\Delta^{\mathcal{B}}(\Omega)$, such that*

$$\zeta(E) = \int_{\Delta^{\mathcal{B}}(\Omega)} \hat{D}(E) d\theta(D), \tag{3}$$

where for any $D \in \Delta^{\mathcal{B}}(\Omega)$, $\hat{D}$ is the corresponding product measure over $\hat{\Omega}$.

Remark 1. *If ζ is exchangeable, then θ is unique.*

Exchangeable processes were first characterized by de Finetti (1931, 1937) and later extended by Hewitt and Savage (1955). Their fundamental result states that a process $\zeta \in \Delta^{\mathcal{B}}(\hat{\Omega})$ is exchangeable if and only if for any finite permutation $\pi : \mathbb{N} \to \mathbb{N}$ and event $E = \prod_{n \in \mathbb{N}} E_n$, we have

$$\zeta(E) = \zeta(\prod_{n \in \mathbb{N}} E_{\pi(n)}). \tag{4}$$

Exchangeable processes are of clear statistical importance, in particular within the subjectivist paradigm (see, for example Schervish (2012)). From the economic vantage, a DM who understands there to be an exchangeable process governing the outcome of actions would be considered Bayesian.[5] This is because, given the representation in Eq. 3, the DM (acts as if she) entertains a second order distribution, which she updates following every observation. As the number of observations increase, her belief will converge (and assuming she placed positive second order probability on the true distribution, her beliefs will converge thereto).

5.1 Independent Arms

The simplest possible structures are those in which the DM's beliefs regarding the outcome of action a depends only on her previous observations regarding action a. The observed outcome of other actions contains no information regarding a. Under the interpretation of actions as the arms of a bandit problem, this corresponds to the arms being independent.

[5] It is possible to consider more general Bayesian models than exchangeable processes. At least for the case of independent actions, for example, it is not hard to adapt a local consistency axiom as in Lehrer and Teper (2015) that will imply that beliefs follow a general martingale process.

In the same spirit of $\mathcal{A}$-equivalence, we say two histories are a-equivalent whenever they correspond to the same sequence of observations regarding action a, and ignoring all other actions.

Definition. *Let $a(h)$ denote the unique history of a-observations given h, defined recursively according to, $a(\varnothing) = \varnothing$ and $a(h, p, (b, f), x)$ as $a(h)$ if $b \neq a$ and $(a(h), x)$ if $b = a$.*

Then independence of arms is clearly given by the dictate that beliefs regarding action a only depend on the equivalence classes generated by a-observations. Formally,

A7. (IA). *For all non-null h, h' if $a(h) = a(h')$ then*

$$a.f \succcurlyeq_h a.g \iff a.f \succcurlyeq_{h'} a.g,$$

for all $f, g : X \to \Sigma$.

With the addition of IA, the DM's problem becomes much simpler; it is well known that independent arms allow for the use of index rules to generate optimal behavior. Moreover, our job (as modelers), becomes simpler as well. Because each action bears no informational content regarding other actions, the family of marginal distributions identified in Theorem 2 completely describes the decision maker's beliefs. Another consequence of IA is that the stochastic process governing each action is *time* independent. Because the resolution of the outcome of a given action does not depend on the outcomes of other actions, it naturally does not depend on how many times other actions have been taken.

Of particular interest is the special case of the above, where not only is independence preserved across actions, but also *within* actions. That is, the resolution of the outcome of each is dictated by an i.i.d. process. Of course, the DM might entertain uncertainty about which i.i.d. process governs the resolution of the action, and learning takes place with respect to this "second-order" distribution. This special case corresponds exactly to the well studied class of bandit models in which each arm provides prizes according to an unknown i.i.d. process, where following every pull of an arm the DM Bayesian updates her belief regarding the true distribution underlying that arm. To deliver this requires the following axiom.

A8. (WA-Sym). *Fix $n \in \mathbb{N}$, $a \in \mathcal{A}$. Let $p \in P$ be such that a is assigned unambiguously for the first n periods.*[6] *Then, for all $q, r, r' \in P$, and for all $h, h' \in \mathcal{H}(p, n)$, such that $a(h)$ is a permutation of $a(h')$ then*[7]

$$p_{-n}q \succcurlyeq (p_{-n}r)_{-h}r' \iff p_{-n}q \succcurlyeq (p_{-n}r)_{-h'}r'.$$

It is easiest to understand the restriction imposed by WA-Sym when thinking of $q, r, r' \in \Sigma$, as a set of fixed continuation utilities. If the decision maker believes history h will occur with probability α, then the left hand side of the implication is true whenever $U(q) \geqslant \alpha U(r) + (1 - \alpha)U(r')$. Then, WA-Sym requires this is also true regarding the same bet on h'. So the probability of h' must also be α. Further, since this also must hold when $q, r, r' \notin \Sigma$, it states that DM's assessment of arbitrary plans must be the same after h and h'. In other words, both the ex-ante probability of h and h', and the informational content of h and h' coincide. This is exactly the symmetry property that characterizes exchangeability within each action.

[6] That is, for every $h \in \mathcal{H}(p, n)$, $\text{proj}_{\mathcal{A}^n} h = (a, a, \ldots, a)$.

[7] Recall, $p_{-n}q \equiv \bigcup_{h \in \mathcal{H}(p,n)} p_{-h}q$.

Theorem 3 (Independent Arms, Exchangeable Process)**.** *Let $\succcurlyeq$ admit an SEE representation with beliefs $\{\mu_{h,a}\}_{h\in\mathcal{H},a\in\mathcal{A}}$. Then $\succcurlyeq_h$ satisfies IA and WA-Sym if and only if for every $a \in \mathcal{A}$, $\mu_{h,a}$ depends only on $a(h)$ and $\big\{\{\mu_{a(h),a}\}_{a(h)\in\mathcal{H}/\overset{a}{\sim}}\big\}$ is consistent with a unique exchangeable process.*

Proof. In Appendix B. ∎

Of course, because the belief regarding each action is an exchangeable process, and because the actions are independent, we can construct the unique joint distribution that is the product of each action-indexed belief. It is immediate that this joint distribution is the unique exchangeable process with which SEE beliefs are consistent.

5.2 Correlated Arms

If we relax the assumption that actions are uncorrelated, (i.e., when IA does not hold), the same technique as above can still guarantee that SEE beliefs are consistent with some exchangeable process. This requires only a slight reformulation of WA-Sym:

Definition. *Let π be an n-permutation and $p, q \in P$. We say that q is π-permutation of p if for all $h \in \mathcal{H}(p,n)$, $h' \in \mathcal{H}(q,n)$, $\mathrm{proj}_{\mathcal{A}^n} h = \pi\big(\mathrm{proj}_{\mathcal{A}^n} h'\big)$.*

If p admits any π-permutations it must be that the first n actions are assigned unambiguously (i.e., it does not depend on the realization of the prior outcomes nor the objective randomization).

***A8**. (Sym). *Let π be an n-permutation and $p, p' \in P$ with p' a π-permutation of p. Then, for all $a \in \mathcal{A}$, $q, r, r' \in P$, and $h \in \mathcal{H}(p,n)$, $h' \in \mathcal{H}(p',n)$, if h is a permutation[8] of h' then*

$$p_{-n}q \succcurlyeq (p_{-n}r)_{-h}r' \iff p_{-n}q \succcurlyeq (p_{-n}r)_{-h'}r'.$$

Sym requires a similar condition as WA-Sym to hold for permutations across different actions As before, Sym implies the ex-ante probability of h and h' coincide.

Proposition 4 (Correlated Arms, Exchangeable Process)**.** *Let $\succcurlyeq$ admit an SEE representation with beliefs $\{\mu_{h,a}\}_{h\in\mathcal{H},a\in\mathcal{A}}$. Then $\succcurlyeq_h$ satisfies Sym if and only if $\big\{\{\mu_{h,a}\}_{h\in\mathcal{H}}\big\}_{a\in\mathcal{A}}$ is consistent with an exchangeable process.*

Proposition 4 will be an obvious corollary of Theorem 7, and so, is stated now without proof.

6 Strong Exchangeability

Unfortunately, without the structure provided by across-action independence, Sym is not sufficient to obtain a unique exchangeable process consistent with an SEE representation. This lack of identification stems directly from the inability to observe the DM's belief regarding *contemporaneous* correlations. Consider two coins, a and b, which can both take values in $\{H, T\}$. Both coins are flipped each period. Consider the following two governing processes, which are i.i.d. across time periods. (1) the coins are perfectly correlated

[8]Ultimately, due to the definition of $\overset{\mathcal{A}}{\sim}$, the relevant aspect of each history is the list of action-outcome realizations (Lemma 2): for example, action a and outcome x, followed by action b and outcome y. A permutation of this history would be action b and outcome y, followed by action a and outcome x. Also, note, histories do not specify the joint realization of all actions and thus the standard symmetry axiom cannot be employed; directly resorting to de Finetti's theorem is not possible.

(with equal probability on HH and TT), or (2) the coins are identical and independent (and both have equal probability on H and T). Notice, the two cases induce the same marginal distributions over each coin *individually*. So, if the modeler has access only to the DM's marginal beliefs, the two processes are indistinguishable.

In this section we introduce a strengthening of exchangeability, which we aptly call *strongly-exchangeable*, under which independence is preserved both inter-temporally (as in vanilla exchangeability) and *contemporaneously*.[9] Following the intuition above, it should come as no surprise that under Sym strongly-exchangeability can never be ruled out. In other words, there is no behavior (i.e., no preference over PoAs) that distinguishes exchangeability from strongly-exchangeability.

Definition. *A process $\zeta \in \Delta^{\mathcal{B}}(\mathcal{S})$ is **strongly-exchangeable** if there exists a probability measure, θ over $\Delta^{IN} \equiv \prod_{a \in \mathcal{A}} \Delta(S_a)$, such that*

$$\zeta(E) = \int_{\Delta^{IN}} \hat{D}(E) d\theta(D),$$

where for any $D \in \Delta^{IN}$, $\hat{D}$ is the corresponding product measure over $\mathcal{S}$.

So the outcomes of actions that occur at the same time are independently resolved. Of course, this does not impose that there is no informational cross contamination between actions. Information regarding the distribution of action a is informative about the underlying parameter governing the exchangeable process, and therefore, also about the distribution of action b.

Just as exchangeability can be characterized by the invariance to permutations of ex-ante probability, so too can strong exchangeability.

Theorem 5. *The process $\zeta \in \Delta^{\mathcal{B}}(\mathcal{S})$ is strongly-exchangeable if and only if for any set of finite permutations $\{\pi_a : \mathbb{N} \to \mathbb{N}\}_{a \in \mathcal{A}}$ and event $E = \prod_{n \in \mathbb{N}} \prod_{a \in \mathcal{A}} E_{n,a}$, we have*

$$\zeta(E) = \zeta(\prod_{n \in \mathbb{N}} \prod_{a \in \mathcal{A}} E_{\pi_a(n),a}). \tag{5}$$

Proof. In Appendix B. ■

Strongly exchangeable process are ones where each dimension can be permuted independently. If $\pi_a = \pi_b$ for all $a, b \in \mathcal{A}$, the condition is exactly exchangeability. Strongly exchangeable process are especially relevant with respect to the current focus because they act as representative members to the equivalence classes generated by SEE-equivalence.

Theorem 6. *Let $\xi \in \Delta^{\mathcal{B}}(\mathcal{S})$ such that $\xi(h) = \xi(\pi(h))$, for any $h \in \mathcal{H}(n)$ and n-permutation π (where $\pi(h)$ is the permutation of entry of the histories). Then there exists a unique strongly exchangeable process that is SEE-equivalent to ξ.*

Proof. In Appendix B. ■

Putting these results together provides us with a precise limit to the identification of an SEE belief structure under exchangeability.

[9] We feel reasonably certain that strong exchangeability must have been studied previously in the statistics literature. However, we have found no references.

Theorem 7 (Correlated Arms, Strongly Exchangeable Process). *Let $\succcurlyeq$ admit an* SEE *representation with beliefs $\{\mu_{h,a}\}_{h\in\mathcal{H},a\in\mathcal{A}}$. Then $\succcurlyeq_h$ satisfies* Sym *if and only if $\big\{\{\mu_{h,a}\}_{h\in\mathcal{H}}\big\}_{a\in\mathcal{A}}$ is consistent with a unique strongly exchangeable process.*

Although we do not have direct access to the DM's conditional beliefs, (i.e., the belief, after history h regarding action a, *conditional* on the outcome of arm b: $\zeta(s_{n,a} = x|h, s_{n,b_1})$), we can nonetheless reconstruct them by assuming conditional independence. Theorem 7 states that this is without loss of generality as it can always be done, and Theorem 6 states that this fully captures the information available, and is unique.

7 Further Discussion

7.1 Related Literature

Within decision theory, the literature on learning broadly considers how a DM incorporates new information, generally via notions of Bayesianism and Exchangeability, and often in the domain of uncertainty: see Epstein and Le Breton (1993); Epstein and Seo (2010); Klibanoff et al. (2013); Lehrer and Teper (2015). Recently, there has been an interest in subjective learning, or, the identification of the set of possible "signals" that the DM believes she might observe. At it's most simple, this is the elicitation of the set of potential tastes (often referred to as subjective states) the decision maker anticipates, accomplished by examining the DM's preference over *menus* of choice objects: see Kreps (1979); Dekel et al. (2001). By also incorporating consumption goods that contract on an objective state space, the modeler can interpret the DM's preference for flexibility as directly stemming from her anticipation of acquiring information regarding the likelihood of states, as in Dillenberger et al. (2014); Krishna and Sadowski (2014).

There is also a small but highly relevant literature working on the identification of responsive learning. Hyogo (2007) considers a two-period model, with an objective state space, in which the DM ranks action-menu pairs. The action is taken in the first period and provides information regarding the likelihood of states, after the revelation of which, the DM choose a state-contingent act from the menu. The identification of interest is the DM's subjective interpretation of actions as signals. Similarly, Cooke (2016) entertains a similar model without the need for an objective state-space, and in which the consumption of a single object in the first period plays the role of a fully informative action. Cooke, therefore, identifies both the state-space and the corresponding signal structure. Piermont et al. (2015) consider a recursive and infinite horizon version of Kreps' model, where the DM deterministically learns about her preference regarding objects she has previously consumed. Dillenberger et al. (2015) consider a different infinite horizon model where the DM makes separate choices in each period regarding her information structure and current period consumption. It is worth pointing out, all of these models, unlike the this paper, capitalize on the "preference for flexibility" paradigm to characterize learning. We are able to identify subjective learning without appealing to the menu structure because of the purely responsive aspect of our model. In other words, flexibility is "built in" to our setup, as a different action can be taken after every possible realization of the signal (action).

7.2 Subjective Learning with Endogenous and Exogenous Information

As witnessed the literature covered above, there seems to be a divide in the literature regarding subjective learning. In one camp, are models that elicit the DM's perception of exogenous flows of information (as a canonical example, take Dillenberger et al. (2014)), and in the other are models that assume information is acquired only via actions taken by the DM (where this paper lies). Realistically, neither of these information structures capture the full gamut of information transmission in economic environments.

Consider the following example within the setup of the current paper. A firm is choosing between two projects (actions), a and b. Assume that each project has a high-type and a low type. The firm believes (after observing h) the probability that each project is the high-type is $\mu_{h,a}$ and $\mu_{h,b}$, respectively. By experimenting between a and b the firm's beliefs and preferences will evolve.

But, what happens if the firm anticipates the release of a comprehensive report regarding project a just before period 1? This report will declare project a high quality with probability $\alpha^h > \frac{1}{2}$ if the projects true type is high and with probability $\alpha^l < \frac{1}{2}$ if it is low. Hence, the report is an informative signal. Now, if the firms belief after observing h in period 0 is given by $[\mu_{h,a}, \mu_{h,b}]$ then, according to Bayes rule, the firms belief regarding project a being the high-type, at the beginning of period 1 will be $\mu^+_{h,a} = \frac{\alpha^h \cdot \mu_{h,a}}{\alpha^h \cdot \mu_{h,a} + \alpha^l(1-\mu_{h,a})}$, if the report is positive, and $\mu^-_{h,a} = \frac{(1-\alpha^h)\cdot \mu_{h,a}}{(1-\alpha^h)\cdot \mu_{h,a} + (1-\alpha^l)\cdot(1-\mu_{h,a})}$ if the report is negative.

Unfortunately, however, the ex-ante elicitation of preferences in our domain cannot capture the anticipation of information. The firm is ranking PoAs according to its aggregated belief from the ex-ante perspective, and thus, so as to maximize its expected belief:

$$\big(\alpha^h \mu_{h,a} + \alpha^l(1-\mu_{h,a})\big)\mu^+_{h,a} + \big((1-\alpha^h)\mu_{h,a} + (1-\alpha^l)(1-\mu_{h,a})\big)\mu^-_{h,a} = \mu_{h,a}.$$

Because of the Bayesian structure, the DM's beliefs must form a martingale, so her expectation of her anticipated beliefs are exactly her ex-ante beliefs. This fact, coupled with the linearity of expected utility, imply that the DM's ex-ante preference over PoAs is unaffected by her anticipation of exogenous information arrival.

All hope is not lost, however, of fully characterizing the DM's subjective information structure. The approach of Dillenberger et al. is orthogonal to our's, leading us to conjecture that the two models can co-exist and impart a clean separation between exogenous and endogenous information flows. Going back to the example, imagine there are two PoAs, p and q such that p is preferred to q under beliefs μ^+_h, and q to p under μ^-_h. The DM would therefore strictly desire flexibility after period 0, even after she is able to condition her decision on h. Of course, because the report is released after period 0, irrespective of the action taken by the DM, for any 0-period history h', there must exist some other PoAs, p' and q', for which flexibility is strictly beneficial (after h').

7.3 A Comment on Bayesianism in Environments of Experimentation

The results in Sections 5.2 and 6 have two related implications to Bayesianism in general models of experimentation. First, it is well known that the beliefs of two Bayesians observing the same sequence of signals will converge in the limit. Our results imply that in a setup of experimentation, different Bayesians obtaining the same information, might still hold different views of the world in the limit. Their beliefs over the uncertainty underlying each action will be identical, but they can hold different beliefs over the joint distribution.

The second point has to do with the possible equivalence with non-Bayesian DMs. Proposition 4 states that Sym is necessary and sufficient for an SEE belief system to be consistent with some exchangeable process. As discussed in the Introduction, Sym projected to stochastic processes is weaker than the standard symmetry axiom applied in the literature, because the standard assumption requires that histories fully specify the evolution of the state, while in our setup, histories can only specify cylinders. Because Sym is a weaker assumption, de Finetti's theorem implies that processes satisfying such an assumption need not be exchangeable and have a Bayesian representation as in Eq. (3).

Consider the following example of a stochastic process. In every period two coins are flipped. In odd periods the coins are perfectly correlated (with equal probability on HH and TT), and in even periods the coins are identical and independent (and both have equal probability on H and T). This process satisfies Sym, but is clearly not exchangeable. Nevertheless, Theorem 6 guarantees that there is a (unique) strongly-exchangeable process that is SEE-equivalent to this process. In this case it is easy to see that that process would be the one in which every period we toss two coins that are identical and independent (and both have equal probability on H and T).

A Lemmas.

Lemma 1. *If $\succcurlyeq_h$ satisfies vNM and IT, then $\succcurlyeq_h$ satisfies weak consequentialism:*

A9. (CSQ). *For all $a \in \mathcal{A}$ and $f, f', g, g' : X \to P$, such that, for all $x \in X$, either (i) $f(x) = f'(x)$ and $g(x) = g'(x)$ or (ii) $f(x) = g(x)$ and $f'(x) = g'(x)$. Then,*

$$(a, f) \succcurlyeq_h (a, g) \iff (a, f') \succcurlyeq_h (a, g').$$

Proof. Assume this was not true and, without loss of generality, that $(a, f) \succcurlyeq_h (a, g)$ but $(a, g') \succ_h (a, f')$. Now notice, when mixtures are taken point-wise, $\frac{1}{2}f + \frac{1}{2}g' = \frac{1}{2}g + \frac{1}{2}f'$. Therefore,

$$\begin{aligned}\big(\frac{1}{2}(a, f) + \frac{1}{2}(a, g')\big) &\succ_h \big(\frac{1}{2}(a, g) + \frac{1}{2}(a, f')\big) \\ &\sim_h (a, \frac{1}{2}g + \frac{1}{2}f') = (a, \frac{1}{2}f + \frac{1}{2}g') \\ &\sim_h \big(\frac{1}{2}(a, f) + \frac{1}{2}(a, g')\big),\end{aligned}$$

where the first line follows from vNM, and the indifference conditions from IT. This is a contradiction. ■

Lemma 2. *If $\succcurlyeq_h$ satisfies vNM and IT for all $h \in \mathcal{H}$, then, if $h \overset{\mathcal{A}}{\sim} h'$ then $\succcurlyeq_h = \succcurlyeq_{h'}$.*

Proof. We will show the claim on induction by the length of the history. So let $h, h' \in \mathcal{H}(1)$ such that $h \overset{\mathcal{A}}{\sim} h'$. Therefore, $h = (p, (a, f), x)$ and $h' = (p', (a, g), x)$. Notice, by definition we have, $p = \alpha(a, f) + (1 - \alpha)r$ and $p' = \alpha'(a, g) + (1 - \alpha')r'$, for some $\alpha, \alpha' \in (0, 1]$ and $r, r' \in P$.

Let $q, q' \in P$; we want to show that $q \succcurlyeq_h q' \iff q \succcurlyeq_{h'} q'$. So let $q \succcurlyeq_h q'$, or by definition, $p_{-h}q \succcurlyeq p_{-h}q'$, which by the above observation is equivalent to

$$\alpha(a, f)_{-((a,f),(a,f),x)}q + (1 - \alpha)r \succcurlyeq \alpha(a, f)_{-((a,f),(a,f),x)}q + (1 - \alpha)r.$$

By independence (i.e., vNM) this is if and only if $(a, f)_{-((a,f),(a,f),x)}q \succcurlyeq (a, f)_{-((a,f),(a,f),x)}q'$, which by CSQ is if and only if $(a, g)_{-((a,g),(a,g),x)}q \succcurlyeq (a, g)_{-((a,g),(a,g),x)}q'$. Using independence again, this is if and only if $p'_{-h'}q \succcurlyeq p'_{-h'}q'$. This completes the base case.

So assume the claim holds for all histories of length n. So let $h, h' \in \mathcal{H}(n+1)$ such that $h \overset{\mathcal{A}}{\sim} h'$. Therefore, $h = (h_n, p, (a, f), x)$ and $h' = (h'_n, p', (a, g), x)$, for some $h_n, h'_n \in H(n)$ such that $h_n \overset{\mathcal{A}}{\sim} h'_n$. By the inductive hypothesis $\succcurlyeq_{h_n} = \succcurlyeq_{h'_n}$.

Let $q, q' \in P$, and $q \succcurlyeq_h q'$, or by definition, $p_{-(p,(a,f),x)}q \succcurlyeq_{h_n} p_{-(p,(a,f),x)}q'$. By independence and weak consequentialism this is if and only if $(a,g)_{-((a,g),(a,g),x)}q \succcurlyeq_{h_n} (a,g)_{-((a,g),(a,g),x)}q'$, which by independence again (and the equivalence of $\succcurlyeq_{h_n}$ and $\succcurlyeq_{h'_n}$), is if and only if $p'_{-(p',(a,g),x)}q \succcurlyeq_{h'_n} p'_{-(p',(a,g),x)}q'$. ■

B Proof Of Main Theorems

B.1 (Decision Theoretic) Representation Results

Proof of Theorem 2.

[Step 0: Value Function.] Since $\succcurlyeq_h$ satisfies vNM, there exists a $v_h : \hat{P} \to \mathbb{R}$ such that

$$U_h(p) = \mathbb{E}_p\left[v_h(a,f)\right] \tag{6}$$

represents $\succcurlyeq_h$, with v_h unique un to affine translations. □

[Step 1: Recursive structure.] To obtain the skeleton of the representation, lets consider $\hat{\succcurlyeq}$, the restriction of $\succcurlyeq$ to Σ (i.e., using the natural association between streams of lotteries and degenerate trees). The relation $\hat{\succcurlyeq}$ satisfies vNM (it is continuous by the closure of Σ in P). Hence there is a linear and continuous representation: i.e., an index $\hat{u} : X \times \Sigma \to \mathbb{R}$ such that:

$$\hat{U}(\sigma) = \mathbb{E}_\sigma\left[\hat{u}(x,\rho)\right] \tag{7}$$

unique upto affine translations.

Following Gul and Pesendorfer (2004), (henceforth GP), fix some $(x',\rho') \in \Sigma$. From SEP we have $\hat{U}(\frac{1}{2}(x,\rho)+\frac{1}{2}(x',\rho')) = \hat{U}(\frac{1}{2}(x,\rho')+\frac{1}{2}(x',\rho))$, and hence, $\hat{u}(x,\rho) = \hat{u}(x,\rho')+\hat{u}(x',\rho)-\hat{u}(x',\rho')$. Then setting $u(x) = \hat{u}(x,\rho') - \hat{u}(x',\rho')$ and $W(\rho) = \hat{u}(x',\rho)$, we have,

$$\hat{U}(\sigma) = \mathbb{E}_\sigma\left[u(x) + W(\rho)\right] \tag{8}$$

Now, consider $p' = (x',\rho)$. Notice that p' has unique 1-period history: $h = (p',p',x')$. By NT, h cannot be null. So, by SST, $\hat{\succcurlyeq}_h = \hat{\succcurlyeq}$. This implies, of course that $W = \delta U + \beta$ for some $\delta > 0$ and $\beta \in \mathbb{R}$. Following Step 3 of Lemma 9 in GP exactly, we see that $\delta < 1$ and without loss of generality we can set $\beta = 0$:

$$\hat{U}(\sigma) = \mathbb{E}_\sigma\left[u(x) + \delta\hat{U}(\rho)\right] \tag{9}$$

Both representing $\hat{\succcurlyeq}$ and being unique up to affine translations, we can normalize each U_h to coincide with $\hat{U}$ over Σ. □

[Step 2: The existence of subjective probabilities.] For each $a \in \mathcal{A}$ consider

$$\mathcal{F}(a) = a \otimes \Sigma$$

i.e., the elements of $\hat{P}$ that begin with action a and from period 2 onwards are in Σ. Associate $\mathcal{F}(a)$ with the set of "acts": $f : S_a \to \Sigma$, in the natural way. For any acts f, g let $f_{-x}g$ denote the act that coincides with f for all $x' \in S_a$, $x' \neq x$, and coincides with g after x. For each $h \in \mathcal{H}$, and acts $f, g \in \mathcal{F}(a)$, say $f \dot{\succcurlyeq}_{h,a} g$ if and only if $(a,f) \succcurlyeq_h (a,g)$.

It is immediate that $\dot{\succcurlyeq}_{h,a}$ is a continuous weak order (where, as before, continuity follows from the closure of $\mathcal{F}$ in P). Further, $\dot{\succcurlyeq}_{h,a}$ satisfies independence. Indeed: fix $f, g, h \in \mathcal{F}(a)$ with $f \dot{\succcurlyeq}_{h,a} g$. Then

$$\begin{aligned} f \dot{\succcurlyeq}_{h,a} g &\implies (a, f) \succcurlyeq_h (a, g) \\ &\implies \alpha(a, f) + (1-\alpha)(a, h) \succcurlyeq_h \alpha(a, g) + (1-\alpha)(a, h) \\ &\implies (a, \alpha f + (1-\alpha)h) \succcurlyeq_h (a, \alpha g + (1-\alpha)h) \\ &\implies \alpha f + (1-\alpha)h \dot{\succcurlyeq}_{h,a} \alpha g + (1-\alpha)h, \end{aligned}$$

where the third line uses IT. Lastly, $\dot{\succcurlyeq}_{h,a}$ satisfies monotonicity, a direct consequence of SST. Hence, we have state-independence which gives us the full set of Anscombe and Aumann axioms for an SEU representation of $\dot{\succcurlyeq}_{h,a}$ with state space S_a. That is, a belief $\mu_{h,a} \in \Delta(S_a)$ and a utility index from $\Sigma \to \mathbb{R}$ (which is of course, $\hat{U}$, and so will be denoted as such), such that

$$\hat{V}_{h,a}(f) = \mathbb{E}_{\mu_{h,a}}\left[\hat{U}(f(x))\right] \tag{10}$$

represents $\dot{\succcurlyeq}_{h,a}$. □

[Step 3: Proportional Actions.] Now, fix some $h \in \mathcal{H}$ and consider an arbitrary $(a, f) \in \hat{P}$. Let $\rho \in \Sigma$ be such that $\text{marg}_X \rho = \mu_{h,a}$. We claim, (a, f) and ρ are h-proportional. Fix some $g, g' : X \to \Sigma$. From (10), we know

$$(a, g) \succcurlyeq (a, g') \iff \mathbb{E}_{\mu_{h,a}}\left[\hat{U}(g(x))\right] \geqslant \mathbb{E}_{\mu_{h,a}}\left[\hat{U}(g'(x))\right] \tag{11}$$

From (9) we have

$$\begin{aligned} \hat{U}(\rho.g) &= \mathbb{E}_\rho\left[u(x) + \delta\hat{U}(g(x))\right] \\ &= \mathbb{E}_{\text{marg}_X \rho}\left[u(x) + \delta\hat{U}(g(x))\right] \\ &= \mathbb{E}_{\mu_{h,a}}\left[u(x)\right] + \delta\, \mathbb{E}_{\mu_{h,a}}\left[\hat{U}(g(x))\right] \end{aligned}$$

In corresponding fashion we obtain the analogous representation for $\hat{U}(\rho.g')$, and hence

$$\rho.g \succcurlyeq \rho.g' \iff \mathbb{E}_{\mu_{h,a}}\left[\hat{U}(g(x))\right] \geqslant \mathbb{E}_{\mu_{h,a}}\left[\hat{U}(g'(x))\right] \tag{12}$$

Combining the implications of (11) and (12), we see that (a, f) and ρ are h-proportional. □

[Step 4: Proportional Plans.] We now claim that for any $h \in \mathcal{H}$ and $p \in P$ there exists some $\sigma \in \Sigma$ such that $p \sim_h \sigma$. Fix some $p \in P$, and for each $n \in \mathbb{N}$ define p^n to be any PoA that agrees with p on the first n periods, then provides elements of Σ unambiguously. Note that $p_n \to p$ point-wise and hence converges in the product topology. Therefore, the claim reduces to finding a convergent sequence $\{\sigma_n\}_{n \in \mathbb{N}} \subset \Sigma$ such that $\sigma^n \sim_h p^n$, as continuity ensures the limits are indifferent.

We will prove the subsidiary claim by induction. Consider p^1, for each $(a, f) \in supp[p^1]$, note, by assumption, $f : X \to \Sigma$. Let $\tau^{1,(a,f)} \in \Sigma$ be such that $\text{marg}_X \tau^{1,(a,f)} = \mu_{h,a}$. By [Step 3], (a, f) and $\tau^{1,(a,f)}$ are h-proportional. And thus, $\tau^{1,(a,f)}.f \sim_h (a, f).f = (a, f)$, by PRP. Let $\sigma^1 \in \Sigma$ be such that $\sigma^1[E] = p^1[\{(a, f) | \tau^{1,(a,f)}.f \in E\}]$. Therefore,

$$\begin{aligned} U_h(p^1) &= \mathbb{E}_{p^1}\left[v_h(a, f)\right] \\ &= \mathbb{E}_{p^1}\left[\hat{U}(\tau^{1,(a,f)}.f)\right] \\ &= \mathbb{E}_{\sigma^1}\left[\hat{U}(\rho)\right] \\ &= \hat{U}(\sigma^1) \end{aligned}$$

where the third line comes from the change of variables formula for pushforward measures. This completes the base case.

Now, assume the claim hold for all h and $m \leqslant n-1$ for some $n \in \mathbb{N}$. Consider p^n. Note that for all h' of the form $h(x) = (h, p^n, (a, f), x)$, the implied continuation problem $p^n(h')$ satisfies the inductive hypothesis. Therefore, there exists a $\sigma^{n-1,h'} \sim_{h'} p(h')$ for all such h'.

Let $\star$ denote the mapping: $(a, f) \mapsto (a, f)^\star = (a, x \mapsto \sigma^{n-1,h(a,x)})$, where $h(a, x) = (h, p^n, (a, f), x)$. By construction, for each (a, f) in $supp(p^n)$, and $x \in S_a$ we have $(a, f) \sim_h (a, f_{-x}\sigma^{n-1,h(a,x)})$ (using the notation from [Step 2]). Employing CSQ we have $(a, f) \sim_h (a, f)^\star$ (i.e., enumerating the outcomes in S_a and changing f one entry at a time, where CSQ ensures that each iteration is indifferent to the last).

Let $\hat{p}^n \in P$ be such that $\hat{p}^n[E] = p^n[\{(a, f)|(a, f)^\star \in E\}]$. So,

$$\begin{aligned} U_h(p^n) &= \mathbb{E}_{p^n}\left[v_h(a, f)\right] \\ &= \mathbb{E}_{p^n}\left[v_h((a, f)^\star)\right] \\ &= \mathbb{E}_{\hat{p}^n}\left[v_h(b, g)\right] \\ &= U_h(\hat{p}^n) \end{aligned}$$

Applying the base case to $\hat{p}^n$ concludes the inductive step. Notice also, the convergence of $\{\sigma^n\}_{n\in\mathbb{N}}$ is easily verified, following the fact that the marginals on P_n are fixed for any σ^m with $m \geqslant n$. ▫

[Step 5: Representation.] Consider any $(a, f) \in \hat{P}$. We claim that there exists an $(a, f') \in \mathcal{F}(a)$ such that $(a, f) \sim_h (a, f')$. Indeed, by [Step 4], $x \in S_a$, there exists some $\rho(a, x)$ such that $\rho(a, x) \sim_{h(a,x)} f(x)$, where $h(a, x) = (h, (a, f), (a, f), x)$. Define $f' \in \mathcal{F}(a)$ as $x \mapsto \rho(a, x)$. It follows from CSQ that $(a, f) \sim_h (a, f')$.

We know by [Step 3] that there exists a $\rho \in \Sigma$, h-proportional to (a, f), with $\text{marg}_X \rho = \mu_{h,a}$. Hence $(a, g) = (a, f).g \sim_h \rho.g$ for all $g : X \to \Sigma$. We have,

$$\begin{aligned} v_h(a, g) &= \hat{U}(\rho.g) \\ &= \mathbb{E}_{\mu_{h,a}}\left[u(x) + \delta\hat{U}(g(x))\right], \end{aligned}$$

and so, for (a, f'):

$$v_h(a, f') = \mathbb{E}_{\mu_{h,a}}\left[u(x) + \delta\hat{U}(\rho(a, x))\right].$$

By the indifference condition $\rho(a, x) \sim_{h(a,x)} f(x)$,

$$v_h(a, f) = \mathbb{E}_{\mu_{h,a}}\left[u(x) + \delta U_{h(x)}(f(x))\right]. \tag{13}$$

Notice, $h(a, x) \overset{\mathcal{A}}{\sim} h'(a, x) = (h, p, (a, f), x)$, so by Lemma 2, $\succcurlyeq_{h(a,x)} = \succcurlyeq_{h'(a,x)}$. Applying this fact, and plugging (13) into (6) provides

$$U_h(p) = \mathbb{E}_p\left[\mathbb{E}_{\mu_{h,a}}\left[u(x) + \delta U_{h'(a,x)}(f(x))\right]\right] \tag{14}$$

as desired. ▫

■

Proof of Theorem 3.

That IA is equivalent to $\mu_{h,a}$ depending only on $a(h)$ follows from [Step 2] of the proof of Theorem 2.

Indeed, if $h, h' \in \mathcal{H}$ are such that $a(h) = a(h')$, then by IA the must induce the same ordering over $a \otimes \Sigma$: $\dot{\succcurlyeq}_{h,a} = \dot{\succcurlyeq}_{h',a}$, and hence the same beliefs.

So, for each, $a \in \mathcal{A}$, let $\zeta^a \in \Delta(S_a^{\mathbb{N}})$ (where $S_a^{\mathbb{N}}$ is endowed with the σ-algebra generated by all elements of finite length) denote the unique countably additive measure consistent with each $\mu_{a(h)}$. We will show WA-Sym implies that $\zeta^a(x_1, \ldots x_n) = \zeta^a(x_{\pi_1}, \ldots x_{\pi_n})$, for all $(x_1, \ldots x_n) \in X^n$ and permutation π. From this, we can invoke classic result of Hewitt and Savage (1955), which states that there exists a unique probability measure, ψ^a over $\Delta(X)$, such that

$$\zeta^a(E) = \int_{\Delta^{\mathcal{B}}(X)} \hat{D}(E) d\theta^a(D).$$

Towards this subsidiary claim, fix some $(x_1, \ldots x_n) \in X^n$ and permutation π. Now, let $p \in P$ denote any PoA such that a is unambiguously assigned for the first n periods. Then there exists unique $h, h' \in \mathcal{H}(p, n)$ that correspond to $(x_1, \ldots x_n)$ and $(x_{\pi_1}, \ldots x_{\pi_n})$, respectively. Let $\alpha = \zeta^a(x_1, \ldots x_n) = \mu_{\varnothing,a}(x_1)\mu_{x_1,a}(x_2) \ldots \mu_{x_1,\ldots x_{n-1},a}(x_n)$. Let $\rho, \rho' \in \Sigma$ be such that $U_h(\rho') = 1$ and $U_h(\rho) = 0$. Then, by (SEE) we have

$$p_{-n}(\alpha\rho + (1-\alpha)\rho') \sim (p_{-n}\rho)_{-h}\rho'$$

so, by WA-Sym, we have,

$$p_{-n}(\alpha\rho + (1-\alpha)\rho') \sim (p_{-n}\rho)_{-h'}\rho'$$

which implies, again by (SEE), $\alpha = \mu_{\varnothing,a}(x_{\pi_1}) \ldots \mu_{x_{\pi_1},\ldots x_{\pi_{n-1}},a}(x_{\pi_n}) = \zeta^a(x_{\pi_1}, \ldots x_{\pi_n})$. ■

Proof of Theorem 7.

Let $\zeta \in \Delta(\mathcal{S})$ denote any process consistent with each $\mu_{h,a}$. I.e., $\frac{\zeta(h,a,x)}{\zeta(h)} = \mu_{h,a}(x)$. The result follows (give Theorems 6 and 5) if for any permutation π, $\zeta(h) = \zeta(\pi h)$. The proof of this is almost identical to the similar claim in the proof of Theorem 3, and so, omitted.

■

B.2 Strong Exchangeability

Proof of Theorem 5.

First we show, if a strongly exchangeable process ζ over $\mathcal{S}$ is induced by an i.i.d distribution D over $\mathcal{S}_{\mathcal{A}}$, then it must be that the marginals of D (on $\{S_a\}_{a \in \mathcal{A}}$) are independent, that is $D \in \Delta^{IN}$. Indeed, consider two non-empty, disjoint collection of actions, $\hat{\mathcal{A}}, \hat{\mathcal{A}}' \subset \mathcal{A}$. Let $E, F \in S_{\hat{\mathcal{A}}}$, $E', F' \in S_{\hat{\mathcal{A}}'}$, be measurable events. Identify E^n with the cylinder it E generates in $\mathcal{S}$ when in the n^{th} coordinate: $E^n = \{s \in \mathcal{S} | s_{n,\mathcal{B}} \in E\}$. Since ζ is strongly exchangeable we have that

$$\zeta\left(E^n \cap E'^n \cap F^{n+1} \cap F'^{n+1}\right) = \zeta\left(E^n \cap F'^n \cap F^{n+1} \cap E'^{n+1}\right). \tag{2Sym}$$

We will refer to the latter weaker property as *two symmetry.* Now, since ζ is i.i.d generated by D, we have that (abusing notation by identifying E with the cylinder it generates in $S_{\mathcal{A}}$)

$$D(E \cap E') \cdot D(F \cap F') = D(E \cap F') \cdot D(F \cap E').$$

Substituting via the rule of conditional probability:

$$D(E|E') \cdot D(E') \cdot D(F|F') \cdot D(F') = D(E|F') \cdot D(F') \cdot D(F|E') \cdot D(E').$$

This implies that

$$\frac{D(E|E')}{D(E|F')} = \frac{D(F|E')}{D(F|F')}.$$

Since this is true for all events, we have that $D(E|E') = D(E|F')$ for every $E \in S_{\hat{\mathcal{A}}}$ and $E', F' \in S_{\hat{\mathcal{A}}'}$, implying $\hat{\mathcal{A}}$ and $\hat{\mathcal{A}}'$ are independent.

We now move to show that strong exchangeability is sufficient for the representation specified in the statement of the result. Since strong exchangeability implies exchangeability, we can apply de Finetti's theorem and represent the process ζ by

$$\zeta(\cdot) = \int_{\Delta(S_{\mathcal{A}})} \hat{D}(\cdot) d\psi(D).$$

We need to show that ψ's support lies in Δ^{IN}.

For $s \in \mathcal{S}$ and $t \in \mathbb{N}$ let s_t be the projection of s into the first t periods. Now, let $\zeta(\cdot|s_t) : S_{\mathcal{A}} \to [0,1]$ be the *one period ahead predictive probability*, given that the history of realizations in the first t periods is s_t. Since ζ is exchangeable, $\zeta(\cdot|s_t)$ converges (as $t \to \infty$) with ζ probability 1. Moreover, the set of all limits is the support of ψ. Denote the limit for a particular s by D_s. Of course, the exchangeability of ζ also guarantees that $\zeta(\cdot,\cdot|s_t) : S_{\mathcal{A}} \times S_{\mathcal{A}} \to [0,1]$, that is the *two period ahead predictive probability*, converges to $D_s \times D_s$. Furthermore, ζ is strongly exchangeable; the limit itself satisfies (2Sym), and the arguments above imply that $D_s \in \Delta^{IN}$ with ζ probability 1. ■

Proof of Theorem 6.

We will construct a pre-measure $\hat{\zeta}$ by dictating its measure on each cylinder. Set $\zeta(\emptyset) = 0$ and $\hat{\zeta}(\mathcal{S}) = 1$. Let $E \neq \mathcal{S}$ be an arbitrary cylinder, i.e., $E = \prod_{n\in\mathbb{N}} \prod_{a\in\mathcal{A}} E_{n,a}$, such that for only finitely many (n,a), is $E_{n,a} \neq S_a$. Clearly, there are a finite number of $a \in \mathcal{A}$ such that $E_{k,a} \neq S_a$ for any k. Order these $a_1 \dots a_n$. For each a_i let m_i denote the number of components such that $E_{k,a_i} \neq S_{s_i}$, and for $j = 1 \dots m_i$, let $k_{i,j}$ denote the j^{th} such component. Finally, for each a_i, let π_{a_i} denote any permutation such that $\pi_{a_i}(k_{i,j}) = j + \sum_{i'<i} m_{i'}$ (notice that the resulting event is the same for any such permutation). Consider $\hat{E} = \prod_{n\in\mathbb{N}} \prod_{a\in\mathcal{A}} E_{\pi_a(n),a}$, where $\pi_a = \pi_{a_i}$ if $a \in a_1 \dots a_n$ and the identity otherwise. That is, for $m \in 1 \dots m(1)$, $\hat{E}_{m,a} = S_a$ for all a except a_1, for $m \in m(1) \dots m(1) + m(2)$, $\hat{E}_{m,a} = S_a$ for all a except a_2, etc. Therefore, $E \subset \mathcal{H}$. Set $\hat{\zeta}(E) = \xi(\hat{E})$. For the remainder of this proof, for any cylinder E, $\hat{E}$ denotes the corresponding cylinder (and subset of $\mathcal{H}$), generated by the above process, fixing any ordering over $\mathcal{A}$.

We need to show that $\hat{\zeta}$ is countably additive. Towards this, assume that E, E' are disjoint cylinders such that $E \cup E'$ is a cylinder. Then it must be that there exists a unique (n,a) such that $E_{n,a} \cap E'_{n,a} = \emptyset$ and for all other (m,b), $E_{m,b} = E'_{m,b}$. Indeed, if this was not the case, then there exists some (m,b) and some x such that (WLOG) $x \in E_{m,b} \backslash E'_{m,b}$. But then, for all $s \in E \cup E'$, $s_{m,b} = x \implies s_{n,a} \in E_{n,a} \neq (E \cup E')_{n,a}$ a contradiction to $E \cup E$ being a cylinder. But this implies $\hat{E}$ and $\hat{E}'$ are disjoint subsets of $\mathcal{H}$ that differ on on a single coordinate, and therefore, that $\hat{E} \cup \hat{E}' = \widehat{(E \cup E')}$. Hence, $\hat{\zeta}$ is finitely additive, since ξ is.

Since $\hat{\zeta}$ is finitely additive, countable additivity follows if we show that for all decreasing sequences of cylinders $\{E^k\}_{k\in\mathbb{N}}$, such that $\inf_k \hat{\zeta}(E^k) = \epsilon > 0$, we have $\bigcap_{k\in\mathbb{N}} E^k \neq \emptyset$. But this follows immediately from the finiteness of S_a. Since $E^{k+1} \subseteq E^k$, it must be that $E^k_{n,a} \subseteq E^k_{n,a}$. But each $E^k_{n,a}$ is finite, hence compact, and nonempty, because $\zeta(E^k) \geqslant \epsilon$. Therefore $\bigcap_{k\in\mathbb{N}} E^k_{n,a} \neq \emptyset$. The result follows by noting that the intersection of cylinder sets is the cylinder generated by the intersection of the respective generating sets. Let ζ denote the unique extension of $\hat{\zeta}$ to the σ-algebra on $\mathcal{S}$.

We need to show that ζ is strongly exchangeable. But this is immediate from the condition that $\xi(h) = \xi(\pi(h))$ for any $h \in \mathcal{H}(n)$ and n-permutation π. Indeed, if E^π is any permutation of E, then it is clear that $\hat{E}^\pi$ must be a permutation of $\hat{E}$. Likewise, we can show that ζ is *SEE-equivalent* to ξ. For any history $h \in \mathcal{H}$, $\hat{h}$ is a permutation of h. Finally, the same logic show that ζ is unique. Towards a contradiction, assume there was some distinct, strongly-exchangeable and SEE-equivalent ζ'. Then there must be some cylinder such that $\zeta(E) \neq \zeta'(E)$. But, by strong exchangeability, $\zeta(\hat{E}) = \zeta(E)$ and $\zeta'(\hat{E}) = \zeta'(E)$, so $\zeta(\hat{E}) \neq \zeta'(\hat{E})$ –a contradiction to their SEE-equivalence. ■

C On the construction of Plans.

C.1 Generalized Plans

We will begin by constructing a more general notion of Plans (reminiscent of IHCPs, first constructed in Gul and Pesendorfer (2004), and then refine our notion to capture only the elements of interest. This methodology serves two purposes. First, the more general approach allows us to use standard techniques for the construction of infinite horizon choice objects. Second, generalized plans may be of direct interest in future work, when, for example, denumerable support is not desirable. To begin, let $Q_0 = \Delta^{\mathcal{B}}(\mathcal{A})$ and, for define recursively for each $n \geqslant 1$

$$Q_n = \Delta^{\mathcal{B}}(\mathcal{A} \times \mathcal{K}(X \times Q_{n-1})).$$

Finally, define $Q^* = \prod_{n\geqslant 0} Q_n$. Q^* is the set of generalized plans.

C.2 Consistency

For the first step, we will follow closely GP of consistent IHCPs, but with enough difference that it makes sense to define things outright. Formally, let $G_1 : \mathcal{A} \times \mathcal{K}(X \times Q_0) \to \mathcal{A}$ as the mapping $(a, \{x, q_0\}) \mapsto a$. Let $F_1 : Q_1 \to Q_0$ as the mapping $F_1 : q_1 \mapsto \big(E \mapsto q_1(G_1^{-1}(E))\big)$, for any $E \in \mathcal{B}(\mathcal{A})$. Therefore, for any $E \in \mathcal{B}(\mathcal{A})$, $F_1(p_1)(E)$ is the probability of event E in period 0 as implied by p_1; $F_1(p_1)$ is the distribution over period 0 actions implied by p_1. From here we can recursively define $G_n : \mathcal{A} \times \mathcal{K}(X \times Q_n) \to \mathcal{A} \times \mathcal{K}(X \times Q_{n-1})$ as:

$$G_n : (a, \{x, q_{n-1}\}) \mapsto (a, \{x, F_{n-1}(q_0)\}))$$

and $F_n : Q_n \to Q_{n-1}$ as:

$$F_n : q_n \mapsto \big(E \mapsto q_n(G_n^{-1}(E))\big)$$

for any E in $\Delta^{\mathcal{B}}(\mathcal{A} \times \mathcal{K}(X \times Q_{n-1}))$. A consistent generalized plan is one such that

$$F_n(q_n) = q_{n-1},$$

for all n. Let Q denote all such generalized plans.

C.3 $P \cong \Delta(\hat{P})$.

Proposition 8. *There exists a homeomorphism, $\xi : Q \to \Delta^B(\mathcal{A} \times \mathcal{K}(X \times Q))$ such that*

$$\text{marg}_{\mathcal{A}\times\mathcal{K}(X\times Q_{n-1})}(\xi(q)) = q_n. \tag{15}$$

Proof. [Step 1: Extension Theorem.] Let $C_n = \{(q_0, \ldots q_n) \in \prod_{k=0}^n Q_k | q_k = F_{k+1}(q_{k+1}), \forall k = 1 \ldots n-1\}$, and $T_n = \mathcal{K}(X \times C_n)$ for $n \geqslant 0$. Let $T^* = \prod_{n=0}^{\infty} T_n$ and $T = \{t \in T^* | (\text{proj}_{T_n} t_{n+1} = t_n\}$. Let $Y_0 = \Delta^{\mathcal{B}}(\mathcal{A})$ and for $n \geqslant 1$ let $Y_n = \Delta^{\mathcal{B}}(\mathcal{A} \times T_0 \times \ldots \times T_n))$. We say the the sequence of probability measures $\{\nu_n \in Y_n\}_{n\geqslant 0}$ is consistent if $\text{marg}_{A\ldots T_{n-1}} \nu_{n+1} = \nu_n$ for all $n \geqslant 0$. Let Y^c denote the set of all consistent sequences. Then we know by Brandenburger and Dekel (1993), for every $\{\nu_n\} \in Y^c$ there exists a unique $\nu \in \Delta^{\mathcal{B}}(\mathcal{A} \times T^*)$ such that $\text{marg}_{\mathcal{A}} \nu = \nu_0$ and $\text{marg}_{A\ldots T_n} \nu = \nu_n$. Moreover, the map $\psi : Y^c \to \Delta^{\mathcal{B}}(\mathcal{A} \times T^*)$:

$$\psi : \{\nu_n\} \mapsto \nu$$

is a homeomorphism. ▫

[Step 2: Extending Backwards.] Let $D_n = \{(t_0, \ldots t_n) \in \times_{n=0}^n T_n | t_k = \text{proj}_{T_n}(t_{k+1}), \forall k = 1 \ldots n-1\}$. Let $Y^d = \{\{\nu_n\} \in Y^c | \nu_n(\mathcal{A} \times D_n) = 1, \forall n \geqslant 0\}$. We will now show, for each $q \in Q$, there exists a unique $\{\nu_n\} \in Y^d$, such that $\nu_0 = q_0$ and $\text{marg}_{\mathcal{A}\times\mathcal{K}(X\times Q_{n-1})}(\text{marg}_{\mathcal{A}\times T_{n-1}}(\nu_n)) = q_n$ for all $n \geqslant 1$. Indeed, let m_0, m_1 be the identify function on $\mathcal{A}$ and $\mathcal{A} \times \mathcal{K}(X \times Q_0)$, respectively. Then for each $n \geqslant 2$ let $m_n : \mathcal{A} \times D_{n-1} \to \mathcal{A} \times \mathcal{K}(X \times Q_{n-1})$ as follows:

$$m_{n+1} : (a, \{x^0, q_0^0\}, \{x^1, q_0^1, q_1^1\} \ldots \{x^n, q_0^n \ldots q_n^n\}^n) \mapsto (a, \{x^n, q_n^n\}).$$

Note: for $n \geqslant 0$, each m_n is a Borel isomorphism. Indeed, continuity of m_n is obvious, and measurability follows immediately from the fact that canonical projections are measurable in the product σ-algebra. It is clear that m_n is surjective, and —since (given F_k for $k \in 1 \ldots n$) q_n uniquely determines $q_0 \ldots q_{n-1}$, which, (given the projection mappings) uniquely determines $T_0 \ldots T_{n-1}$— m_i is also injective. As for, m_n^{-1}, continuity follows from the continuity of F_k for $k \in 1 \ldots n$ and the projection mappings. Lastly, measurability of m_n^{-1} comes from the fact that a continuous injective Borel function is a Borel isomorphism (see Kechris (2012) corollary 15.2).

So, let $\psi : Q \to Y^d$ as the map

$$\phi : q \mapsto \{E_n \mapsto q_n(m_n(E_n))\}_{n\geqslant 0},$$

for any $E_n \in \mathcal{B}(A \times T_0 \times \ldots \times T_n)$. The continuity of ϕ and ϕ^{-1} follow from the fact that they are constructed from the pushforward measures of m_n^{-1} and m_n, respectively, which are themselves continuous (or, explicitly, see GP lemma 4).

Finally, let $\Gamma_n = \mathcal{A} \times D_n \times_{k=n+1}^{\infty} T_k$. Let $\nu = \psi(\{\nu_n\})$ for some $\{\nu_n\}$ in Y^d. Then $\nu(\Gamma_n) = \nu(\mathcal{A} \times D_n) = 1$. So, $\nu(\mathcal{A}\times T) = \nu(\cap_{n\geqslant 0}\Gamma_n) = \lim \nu(\Gamma_n) = 1$. Also, note, if $\nu(\mathcal{A}\times T) = 1$, then $\nu(\Gamma_n) = 1$ for all $n \geqslant 0$. Putting these together, $\nu \in Y^d$ if and only if $\nu(\mathcal{A} \times T) = 1$, in other words, $\psi(Y^d) = \{\nu \in \Delta^{\mathcal{B}}(\mathcal{A} \times T^*) | \nu(\mathcal{A} \times T) = 1\}$. ▫

[Step 3: Extending Forwards.] Let τ denote the map from $\mathcal{A} \times \mathcal{K}(X \times Q) \to \mathcal{A} \times T$ as

$$\tau : (a, \{x, q\}) \mapsto (a, (\{x, q_0\}, \{x, q_0, q_1\}, \ldots))$$

That τ it is a bijection follows from the consistency conditions on Q, T, and C_n for $n \geqslant 1$. Now takes some measurable set $E \subseteq T$. Then $\tau^{-1}(E) = \bigcap_{n\geqslant 0} \{\{x, q_0, \ldots q_n \times_{k=n_1^\infty} Q_k\} \in K(X \times Q^*)\}$, the countable intersection of measurable sets, and hence measurable. That τ and τ^{-1} are continuous is immediate. Therefore, by the same argument as in [Step 2], τ is a Borel isomorphism and $\kappa : \Delta^{\mathcal{B}}\big(\mathcal{A} \times T\big) \to \Delta^{B}\big(\mathcal{A} \times \mathcal{K}(X \times Q)\big)$,

$$\kappa : \nu \mapsto \big(E \mapsto \nu(\tau(E))\big)$$

for all E in $\Delta^{\mathcal{B}}\big(\mathcal{A} \times \mathcal{K}(X \times Q)\big)$. Clearly, $\mathrm{marg}_{\mathcal{A}}(\kappa(\nu)) = \mathrm{marg}_{\mathcal{A}}(\nu)$ and $\mathrm{marg}_{\mathcal{A} \times \mathcal{K}(X \times Q_{n-1})}(\kappa(\nu)) = \mathrm{marg}_{\mathcal{A} \times \mathcal{K}(X \times Q_{n-1})}(\mathrm{marg}_{\mathcal{A} \times T_{n-1}}(\nu))$ for all $n \geqslant 1$.

□

Behold, $\lambda = \kappa \circ \psi \circ \phi$ is the desired homeomorphism. ■

C.4 Refining Generalized Plans

Definition. *Let $R_0 = Q_0$ and $R_1 = \{r_1 \in Q_1 | r_1(\mathcal{A} \otimes R_0) = 1\}$. Then, recursively let $R_n = \{r_n \in Q_n | r_n(\mathcal{A} \otimes R_{n-1}) = 1\}$. Set $R = \prod_{n=0}^{\infty} R_n$.*

Corollary 8.1. *λ is a homeomorphism between R and $\Delta^{\mathcal{B}}(\mathcal{A} \otimes R)$.*

Proof. Identify $\Delta^{\mathcal{B}}(\mathcal{A} \otimes R)$ with $\big\{\nu \in \Delta^{\mathcal{B}}(\mathcal{A} \times K(X \times Q)) | \nu(\mathcal{A} \otimes R) = 1\big\}$. Let $r \in R$. For each $n \geqslant 0$ let $\Gamma_n^r = \{(a, \{x, q\}) \in \mathcal{A} \otimes Q | q_k \in R_k,\ k = 0 \ldots n\}$. Then $\lambda(r)(\Gamma_n^r) = \mathrm{marg}_{A \times K(X \times Q_n)}(\lambda(r))(\mathcal{A} \otimes R_n) = r_{n+1}(\mathcal{A} \otimes R_n) = 1$ for all $n \geqslant 1$. So $\lambda(r)(\mathcal{A} \otimes R) = \lambda(r)(\cap_{n \geqslant 0} \Gamma_n^r) = \lim \lambda(r)(\Gamma_n^r) = 1$. Now, fix $q \in Q$ with $\lambda(q)(\mathcal{A} \otimes R) = 1$, then $q_n(\mathcal{A} \otimes R_{n-1}) = \mathrm{marg}_{A \times K(X \times Q_{n-1})}(\lambda(q))(\mathcal{A} \otimes R_{n-1}) = \lambda(r)(\Gamma_n^r) \geqslant \lambda(r)(\mathcal{A} \otimes R) = 1$ for all $n \geqslant 0$ and so $q \in R$. ■

Definition. *For a metric space, M, let $\Delta(M) \subseteq \Delta^{\mathcal{B}}(M)$ denote the set of all distributions with countable support. I.e., for all $\nu \in \Delta(M)$, there exists a countable set S_ν such that $m \notin S_\nu \implies \nu(m) = 0$, and $\sum_{m \in S_\nu} \nu(m) = 1$.*

Definition. *Set $W : \mathcal{P}(R) \to \mathcal{P}(R)$ as the function:*

$$W : E \mapsto \{r' \in R | (x', r') \in \{x'', r''\} \text{ for some } (a'', \{x'', r''\}) \in \mathrm{supp}(\lambda(r)), r \in E\}$$

Definition. *Let $P_0 = \Delta(\mathcal{A})$ and $P_1 = \{p_1 \in R_1 | p_1 \in \Delta(\mathcal{A} \otimes P_0)\}$. Then, recursively let $P_n = \{p_n \in R_n | p_n \in \Delta(\mathcal{A} \otimes P_{n-1})\}$. Set $P = \big\{p \in \prod_{n=0}^{\infty} P_n | \times_{n=0}^{\infty} \lambda(W^n(r)) \subset \times_{n=0}^{\infty} \Delta(\mathcal{A} \otimes R)\big\}$.*

Proof of Theorem 1. We show that λ is a homeomorphism between P and $\Delta(\mathcal{A} \otimes P)$. First note, by construction, for all $r \in R$, $\lambda(r) \in \Delta^{B}(A \otimes W(r))$. Let $p \in P$; by the conditions on P, $\lambda(p) \in \Delta(\mathcal{A} \otimes R)$. Therefore, it suffices to show that for any $p \in P$, and $r \in W(p)$, $r \in P$. So fix some $r \in W(p)$. It follows from an analogous argument to Corollary 8.1 that $r \in \prod_{n=0}^{\infty} P_n$. Finally, note that $W^{n-1}(r) \subseteq W^n(p)$, for all $n \geqslant 2$. ■

References

Dirk Bergemann and Juuso Välimäki. Bandit Problems. In Steven N. Durlauf and Lawrence E. Blume, editors, *The New Palgrave Dictionary of Economics*. Second edition, 2008.

Donald A Berry and Bert Fristedt. *Bandit problems: sequential allocation of experiments (Monographs on statistics and applied probability)*. Springer, 1985.

Adam Brandenburger and Eddie Dekel. Hierarchies of Beliefs and Common Knowledge. *Journal of Economic Theory*, 59(1):189–198, 1993.

Kevin Cooke. Preference Discovery and Experimentation. 2016. *Working paper.*

Bruno de Finetti. Funzione caratteristica di un fenomeno aleatorio. 1931.

Bruno de Finetti. La prévision: ses lois logiques, ses sources subjectives. In *Annales de l'institut Henri Poincaré*, volume 7, pages 1–68, 1937.

Eddie Dekel, Barton L. Lipman, and Aldo Rustichini. Representing Preferences with a Unique Subjective State Space. *Econometrica*, 69(4):891–934, 2001.

David Dillenberger, Juan Sebastián Lleras, Philipp Sadowski, and Norio Takeoka. A theory of subjective learning. *Journal of Economic Theory*, pages 1–31, 2014.

David Dillenberger, R. Vijay Krishna, and Philipp Sadowski. Dynamic Rational Inattention. 2015. *Working paper.*

Larry G Epstein and Kyoungwon Seo. Symmetry of evidence without evidence of symmetry. *Theoretical Economics*, 5(3):313–368, 2010.

L.G. Epstein and M. Le Breton. Dynamically consistent beliefs must be Bayesian. *Journal of Economic Theory*, 61:1–22, 1993.

J. Gittins and D. Jones. A dynamic allocation index for the sequential allocation of experiments. In J Gani, K Sarkadi, and I Vincze, editors, *Progress in Statistics.* 1974.

Faruk Gul and Wolfgang Pesendorfer. Self-Control and the Theory of Consumption. *Econometrica*, 72(1): 119–158, 2004.

Edwin Hewitt and Leonard J. Savage. Symmetric measures on Cartesian products. *Transactions of the American Mathematical Society*, 80(2):470, 1955.

Kazuya Hyogo. A subjective model of experimentation. *Journal of Economic Theory*, 133(1):316–330, 2007.

Alexander Kechris. *Classical descriptive set theory*, volume 156. Springer Science & Business Media, 2012.

Peter Klibanoff, Sujoy Mukerji, and Kyoungwon Seo. Perceived Ambiguity and Relevant Measures. *Econometrica*, 82(5):1–46, 2013.

David M Kreps. A Representation Theorem for "Preference for Flexibility". *Econometrica*, 47(3):565–577, 1979.

R. Vijay Krishna and Philipp Sadowski. Dynamic Preference for Flexibility. *Econometrica*, 82(133):655–703, 2014.

Ehud Lehrer and Roee Teper. Who is a Bayesian? 2015. *Working paper.*

Evan Piermont, Norio Takeoka, and Roee Teper. Learning the Krepsian State: Exploration Through Consumption. 2015. *Working paper.*

Herbert Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–536, 1952.

Leonard J Savage. *The foundations of statistics.* Wiley, 1954.

Mark J Schervish. *Theory of statistics.* Springer Science & Business Media, 2012.