

# Learning the Krepsian State: Exploration Through Consumption\*

Evan Piermont<sup>†</sup>, Norio Takeoka<sup>‡</sup> and Roe Teper<sup>§</sup>

September 19, 2016

## Abstract

We take the Krepsian approach to provide a behavioral foundation for responsive subjective learning processes. In contrast to the standard subjective state space models, the resolution of uncertainty regarding the true state is endogenous and depends on the decision maker's actions. There need not be full resolution of uncertainty between periods. When the decision maker chooses what to consume, she also chooses the information structure to which she will be exposed. When she consumes outcomes, she learns her relative preference between them; after each consumption history, the decision maker's information structure is a refinement of the previous information structure. We provide the behavioral restrictions corresponding to a recursive representation exhibiting such a learning process. Through the incorporation of dynamics we are able to identify the set of preferences the decision maker believes possible after each history of consumption, without appealing to an environment with risk.

*Key words:* responsive learning; subjective learning; conditional preferences; preference for flexibility.

*JEL Classification:* D81; D83.

---

\*The authors wish to thank Kevin Cooke, Eddie Dekel, Kfir Eliaz, Edi Karni, R. Vijay Krishna, Pietro Ortoleva, Wolfgang Pesendorfer, Philipp Sadowski, Marciano Siniscalchi, Francois Forges, an Associate Editor and two anonymous referees for their comments and remarks. Takeoka gratefully acknowledges financial support from a Grant-in-Aid for Scientific Research (C).

<sup>†</sup>University of Pittsburgh, Department of Economics, [ehp5@pitt.edu](mailto:ehp5@pitt.edu)

<sup>‡</sup>Hitotsubashi University, Department of Economics, [norio.takeoka@r.hit-u.ac.jp](mailto:norio.takeoka@r.hit-u.ac.jp)

<sup>§</sup>University of Pittsburgh, Department of Economics, [rteper@pitt.edu](mailto:rteper@pitt.edu)

# 1 Introduction

Uncertainty regarding the state of the world, or one’s preferences, underlies almost every economic environment. In such settings, the agent might have the opportunity to actively explore in order to increase her understanding. A typical example are the models of strategic experimentation [Robbins, 1952, Gittins and Jones, 1979, Weitzman, 1979].<sup>1</sup> When the agent takes an action, she does not only derive utility from it, but also observes its consequence and gains better information pertaining to the underlying uncertainty. It is intuitive, then, that the agent’s understanding (or lack thereof) is a function of her experience, and subsequently, of her previous choices. We refer to such learning models as *responsive*. This paper provides a behavioral foundation and identification techniques for a model of *subjective and responsive learning*.

Kreps [1979] and Dekel et al. [2001] put forth the canonical models of introspective uncertainty, which identify the subjective uncertainty the decision maker (henceforth, DM) is facing regarding her own preferences. Implicitly, these models are static and assume that all uncertainty is realized in a way that is independent of the decision made by the agent. More recently, Takeoka [2007] and Dillenberger et al. [2014] introduce (static) models of gradual learning, and Krishna and Sadowski [2014, 2015] provide the behavioral foundation for a dynamic version of Kreps’ model that allows for tastes shocks. However, in these models, as in Kreps, learning is not *responsive*: the arrival of information is irrespective of the consumption choices made by the decision maker.

In many circumstances, such as a DM acknowledging the possibility of transient taste shocks, this is perfectly reasonable. It is likely the DM’s choice of restaurant in the morning will not affect her ability to discern her preference for fish and steak in the evening. However, if we consider the scenario of a DM learning her ranking over as-of-yet unconsumed alternatives, unresponsive learning is inadequate. For instance, imagine a novice researcher who is deciding on a project, but who is uncertain about her particular talents. She obtains no information without experience; there is no “exogenous” signal that will indicate her skill set. If she explores different research directions it will become apparent in which area she is more talented.

Weitzman [1979] introduced a model that studies the optimal strategy of a decision maker in this situation. The agent is engaged in a dynamic problem, trying to understand which project to invest in. Once she explores a project, she knows exactly what that project yields; this may also convey some information about other projects (in case the projects’ outcomes are correlated). At each point the agent has to decide whether to keep exploring new projects or to commit to already explored and known projects.

We here study the decision theoretic aspects of responsive subjective learning. We follow the decision theoretic learning literature and consider a dynamic constrained choice environment that extends the Krepsian framework. We provide the axiomatic foundation of a recursive utility function exhibiting subjective responsive learning. Initially, just as in Kreps, there is uncertainty regarding preferences. Then, each period, the DM jointly chooses a consumption outcome and a constraint for the following period so as to maximize her (current period) consumption utility and continuation value. The DM takes into account that her choice of consumption today may teach her about her preferences, altering her information structure, and accordingly, her preferences over future constraints. This representation captures the notion of responsive learning by allowing the continuation utility function and information structure to explicitly depend on previous consumption choices. In this respect, our framework expresses the intertemporal tradeoff between the consumption value of the choice today and its future informational value.

---

<sup>1</sup>See Bergemann and Välimäki [2006] for a more recent survey.

The decision maker’s subjective information structure changes *only* in response to her consumption choices, in a manner similar to Weitzman [1979].<sup>2</sup> In particular, our model considers a learning process characterized in part by two constraints. First, the DM’s uncertainty regarding her ranking between  $a$  and  $b$  is resolved once  $a$  and  $b$  are consumed. Second, if the DM is initially uncertain about her ranking between  $a$  and  $b$ , then this uncertainty is not fully resolved unless she consumes outcomes over which her preferences are perfectly correlated with  $a$  and  $b$ .

## 1.1 Results

The learning characteristics described above imply that preferences over already consumed goods are stationary. We begin the analysis by considering a somewhat simple framework combining flexibility and dynamics, where the DM chooses between menus of streams of outcomes. In this framework we state Theorem 1, showing that stationarity in the presence of flexibility is sufficient in order to uniquely identify the subjective uncertainty. It is worth noting, uniqueness is achieved even though we do not consider environments with risk, as is customary in the literature beginning with the uniqueness result provided by Dekel et al. [2001].

We then proceed to develop a complete model of learning in a framework of dynamic programming. In this domain we state and prove Theorem 3, which is the axiomatization of responsive subjective learning. In particular, we are able to elicit from behavior, not only (1) the set of states the DM believes possible, but also (2) how the DM expects to learn conditional on a given path of consumption, and (3) how the decision maker anticipates her preferences, over both future consumption and consequent information, will change after learning any one of the pieces of information she believes possible. The elicitation of (2) and (3) are novel.<sup>3</sup> The main complication in accommodating responsive learning is identifying (3).

What makes this elicitation, and thus the axiomatization, inherently difficult, is that we are interested in states that are not only subjective but also conditional. To understand the DM’s conditional preferences for information, it is first necessary to be able to condition on the relevant (subjective) state. This is not straightforward; the information structure is not directly observable as it is not incorporated into the primitive. A contribution of this paper is developing the tools allowing the modeler to elicit preferences conditional on subjective states.

Finally, Theorem 4 states that the subjective information structure (that is, the initial underlying uncertainty and how the DM expects to learn as she consumes) in our model is uniquely identified. What facilitates the identification in our environment is the dynamic and cumulative structure of learning. Our framework elicits the DM’s (anticipated) understanding of uncertainty at multiple points in the learning process. It is the stationarity of preferences over already consumed goods, and the consistency between the multiple elicitations embodied by the cumulative nature of learning, that allows for such a result. The examples in the subsequent section also expose how the dynamic aspect of choice uniquely determines the information structure.

---

<sup>2</sup>In a more general bandit problem, when learning pertains to distributional parameters, our partitional learning seems perhaps less realistic and noisy signals would be a natural generalization. It is also interesting to think about models that combine responsive learning and exogenous information flows (the latter as in Dillenberger et al. [2014]). These models bear important economic content, but their axiomatizations are far from being close to the one suggested here. For example, Piromont and Teper [2016] axiomatize model of general responsive learning, albeit in a substantially different choice environment.

<sup>3</sup>The concept of conditional preferences of conditional preferences, etc., has been examined in other contexts. For examples see, Gul and Pesendorfer [2005] and Siniscalchi [2011].

## 1.2 Organization

The following section presents a simple model that explores how temporal aspects of choice could identify a Krepsian state space, and discusses the main behavioral implication of learning through consumption in this environment. The complete model is continued in Section 3, beginning with a conceptual discussion, then introducing the framework in Section 3.1, and the axioms in Section 3.3. The main results are presented in Section 4. In particular, the representation theorem is in Section 4.2 and the uniqueness result is discussed and presented in Section 4.3. Section 5 contains a survey of the relevant literature.

Proofs omitted from the main text appear in the Appendix. In addition, Appendix C offers an alternative axiomatization connecting our main axiom to the dominance relation of Kreps [1979]. Finally, Appendix D discusses the additive specification of our general representation.

## 2 Flexibility with Stationary Preferences

In order to elucidate the mechanics behind our identification strategy and main axiom, we introduce the following simple choice environment, which is the minimal intersection between preference for flexibility and dynamic choice.<sup>4</sup> The DM is associated with a preference over menus à la Kreps, but menus contain *sequences* of consumption objects, rather than static prizes.

Let  $X$  denote a finite set of consumption prizes, and let  $\Sigma = \times^{\infty} X$  denote the set of all infinite sequences of consumption, with typical elements  $\sigma, \rho$ , etc. Let  $\sigma_n$  denote the  $n^{\text{th}}$  component of  $\sigma$ . For any  $a \in X$  let  $\mathbf{a}$  denote the constant sequence  $a$ . The decision maker's preference,  $\succsim$ , is therefore a subset of  $\mathcal{K}(\Sigma) \times \mathcal{K}(\Sigma)$ , where  $\mathcal{K}(\cdot)$  denotes the set of all non-empty (compact) subsets.<sup>5</sup> Let  $W, W'$ , etc. denote generic menus of streams. Let  $SO_r(X)$  denote the set of all strict orderings over  $X$ .<sup>6</sup>

Assume we can identify from flexibility the decision maker's perception of uncertainty (regarding her consumption preferences), that is, a subjective state space. In a departure from the standard Krepsian model, we endow the state space with additional structure, capitalizing on the dynamic component. Specifically, assume the state space,  $S$ , is composed of *stationary* preferences. That is,  $S$  is a subset of the possible orderings of  $X$ , and the representations of these states are stationary discounted utility functions as in Koopmans [1960].

**Definition.** *A preference  $\succsim$  has a **stationary sequential Krepsian representation** if there exists a state space  $S \subseteq SO_r(X)$  with associated utility representations  $u_s : X \rightarrow \mathbb{R}$  for each  $s \in S$ , and a discount rate,  $\delta \in (0, 1)$ , such that*

$$U(W) \equiv \sum_{s \in S} \left[ \max_{\sigma \in W} \sum_{n=1}^{\infty} \delta^{n-1} u_s(\sigma_n) \right]. \quad (\text{SSK})$$

*represents  $\succsim$ .*

There are two assumptions embodied in the formulation of Eq. (SSK). First, states are ordinal rankings over  $X$ . This assumption is in the heart of the paper, and our techniques do not uniquely identify cardinal

<sup>4</sup>We are thankful to an anonymous reviewer for suggesting a formulation of our results in this simplified framework.

<sup>5</sup>We introduce concepts here with some imprecision, as the technicalities (such as defining relevant topologies) are thoroughly discussed in Section 3.1, where the formal model is introduced.

<sup>6</sup>We assume throughout the state space in question does not admit indifferences (that is, states correspond to strict orderings of  $X$ ). This can be relaxed if we include a universal worst outcome, as discussed in Appendix B. But it clutters the otherwise notationally simple proof of Theorem 1.

rankings. Second, states and intertemporal tradeoffs are aggregated in an additive separable fashion. This is merely for expositional purposes. The arguments will follow for every strictly monotonic aggregator.

The remainder of this section exhibits how the additional structure provided by the stationary dynamic-component allows us to (1) uniquely identify the state space  $S$ , and (2) capture the gradual aspects of learning through consumption.

## 2.1 Uniqueness of the State Space

It is well known, in the absence of additional structure, the Krepsian state space (over static consumption alternatives) is not unique. This is shown in the following example borrowed from Dekel et al. [2001]. Let  $X = \{a, b, c\}$ . The DM ranks menus of  $X$  according to the number of elements they contain, that is,

$$\{a, b, c\} > \{a, b\} \sim \{b, c\} \sim \{a, c\} > a \sim b \sim c. \quad (1)$$

The table below specifies two different state spaces representing the same preference as in Eq. (1).

		$X$					$X$			
			$a$	$b$	$c$			$a$	$b$	$c$
$S$	$s_1$	3	2	1		$S'$	$s'_1$	3	1	2
	$s_2$	1	3	2			$s'_2$	2	3	1
	$s_3$	2	1	3			$s'_3$	1	2	3

Notice that  $S$  and  $S'$  induce different sets of ordinal rankings over  $X$ ,

$$S = \begin{cases} a >_{s_1} b >_{s_1} c \\ b >_{s_2} c >_{s_2} a \\ c >_{s_3} a >_{s_3} b \end{cases}, \text{ and } S' = \begin{cases} a >_{s'_1} c >_{s'_1} b \\ b >_{s'_2} a >_{s'_2} c \\ c >_{s'_3} b >_{s'_3} a \end{cases}.$$

Now, assume the DM's entertains a stationary sequential Krepsian representation which induces the static preferences outlined above. Clearly, by stationarity, her preferences over constant streams is dictated by Eq. (1). However, by examining non-constant streams, the state space can be identified.

Consider the following menus of infinite streams:

$$W = \{aaba, bcba, acca\}, \quad W' = W \cup aaca. \quad (2)$$

Since every stream in  $W$  and  $W'$  provides  $a$  from period 4 onward, the DM's preference between the menus will be dictated by the initial 3 periods. First, assume the DM's true model of uncertainty is embodied by state space  $S$ . While it can be calculated via the functional representation that  $W \sim W'$ , there is a more illuminating way of deriving this condition.

Notice, every stream in  $W$  and  $W'$  provides either  $a$  or  $b$  in the first period. Likewise, they all provide either  $a$  or  $c$  in the second,  $b$  or  $c$  in the third, and  $a$  in any subsequent period. If the true state was  $s_1$  then  $a$  is preferred to  $b$ ,  $a$  to  $c$  and  $b$  to  $c$ , and therefore,  $aaba$  is strictly preferred to any stream that provides  $a$  or  $b$  in the first period,  $a$  or  $c$  in the second,  $b$  or  $c$  in the third, and  $a$  from four onward. In particular, this implies  $aaba$  is the argmax of  $\sum_{n=1}^{\infty} \delta^{n-1} u_{s_1}(\cdot)$  over  $W$  and  $W'$ . Similarly,  $bcba$  maximizes  $\sum_{n=1}^{\infty} \delta^{n-1} u_{s_2}(\cdot)$  and  $acca$ ,  $\sum_{n=1}^{\infty} \delta^{n-1} u_{s_3}(\cdot)$ . Given this observation, it is immediate that the additional flexibility provided by  $W'$  relative to  $W$  is of no benefit to the DM who entertains state space  $S$ . By the same argument, we can

see that  $aaca$  is the (unique) optimal stream from  $W'$  if the true state is  $s'_1$ . Hence, a DM who entertains  $S'$  strictly prefers  $W'$  to  $W$ . Thus, by examining the preference over  $W$  and  $W'$  we can separate DMs who entertain  $S$  from those who entertain  $S'$ , something that could not be done in the static environment.

In the static case, the differences between states may wash out in the aggregation process. But, when dynamics are introduced, we can construct streams that are dominant for a given state by appropriately selecting streams so that in each period the difference between all streams depends only on the preference between two fixed consumption alternatives (in the example before, the first period is only a choice between  $a$  or  $b$ ). Then, due to stationarity, these pairwise preference can be aggregated to reveal the entirety of the utility function, and hence the state space. Of course, this intuition generalizes to all stationary sequential Krepsian preferences.<sup>7</sup>

**Theorem 1.** *If  $\succsim$  admits a stationary sequential Krepsian representation, then the state space is unique.*

*Proof.* Towards a contradiction, assume  $\succsim$  admits a stationary sequential Krepsian representation with state spaces  $S$  and  $S'$ . Since  $S \neq S'$ , there exists some state  $u_{\bar{s}}$  (w.l.o.g.) in  $S$  and not in  $S'$ . Let  $a \in X$  be an arbitrary consumption prize. Consider the set of all pairs of distinct elements of  $X$ , enumerated by  $\gamma_1 \dots \gamma_{\binom{|X|}{2}}$ . For any  $s \in S$  define  $\sigma^s$  as

$$\sigma_n^s = \begin{cases} \operatorname{argmax}_{b \in \gamma_n} u_s(b) & \text{if } n \leq \binom{|X|}{2} \\ a & \text{otherwise.} \end{cases} \quad (3)$$

Let  $W = \{\sigma^s | s \in S\}$  and  $W' = \{\sigma^{s'} | s' \in S'\}$ . According to (SSK), when the state space is  $S$ ,  $W \cup W' > W'$ , but when the state space is  $S'$ ,  $W \cup W' \sim W'$ , a contradiction. ■

## 2.2 Strategic Planning

This section uses the above framework to explicate *strategic planning*, the key behavioral restriction underlying the process of learning through consumption. The above identification result allows us to infer what the DM anticipates learning by the time when she must choose out of the menu, but does not capture the responsive aspect of learning through consumption. Since the process of interest is dictated by previously consumed outcomes, we here consider a *family* of SSK representations which are indexed by possible exploration histories (i.e., subsets of  $X$  which were previously explored by the DM). For any non-empty  $h \subseteq X$ ,  $\succsim_h$  expresses the flexibility the DM prefers when she will have the opportunity to explore elements in  $h$ ; this captures the information she expect to learn from her exploration of  $h$ .

The preference relations  $\{\succsim_h\}_{h \subseteq X}$  give rise to an associated family of SSK state spaces,  $\{S_h\}_{h \subseteq X}$ . Recall, the learning process of interest is principally characterized by the constraint that the DM's uncertainty is resolved by (and only by) learning her ordinal ranking between outcomes. Since the DM anticipates all uncertainty to be resolved after consuming every alternative, the state space  $S_X$  represents the DM's understanding of the entire space of uncertainty. The same idea applies to every exploration history. Following the exploration of  $h$ , the DM can distinguish the events in  $S_X$  described by her ranking of the alternatives  $h$ , which she has consumed. So, the learning through consumption process implies the family of state spaces  $\{S_h\}_{h \subseteq X}$  should correspond to a set of partitions,  $\{\mathcal{P}(h)\}_{h \subseteq X}$ , of  $S_X$ . In particular, a cell in the partition  $\mathcal{P}(h)$  is formed by the states (in  $S_X$ ) which coincide on their ranking over  $h$ .

<sup>7</sup>Kochov [2015] also uses the temporal aspect of choice problems, in lieu of lotteries, to facilitate identification. He identifies the set of probability measures in a Max-Min environment.

For example, a DM is considering four consumption objects: two laptops: Mac ( $m$ ) and PC ( $p$ ), and two phones: iPhone ( $i$ ) and Android ( $a$ ). Imagine, the DM believes that the true uncertainty is captured by the state space

$$S_X = \begin{cases} m \succ_{1^*} p \succ_{1^*} i \succ_{1^*} a, \\ m \succ_{2^*} p \succ_{2^*} a \succ_{2^*} i, \\ p \succ_{3^*} m \succ_{3^*} a \succ_{3^*} i, \end{cases} \quad (4)$$

where  $\succ_{i^*}$  is associated with the state  $s_{i^*}$ . The preference relation  $\succsim_{mp}$  represents the DM's preference over menus of sequences when she is to consume  $m$  in the first period,  $p$  in the second, and then choose a sequence from the menu after realizing her preference over  $m$  and  $p$ . As such, the DM's state space  $S_{mp}$  will have two states corresponding to the events in the partition  $\{\{s_{1^*}, s_{2^*}\}, \{s_{3^*}\}\}$ . Perhaps:

$$S_{mp} = \begin{cases} m \succ_1 p \succ_1 i \succ_1 a, \\ p \succ_2 m \succ_2 a \succ_2 i. \end{cases}$$

Similarly,  $\succ_i$  is associated with the state  $s_i$ . Notice, state  $s_1$  is the DM's aggregated preference in the event  $\{s_{1^*}, s_{2^*}\}$ , and  $s_2$  her aggregated preference in event  $\{s_{3^*}\}$ . Naturally, since the event  $\{s_{1^*}, s_{2^*}\}$  is defined by “ $m$  is preferred to  $p$ ”, then  $m \succ_1 p$ . Likewise  $p \succ_2 m$ ; in fact, since  $\{s_{3^*}\}$  contains only one state,  $\succsim_2 = \succsim_{3^*}$ .

Conversely, imagine that following the consumption of  $m$  and  $p$  the state space was  $S'_{mp} = S_X \cong \{\{s_{1^*}\}, \{s_{2^*}\}, \{s_{3^*}\}\}$ . In this case, after the DM consumes  $m$  and  $p$  she expects to learn her preference between  $a$  and  $i$  even when she learns “ $m$  is preferred to  $p$ ”. In this example, this learning cannot be rationalized by a responsive model, as the information she expects to learn is uncorrelated with her preference over her previous consumption. In the event “ $m$  is preferred to  $p$ ”, the DM's ability to distinguish her preferences over  $a$  and  $i$  (i.e., between the events  $\{s_{1^*}\}$  and  $\{s_{2^*}\}$ ) must be predicated on information unrelated to  $m$  and  $p$ .

It is worth clarifying:  $S_{mp}$  is her perception of what uncertainty will be resolved *by consuming  $m$  and  $p$* , and therefore its states correspond to events which can be distinguished given her experience. Thus, while she believes the true uncertainty to be represented by  $S_X$ , after  $mp$  she could not distinguish the events  $\{s_{1^*}\}$  and  $\{s_{2^*}\}$ . Her preference for flexibility would identify a state space resembling  $S_{mp}$  (by aggregating states  $s_{1^*}$  and  $s_{2^*}$ ).

This restriction, that the state spaces can be associated with a partitional structure, corresponds to the behavioral restriction embodied by *strategic planning*. Strategic planning dictates that the flexibility desired after  $h$  (i.e., by  $\succsim_h$ ) is determined completely by the DM's ranking over  $h$ . Going back to the example, notice,

$$\{\mathbf{a}, \mathbf{i}\} \succ_{mp} \mathbf{a} \quad \text{and} \quad \{\mathbf{a}, \mathbf{i}\} \succ_{mp} \mathbf{i}.$$

That is, the DM prefers flexibility to choose her phone after seeing her preference over laptops. However, the DM will exhibit the rankings

$$\{p\mathbf{a}, m\mathbf{i}\} \sim_{mp} \{p\mathbf{a}, m\mathbf{i}, p\mathbf{i}, m\mathbf{a}\},$$

which implies that the DM does not require the flexibility to choose the phone and laptop separately; she is

	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\sigma$
$a \succ b \succ c$	$a$	$a$	$b$	$aab$
$a \succ c \succ b$	$a$	$a$	$c$	$aac$
$b \succ a \succ c$	$b$	$a$	$b$	$bab$
$b \succ c \succ a$	$b$	$c$	$b$	$bc b$
$c \succ a \succ b$	$a$	$c$	$c$	$acc$
$c \succ b \succ a$	$b$	$c$	$c$	$bcc$

Figure 1: Example of  $\Sigma(h)$  with  $h = \{a, b, c\}$

happy to jointly make this decision. Of course, this is because her preference over phones (as reflected by  $S_{mp}$  after consuming  $mp$ ) is completely determined by her preference over laptops; when choosing from the menu  $\{p\mathbf{a}, m\mathbf{i}, p\mathbf{i}, m\mathbf{a}\}$ , the choice between  $m$  and  $p$  in the first period perfectly reveals the state in  $S_{mp}$ .

This intuition is generalized, capitalizing on our ability to identify states in a manner similar to the proof of Theorem 1. So, let  $\Gamma(h)$  denote the set of all pairs of distinct elements of  $h$ , enumerated as  $\gamma_1 \dots \gamma_{\binom{|h|}{2}}$ . Then, for any  $s \in SOr(h)$  define  $\sigma^s \in h^{\binom{|h|}{2}}$  as

$$\sigma_n^s = \operatorname{argmax}_{b \in \gamma_n} u_s(b), \quad (5)$$

where  $u_s$  is any utility representation of  $s$ . Let  $\Sigma(h) = \{\sigma^s | s \in SOr(h)\}$ .

For a general exploration history  $h$ , the elements of  $\Sigma(h)$  represent the possible strict rankings over the outcomes comprising  $h$ . To make this clear, consider the following example. Let  $h = \{a, b, c\}$ , and take the following ordering over doubletons:  $\{a, b\}, \{a, c\}, \{b, c\}$ . Then,  $\Sigma(\{a, b, c\}) = \{aab, aac, bab, bcb, acc, bcc\}$ . Each of the 6 sequences in  $\Sigma(\{a, b, c\})$  corresponds to one of the six possible strict rankings  $\succ$  over  $\{a, b, c\}$  as shown in Figure 1.

If the DM knows her ranking of the elements of  $h$ , then there is a unique sequence in  $\Sigma(h)$  that maximizes that ranking. Indeed, notice that all the sequences in  $\Sigma(\{a, b, c\})$  give either  $a$  or  $b$  in period 1, either  $a$  or  $c$  in period 2, and either  $b$  or  $c$  in period 3. If  $a \succ b \succ c$ , the consumption stream  $aab$  dominates all the other consumption streams in  $\Sigma(\{a, b, c\})$  by period-by-period comparison, and hence is most preferable among  $\Sigma(\{a, b, c\})$ . Similarly, if  $a \succ c \succ b$ ,  $aac$  is most preferable among  $\Sigma(\{a, b, c\})$ , and so on. Because the comparison is through period-wise dominance, consumption streams in  $\Sigma(h)$  can be ranked only based on single-period rankings without knowing any additional properties regarding intertemporal tradeoffs like discount factors or elasticity of intertemporal substitution.

Clearly, there is a bijective relation between  $\Sigma(h)$  and  $SOr(h)$ , since if the DM knows her ranking over  $h$ , then there is a unique sequence in  $\Sigma(h)$  maximizes that preference. Moreover, there is an injective relation between  $\mathcal{P}(h)$  and  $SOr(h)$ , since each element of the partition corresponds to a ranking over  $h$ . Strategic planning provides the link between these relations, ensuring that the identified state space is consistent with such a partition.

**Definition.** Let  $W \subset \Sigma$  be a finite collection of sequences. Then  $p^W : \Sigma(h) \rightarrow W$  is an **SSK strategic plan** if

$$\bigcup_{\sigma \in \Sigma(h)} \sigma p^W(\sigma) \sim_h \bigcup_{\sigma \in \Sigma(h)} \bigcup_{\rho \in W} \sigma \rho$$



In the example above,  $m \mapsto \mathbf{i}$  and  $p \mapsto \mathbf{a}$  constitutes a strategic plan for  $W = \{\mathbf{i}, \mathbf{a}\}$ . The intuition for the general definition is the same. A strategic plan takes flexibility inherent in making two separate choices (i.e., from  $\Sigma(h)$  and from  $W$ ) and reduces it to a single choice (from  $\Sigma(h)$ ). From the identification of  $\Sigma(h)$  and  $SO_r(h)$ , the existence of a strategic plan (for every finite collection,  $W$ ) implies that the resulting SSK state space  $S_h$  can be identified with the partition  $\mathcal{P}(h)$ .

### 2.3 The Limitation of the Simple Model

The existence of a strategic plan for all exploration histories and all finite menus is necessary for a model of learning through consumption, but alone it is insufficient to guarantee any consistency between the identified state spaces. Our interpretation is that preference for flexibility arises from a cumulative learning process, so the DM’s perceptions of uncertainty after different histories all arise from a single ex-ante view of what information might be realized. In the discussion above, this consistency is informally implied by our assertion that  $S_h$  corresponds to a cell in the partition  $\mathcal{P}(h)$ . This implication is informal as one needs to show how the different states in a cell of  $\mathcal{P}(h)$  are aggregated to the proper state in  $S_h$ .

What seems to be a technical issue stems from a more conceptual problem. Due to the nature of the SSK state spaces, the modeler does not have access to how the DM expects to continue to learn –given the event “ $m$  is preferred to  $p$ ,” what does the DM expect to learn from further consuming  $a$  and  $i$ ? Given exploration history  $h$ , we can only identify the DM’s anticipated change in her preference over consumption streams, but not her preference over *further information* –a key aspect of exploration and exploitation models. This would require the DM to be able to capitalize on further information, providing flexibility in multiple periods.<sup>8</sup>

To illustrate this point, recall the example in the previous subsection. The DM has in mind a subjective state space  $S_X$  as given by (4). The responsive learning implies that after exploration history  $mp$ , the DM anticipates an information structure given by the partition  $\{\{s_{1^*}, s_{2^*}\}, \{s_{3^*}\}\}$ . However, since only the aggregated state matters for evaluating consumption streams, what can be identified within the simple model is not this partition itself but the state space  $S_{mp}$ , elements of which are presumably obtained by aggregating different states in a cell of the partition.

Given  $S_{mp}$ , it is easy to see the DM will display the following indifference

$$\{paia, maii\} \sim_{mp} \{paia, maii, paii, maia\}.$$

When trees of menus (i.e., compound menus) are taken as choice objects, the DM’s uncertainty conditional on a cell of the partition can be identified. For example, the compound menu  $\{paia, mai\{\mathbf{a}, \mathbf{i}\}\}$  allows the DM to delay a decision between  $\mathbf{a}$  and  $\mathbf{i}$  until after consumption  $a$  and  $i$ . If the DM anticipates, after learning  $\{s_{1^*}, s_{2^*}\}$ , that she will further learn which state in the cell is true, she would desire this flexibility. Thus,

$$\{paia, mai\{\mathbf{a}, \mathbf{i}\}\} >_{mp} \{paia, maii, paii, maia\}.$$

This ranking reveals that after exploration history  $mp$ , the DM expects residual uncertainty about the ranking between  $a$  and  $i$  even after she learns “ $m$  is preferred to  $p$ ”. We can deduce that her information structure would look like  $\{\{s_{1^*}, s_{2^*}\}, \{s_{3^*}\}\}$ . Such an inference is not possible within the domain of the simple model.

---

<sup>8</sup> The simplest enrichment, however, looking at streams of *menus*, is not sufficient. Streams of menus do not allow the DM’s (ex-ante) preference for flexibility in later periods to depend on the information that arrived in earlier periods: only aggregated preference can be identified. The needed domain is one in which the choice objects are *trees of menus*.

In the subsequent sections we develop a model of dynamic programming that will fully capture the features of learning through consumption as a model of exploration and exploitation.

### 3 A Complete Model of Learning Through Consumption

To better understand the connection between the snapshots of the DM's information structure (i.e.,  $\{S_h\}_{h \subseteq X}$ ) and the learning process which underlies it, we appeal to a richer, recursive model. The objects of choice are Infinite Horizon Choice Problems (IHCPs) (formally constructed in section 3.1) and are generically denoted by  $z, w$  etc, with  $Z$  the set of all IHCPs. Each IHCP  $z$  induces a menu  $M(z) \subseteq X \times Z$ . Each element,  $i \in M(z)$ , is a pair: a consumption today,  $a_i \in X$ , and a continuation problem tomorrow,  $z_i \in Z$ . Since  $z_i$  is also an IHCP, we can write it too as a menu of pairs of consumptions and continuation problems:  $z_i = \{(b_j, z_j)\}_{j \in M(z_i)}$ . It is helpful to think of an IHCP as an infinite tree, where each node is an IHCP and the branches emanating from a node correspond to the elements of the induced menu, see Figure 2. The menu induced by  $z$ ,  $M(z)$ , is the set of branches emanating from from  $z$ .

Since an IHCP is a tree of menus, at the beginning of each period, the DM faces a menu,  $M(z)$ , from which she must jointly choose that period's consumption and the next periods IHCP. As such, she can exhibit a preference for flexibility at each period, and thus, she can capitalize on her information at multiple points in time.

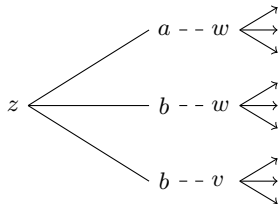


Figure 2: Representation of an IHCP  $z$  as a tree of constraints. Note that in the first period, two pairs offer  $w$  as the next period IHCP and two pairs offer  $b$  as the consumption prize.

Consider a DM who is ranking contingent IHCPs to be consumed after exploration history  $h$ . She is aware that when she chooses a path of consumption from an IHCP she will have the information structure induced by  $h$ : the partition  $\mathcal{P}(h)$ . This imparts an ex-ante preference for flexibility. Hence her value for IHCPs after history  $h$ ,  $U_h$ , takes the form,<sup>9</sup>

$$U_h(z) = \phi_h \left( \max_{i \in M(z)} F[u_P(a_i), U_{h \cup a_i | P}(z_i)]_{P \in \mathcal{P}(h)} \right), \quad (6)$$

where  $\phi_h$  (that aggregates information dependent utilities) and  $F$  (that aggregates consumption and exploration utilities) are strictly increasing aggregators.

The DM is able to make distinct choices contingent on the resolution of uncertainty embodied by the information provided by the consumption of  $h$  and  $a$ . As the DM consumes, both the information structure and the conditioning events evolve. The information structure is refined,  $\mathcal{P}(h) \rightarrow \mathcal{P}(h \cup a)$ , as a direct result

<sup>9</sup>When consumption experience increases, the expected information structure becomes weakly finer:  $\mathcal{P}(h \cup a)$  is a finer partition than  $\mathcal{P}(h)$ . Moreover,  $\mathcal{P}(h \cup a | P)$  stands for the partition of the cell  $P$  generated by the consumption of  $h \cup a$ .

of the addition of new elements to the history. In turn, the conditioning events become more informative,  $P \rightarrow P'$ , as the DM utilizes this available information in each period's consumption choice.

Note that the continuation value does not take the same form. This is because, when choosing from  $M(z)$  she utilized her available information, so subsequent choices will take this into account. Hence, any future choice must be conditioned on being in the true cell of the partition of  $\mathcal{P}(h)$ . The functional,  $U_{h|P}$ , with information structure  $\mathcal{P}(h)$  and conditioning event  $P \in \mathcal{P}(h)$ , takes the following *Learning Through Consumption* functional form:

$$U_{h|P}(z) = \max_{i \in M(z)} F[u_P(a_i), U_{h \cup a_i|P}(z_i)], \quad (\text{LTC})$$

where the aggregation property of information dependent utilities takes the form  $u_P = \phi_{h'|P}(u_{P'})_{P' \in \mathcal{P}(h'|P)}$  for every  $P \in \mathcal{P}(h)$  and  $h' \supset h$ .

### 3.1 Framework

Let  $X$  be a finite set, with some metric  $d_X$ . The set  $X$  represents everything that the decision maker could consume (in a given period) and is fully observable. Each period the decision maker consumes one element of  $X$ ; let lowercase letters near the beginning of the alphabet (i.e.,  $a, b, c$ ) denote elements of  $X$ . For any metric space,  $Y$ , let  $\mathcal{K}(Y)$  denote the set of all compact subsets of  $Y$  endowed with the Hausdorff metric. We can construct the set of IHCPs in a manner similar to Gul and Pesendorfer [2004]. A one period choice problem is a collection of alternatives from which the DM must choose –an element of  $\mathcal{K}(X)$ . The set of all one period choice problems,  $Z_1$ , corresponds to the set of all compact subsets of  $X$  (i.e.,  $\mathcal{K}(X)$ ).

An  $N$  period choice problem is a collection of pairs consisting of an alternative to consume and an  $N - 1$  period choice problem. As such, the set of all  $N$  period choice problems,  $Z_N$ , corresponds to  $\mathcal{K}(X \times Z_{N-1})$ . Note that compactness is defined in the product topology over  $X \times Z_{N-1}$  (where the topology of  $Z_{N-1}$  is itself the product topology defined similarly). We can continue this iteration so as to define  $\hat{Z} = \times_{n=1}^{\infty} Z_n$ . We will restrict ourselves to the set of *consistent* elements of  $\hat{Z}$ , where  $z \in \hat{Z}$  is consistent if for all  $n$ ,  $proj_{Z_{n-1}} z_n = z_{n-1}$  (where,  $proj$  is the projection mapping). Let  $Z \subset \hat{Z}$  be the restriction of  $\hat{Z}$  to consistent elements. Call  $Z$  the set of Infinite Horizon Choice Problems. The primitive of the model is a preference relation,  $\succsim$ , over elements of  $Z$ .

A member of  $Z$  specifies an  $n$ -period problem for each  $n \in \mathbb{N}$ . The restriction of consistency ensures that for each  $z \in Z$ , and each  $m, n \in \mathbb{N}$  with  $n > m$ , the specified  $n$ -period problem is an extension of the specified  $m$ -period problem. That is, the first  $m$ -periods of the  $n$ -period problem coincide with the  $m$ -period itself. Intuitively, this allows us to view each  $z \in Z$  as an *infinite* period problem, by considering the sequence of arbitrarily large, and expanding, finite period problems.

Of course, an infinite period problem will have a particular recursive property: the objects assigned after the first period consumption will themselves be infinite period problems. This is substantiated by the fact that there exists a canonical homeomorphism between  $Z$  and  $\mathcal{K}(X \times Z)$ .<sup>10</sup> By the identification between

<sup>10</sup>This result can be seen as a special case of the similar claim in Gul and Pesendorfer [2004], where we restrict the domain at each iteration to degenerate lotteries. More constructively, it is easy to check that the following functions,  $I$  and  $M$ , are well defined, continuous, and inverses of one another:

Define  $I_0 : \mathcal{K}(X \times Z) \rightarrow Z_1$  via the following map:  $I_0(y) = \{a \in X | (a, z) \in y \text{ for some } z \in Z\}$ . Then for each  $n \geq 1$ , define  $I_n : \mathcal{K}(X \times Z) \rightarrow Z_{n+1}$  via the following map:  $I_n(y) = \{(a, z_n) \in X \times Z_n | (a, z) \in y \text{ for some } z \in Z \text{ such that } proj_n(z) = z_n\}$ . And let  $I : \mathcal{K}(X \times Z) \rightarrow Z$  defined by  $I(y) = \times_{n=0}^{\infty} I_n(y)$ .

$Z$  and  $\mathcal{K}(X \times Z)$ , we can describe each  $z$  as a collection in  $X \times Z$ :  $z = \{(a_i, z_i)\}_{i \in M(z)}$  for some index set  $M(z)$ . We call  $M(z)$  the menu induced by  $z$ .

### 3.2 Notation

**Tree-amalgamating operations.** For any two IHCPs,  $z, w \in Z$ , we can define their union,  $(z \cup w)$  as the IHCP such that  $M(z \cup w) = M(z) \cup M(w)$ . See Figure 3.

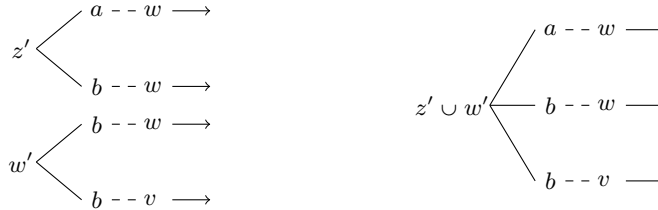


Figure 3: Representation of IHCPs  $z'$  and  $w'$ , and their union  $z' \cup w'$ . Note,  $(b, w)$  is available both in  $z'$  and  $w'$  and appears only once in  $z' \cup w'$ ; and,  $z' \cup w' = z$  as shown in Figure 2.

We will refer to (finite) collections of IHCPs with capital letters near the end of the alphabet ( $W, V$ ). In an abuse of notation, we will identify each collection of IHCPs with the IHCP that is the union of its elements: the collection  $W$  can be treated as the IHCP:  $W = \bigcup_{w \in W} w$ . With this identification, any IHCP,  $z$ , that does not have a degenerate first period choice, could also be described as a (not necessarily unique) collection of smaller IHCPs. As such, the distinction between collections and individual IHCPs is not strict (i.e., for every  $W$  there is an associated  $w$ ) but the notation will be helpful later.

**Tree-traversal operations.** These notations identify particular branches of a given IHCP. For a given,  $z \in Z$ , we can define the set of all feasible (decision) sequences of length  $n$ ,  $n(z)$ . Given  $z$ , a decision is feasible if it is an element of the implied menu:  $i \in M(z)$  (with the corresponding continuation tree  $z_i$ ). A feasible decision sequence of length  $n$ , for IHCP  $z$ , is a sequence of  $n$  choices, where the  $m^{\text{th}}$  choice is a feasible choice out of the continuation tree implied by the  $(m-1)^{\text{th}}$  choice and the first choice feasible out of  $M(z)$ . In other words, each decision sequence corresponds to a particular path up the tree. We will denote such sequences with lower case greek letters:  $\sigma, \rho$ , etc. See Figure 4.<sup>11</sup>

**Tree-pruning and concatenation operations.** For any finite choice problem,  $A \in Z_n$ , we can define  $Az$  as the IHCP that assigns  $z$  as a continuation tree to every feasible decision sequence of length  $n$  (contained in  $A$ ). In particular,  $az$  is the IHCP which begins with consumption of  $a \in X$  followed by the IHCP  $z$ . Likewise, for some finite sequence of consumption,  $\sigma \in \times_{n=1}^N X$ ,  $\sigma z$  is the IHCP which begins with consumption  $\sigma$  followed by  $z$ .

For some  $z$ , let  $\sigma$  be a finite, feasible decision sequence. Then we can define  $z_\sigma$  as the continuation tree after choosing  $\sigma$ . From this, we can define  $z_{-\sigma}w$  as the IHCP that agrees with  $z$  everywhere but after the feasible sequence  $\sigma$ , where  $z_\sigma$  is replaced by  $w$ . See Figure 4. Similarly, for an IHCP  $z$ , a finite feasible

---

Now define  $M_0 : Z \rightarrow \mathcal{K}(X)$  via  $M_0(z) = z_1$ . For each  $n \geq 1$  and  $a \in M_0$  define  $M_{a,n} : Z \rightarrow \mathcal{K}(Z_n)$  via  $M_{a,n}(z) = \{z'_n \in Z_n | (a, z'_n) \in \text{proj}_{n+1}(z)\}$ . Let  $M_a(z) = \bigcap_{n \geq 0} \{z' \in Z | \text{proj}_n(z') = M_{a,n}(z)\}$ . Finally, define  $M : Z \rightarrow \mathcal{K}(X \times Z)$  as  $M(z) = \{(a, z') | a \in M_0(z) \text{ and } z' \in M_a(z)\}$ .

<sup>11</sup>In Figures 2-5, we have suppressed the  $M(z)$  indexing notation in order to make the illustrations cleaner (both visually and conceptually). It is worth noting that the *true* description of  $\sigma$  in Figure 4 is actually  $\sigma = b_2 a'_1$  (under the obvious index) so as to distinguish it from a different path (for example if  $a'$  was available in  $v$ ).

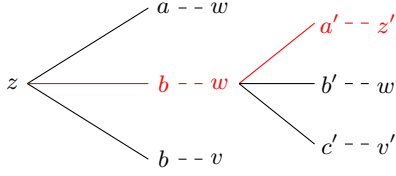


Figure 4: The red path corresponds to the feasible sequence  $\sigma = ba'$ , and  $z_\sigma = z'$ . The IHCP  $z_{-\sigma}z''$  would be identical to the above except with  $z'$  replaced with  $z''$ .

decision sequence  $\sigma$ , and a finite choice problem  $A$ , let  $z_{-\sigma}Az_\sigma$  be the IHCP in which, if the path  $\sigma$  was chosen then the continuation problem is  $Az_\sigma$ . That is,  $z_{-\sigma}$  was concatenated with  $Az_\sigma$ . Finally, we denote by  $z_{-n}Az_n$  the IHCP resulting from the concatenation of  $z_{-\sigma}$  with  $Az_\sigma$  for *every* feasible sequence of length  $n$  (that is,  $\sigma \in n(z)$ ), see Figure 5.

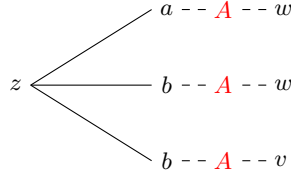


Figure 5: The IHCP  $z_{-n}Az_n$  for  $n = 1$ .

Finally, the use of bold lettering signifies degenerate IHCPs – predetermined sequences of consumption. So for each  $\sigma \in \times_{n=1}^\infty X$ , we identify  $\sigma$  as the IHCP such that there is only one feasible decision sequence of any given length, where the unique sequence coincides with the first  $n$  terms of  $\sigma$ . The above notational conventions are summarized in Table 6.

Notation	Meaning	Notes
$a, b, c$	Single period consumption prizes	
$z, w, v$	IHCPs	
$\sigma, \rho, \pi$	Streams of consumption	Finite unless otherwise noted, also represents feasible sequences when subscripted as in $z_\sigma$ or $z_{-\sigma}$
$W, V$	Collections of IHCPs	Identified with their unions (and so also IHCPs)
$Az$	A finite choice problem followed by an IHCP	
$\sigma, \rho$	Degenerate IHCPs	Specify a predetermined (infinite) sequence of consumption

Figure 6: Table of notational conventions.

### 3.3 Axioms

#### 3.3.1 Exploration Histories

At any given node in an IHCP we can define the set of consumption prizes that must have been consumed to reach that node. Formally, define the exploration history at a node as the support<sup>12</sup> (in  $X$ ) of the (unique) decision sequence that leads to that node. In other words, the exploration history does not remember the index but only remembers the consumption (i.e., ignores the menu from which the consumption was chosen, the order of consumption, and the number of times a prize has been consumed). We assume as a structural “axiom 0” that the preference for continuation trees depends *only* on the history.

[A0: ONLY HISTORIES MATTER (HST)]. *For all  $z, w \in Z$  and  $\sigma, \rho$  such that  $\text{supp}(\sigma) = \text{supp}(\rho)$*

$$\sigma z \succcurlyeq \sigma w \iff \rho z \succcurlyeq \rho w$$

This allows us to simplify notation with the following abuse, letting  $h$  denote both a subset of  $X$  (the set of previously consumed outcomes), and any specified sequence consisting of exactly those elements. As such, we can define the history dependent preferences.

**Definition.** *For each  $h$ , let  $\succcurlyeq_h$  denote the projection of preferences after exploration history  $h$  defined by*

$$z \succcurlyeq_h w \iff hz \succcurlyeq hw.$$

#### 3.3.2 Ex-ante Preferences

We assume that the following axioms hold for all  $h$ . To achieve the basic utility structure we have that  $\succcurlyeq$  is a continuous weak order:

[A1: CONTINUOUS WEAK ORDER (ORD)]. *The binary relation  $\succcurlyeq$  is a continuous weak order.*

The first substantive axiom governs the learning process. We are assuming the DM’s preference does not change unless she learns something (or, more accurately, anticipates learning). If the decision maker has already consumed  $a$  then consuming  $a$  again will not teach her anything new, and her preference will not change. Hence, we postulate that consuming items that are already in the exploration history will not change her preferences.

[A2: H-STATIONARITY (STA)]. *For any  $z, w, v \in Z$ ,  $\rho$  feasible for  $v$ , and  $n > 0$ , if  $A, B$  are finite decision problems:*

1. *if  $\text{supp}(A) \subseteq h \cup \bigcap_{\sigma \in n(z) \cup n(w)} \sigma$ , then  $z \succcurlyeq_h w \iff z_{-n}Az_n \succcurlyeq_h w_{-n}Aw_n$ ; and*
2. *if  $\text{supp}(B) \subseteq h \cup \rho$ , then  $v_{-\rho}B\{z \cup w\} \sim_h v_{-\rho}\{Bz \cup Bw\}$ .*

STA has two components. First, (1) dictates that the decision maker does not change her preference when consuming prizes she has already consumed. The motivation for this is straightforward, as the only mechanism of learning is consumption of new prizes. The formal restriction states that if consuming the

<sup>12</sup>The support of a (finite) decision problem, and in particular a feasible sequence, is all the elements of  $X$  that can be consumed in that problem.

decision problem  $A$  is unavoidable (i.e., it will be reached on *every* path of consumption after  $n$  periods) and if all of the consumption prizes in  $A$  will already have been consumed (i.e.,  $\text{supp}(A) \subseteq h \cup \bigcap_{\sigma \in n(z) \cup n(w)} \sigma$ ), then the inclusion of  $A$  will not influence the DM's preferences between  $z$  and  $w$ . A special case of (1) states that if  $\text{supp}(A) \subseteq h$  then  $z \succsim_h w \iff Az \succsim_h Aw$ , which is reminiscent of the canonical stationarity axiom of Koopmans [1960] (applied to previously consumed prizes). We require the more general formulation to enforce that preferences remain stationary (with respect to previously consumed prizes) even as they change (with respect to novel prizes).

Second, (2) dictates that the decision maker is indifferent to making a choice before or after the consumption of previously consumed prizes. In general, the decision maker would prefer to delay making decisions, as it would allow her to condition on more information. However, if in the interim the decision maker learns no new information (because she consumes no novel prize), then she is just as well off making the decision now.

Without learning, the union of two IHCPs would never be preferred to both; the decision maker knows already which stream she would choose and would (weakly) prefer the menu that contained this stream. When the decision maker anticipates learning then flexibility is strictly beneficial, as it allows her to capitalize of the information she learns. Flexibility allows the DM to make different decisions conditional on the realization of the learning process.

[A3: PREFERENCE FOR FLEXIBILITY (FLX)]. *For all  $v, z, w \in Z$  and for all feasible decision sequences given  $v$ :*

$$v_{-\sigma}(z \cup w) \succsim_h v_{-\sigma}z$$

This axiom looks slightly different than its canonical form. We require that flexibility is not disadvantageous, not only at time zero, but also on any branch of any IHCP. Because we have relaxed stationarity it is not enough to impose that  $z \cup w \succsim_h z$  since this could be true for initial choices but be violated for continuation problems on a particular branch of some IHCP.

Next, we need some notion that the DM is consistent in how she expects to learn. The DM's (time 0) anticipated preference at  $h'$  should be the same as how she anticipates (at time 0) anticipating (at  $h \subset h'$ ) her preference at  $h'$ . In words, breaking her expected learning process into separate pieces should not affect the final preference.

[A4: CONSISTENT FLEXIBILITY (CON)]. *For every  $h' \supseteq h$ , and set of sequences contained in the history,  $A \subseteq \times_{i=1}^N h$ , and functions,  $f, g : A \rightarrow Z$  then*

$$\begin{aligned} \bigcup_{\sigma \in A} \sigma \pi f(\sigma) \sim_h \bigcup_{\sigma \in A} \sigma \pi \{f(\sigma) \cup g(\sigma)\} \\ \iff \\ \bigcup_{\sigma \in A} \sigma f(\sigma) \sim_{h'} \bigcup_{\sigma \in A} \sigma \{f(\sigma) \cup g(\sigma)\} \end{aligned}$$

*for all  $\pi$  such that  $h \cup \pi = h'$ .*

The DM must choose out of  $A$  today, which will determine the IHCP she will receive after consuming  $\pi$  which will create the exploration history  $h'$ . If she sees no benefit to the additional flexibility provided by the assignment  $g$ , then it is because under all possible realizations of uncertainty brought about by her

consumption of  $h'$ , the IHCP assigned by  $f$  is preferred to that by  $g$ . But if this is the case, there must be no benefit to flexibility according to the preference  $\succsim_{h'}$ .

### 3.3.3 Strategic Planning

Our next axiom generalizes the notion of an SSK strategic plans to the setting where choice objects are IHCPs. The motivation is identical as before: to ensure that *the outcome of prior actions (i.e., the rankings of consumed goods) is a sufficient statistic to predict future choice*. If the decision maker is told (ex-ante) the realization of her ranking over the elements of some exploration history  $h$ , she can exactly predict her (interim, after history  $h$ ) preference over all other IHCPs. She has no value in waiting to make a decision because any information she might learn is contained by the information about her ranking over  $h$ .

We construct the choices that will correspond to the ordinal rankings over the exploration history as in section 2.2; recall the following bit of notation:

**Definition.** For every  $h$ , such that  $|h| \geq 2$ , let  $\Gamma(h)$  denote the set of all pairs of distinct elements of  $h$ , enumerated as  $\gamma_1 \dots \gamma_{\binom{|h|}{2}}$ . Then, For any  $s \in \text{SOr}(h)$  define  $\sigma^s \in h^{\binom{|h|}{2}}$  as

$$\sigma_n^s = \operatorname{argmax}_{b \in \gamma_n} u_s(b), \quad (7)$$

where  $u_s$  is any utility representation of  $s$ . Let  $\Sigma(h) = \{\sigma^s | s \in \text{SOr}(h)\}$ . For  $h$  with  $|h| = 0, 1$ , let  $\Sigma(h) = \emptyset$ .

Recall, if the DM knows her ranking over  $h$ , then there is a unique element of  $\Sigma(h)$  that maximizes that preference. So, for any collection of IHCPs  $W \subseteq Z$ , the DM chooses a distinct alternative from the IHCP

$$\{\sigma W | \sigma \in \Sigma(h)\}$$

according to a distinct ranking over  $h$ . Therefore, if we can observe a subsequent choice from  $W$  following a particular consumption stream in  $\Sigma(h)$ , we can infer that such a choice is based on preference conditional on the corresponding ranking over  $h$ .

Our next axiom, **PLN**, dictates that such a mapping between  $\Sigma(h)$  and the conditional preferences exists. It will turn out (as a straightforward consequence of Theorem 4) that  $\Sigma(h)$  uniquely describes the DM's information structure. It partitions the subjective state space into equivalence classes of states that respect the same ranking over the elements of  $h$ . With this in mind it becomes clear that subsequent partitions are refinements of previous ones. To capture this idea more formally and reveal the DM's preference contingent on being in a state, we introduce the notion of assignments and strategic plans.

**Definition.** Given any collection of IHCPs,  $W \subseteq Z$ , and a history,  $h$ , an **assignment** is a function from  $\Sigma(h)$  to  $W$ . Let  $As(W, h)$  be the set of all such assignments.

An assignment for  $W$  and  $h$  is function that maps to each sequence of consumption in  $\Sigma(h)$  and an IHCP out of  $W$ . Then, a *strategic plan* is an assignment such that flexibility is no longer beneficial (beyond choosing from  $\Sigma(h)$ , that is).

**Definition.** Given a history,  $h$ , and  $W \subseteq Z$ , a **strategic plan** is an assignment  $p^W \in As(W, h)$ , such that

$$\bigcup_{\sigma \in \Sigma(h)} \sigma(p^W(\sigma)) \sim_h \bigcup_{\sigma \in \Sigma(h)} \sigma W$$



for all  $\Sigma(h)$ .

A strategic plan for  $W$  assigns to each sequence in  $\Sigma(h)$  a subset of the choice problem (i.e., an element in the collection  $W$ ) in such a way that there is no additional benefit from flexibility. Since each element of  $\Sigma(h)$  corresponds to a subjective state, then  $z = p^W(\sigma)$  implies that  $z$  is the maximizing element of  $W$  in the state that corresponds to  $\sigma$ .

[A5: STRATEGIC PLANNING (PLN)]. For all  $h$  and all  $W = \{z_i\}_{i=1}^N \subset Z$  there exists a strategic plan over  $W$  with respect to  $\succsim_h$ .

Since flexibility is beneficial only when the DM wants to capitalize on learning distinct pieces of information, axiom PLN dictates that the only information arriving is from the realization of her ranking of the elements of  $h$ . Indeed, if there was some additional un-responsive information flow, the DM would have a strict preference for flexibility, even after conditioning on her ranking over  $h$ .

The economic content of the axiom is precisely the following: the existence of a strategic plan implies that everything the DM anticipates learning is described by her ranking over the elements of  $h$ , and that the DM cannot obtain new information regarding elements in the exploration history (preferences over consumed elements are stationary). That is, the axiom rules out (i) any learning (objective or subjective) which is not the direct result of consumption, and (ii) learning models that allow for residual uncertainty over outcomes that have been consumed.

It is also worth noting that although the axiom is existential, it is still falsifiable. Since strategic planning concerns only finite collections,  $W$ , of IHCPs, there is a finite number of assignments in  $As(W, h)$ . Nonetheless, in light of the observability problems associated with existential axioms, we offer an alternative axiomatization in Appendix C.

We also assume the DM is a consequentialist: she cares only about the path of consumption that actually occurs. To impose this, we restrict that the continuation value for IHCPs after learning an event depends only on the consumption in that event, and not what would have been consumed had a different event occurred. Relating back to observables, this implies the maximizing elements of a strategic plan depend only on the conditioning sequence and not on the menus offered following such sequences. The next axiom, closely resembling the Weak Axiom of Revealed Preference from standard choice theory, creates such a restriction.

[A6: CONSEQUENTIALISM (CSQ)]. For all  $W, V \subseteq Z$ , if there exist strategic plans  $p^W$  and  $p^V$ , such that  $p^W(\rho), p^V(\rho) \in W \cap V$  for some  $\rho \in \Sigma(h)$ , then

$$\rho(p^V(\rho)) \cup \bigcup_{\sigma \in \Sigma(h) \setminus \rho} \sigma(p^W(\sigma)) \sim_h \bigcup_{\sigma \in \Sigma(h)} \sigma W$$

CSQ states that if  $w$  is the maximal element in  $W$  conditional on choosing  $\sigma$ , and  $w$  is available in  $V$ , but some other alternative  $v$  was chosen, then  $v$  must be as good as  $w$ .

### 3.3.4 Interim Preferences

Using the language of strategic planning we can define the interim (i.e., state-dependent) preferences. That is, the anticipated preference of the decision maker, conditional on having *learned* that her ranking over  $h$  is the ranking corresponding to the sequence  $\sigma \in \Sigma(h)$ ,  $\succsim_{h|\sigma}$ .

**Definition.** For each  $h$  and  $\sigma \in \Sigma(h)$ , let  $\succsim_{h|\sigma}$  be defined by  $z \succsim_{h|\sigma} w$  if and only if there exists a strategic plan  $p : \Sigma(h) \rightarrow \{z, w\}$  such that  $p(\sigma) = z$ .

	$\{a, b\} = a$	$\{a, c\}$	$\{b, c\}$	$\sigma$
$a \succ b \succ c$	$a$	$a$	$b$	$ab$
$a \succ c \succ b$	$a$	$a$	$c$	$ac$
<del><math>b \succ a \succ c</math></del>	<del><math>b</math></del>	<del><math>a</math></del>	<del><math>b</math></del>	<del><math>ab</math></del>
<del><math>b \succ c \succ a</math></del>	<del><math>b</math></del>	<del><math>c</math></del>	<del><math>b</math></del>	<del><math>cb</math></del>
$c \succ a \succ b$	$a$	$c$	$c$	$cc$
<del><math>c \succ b \succ a</math></del>	<del><math>b</math></del>	<del><math>c</math></del>	<del><math>c</math></del>	<del><math>cc</math></del>

Figure 7: Example of  $\Sigma(h'|\sigma)$

Although we do not have the language to directly observe the DM's hypothetical preferences, we can infer them by constructing  $\succsim_{h|\sigma}$ , which is fully characterized by  $\succsim$ .

As the decision maker continues to consume, her interim preference will change. As such, we can consider preferences of the form  $\succsim_{h'|\sigma}$  where  $\sigma \in \Sigma(h)$  and  $h' \supset h$ . These preferences are defined in the same way as  $\succsim_h$  was defined from  $\succsim$ .

**Definition.** For each  $h, \sigma \in \Sigma(h)$  and  $h' \supset h$  let  $\succsim_{h'|\sigma}$  be defined by

$$z \succsim_{h'|\sigma} w \iff h'z \succsim_{h|\sigma} h'w$$

Since we will want also to define strategic plans from the point of view of interim preferences (to examine the DM's anticipated future learning conditional on past learning), it will be helpful to define  $\Sigma(h'|\sigma)$  for  $h' \supset h$  and  $\sigma \in \Sigma(h)$ .  $\Sigma(h'|\sigma)$  is the subset of  $\Sigma(h')$  that is consistent with  $\sigma$ . That is, defined by orders that respect the ordering that generated  $\sigma$  and leaving out the redundant pairs. For example, if  $h = \{a, b\}$  and  $h' = \{a, b, c\}$  and  $\sigma = a$  then  $\Sigma(h'|\sigma) = \{ab, ac, cc\}$ , as shown in Figure 7.

Our first result, which is instrumental in constructing the LTC representation, provides the recursive structure of the preferences. The conditional preference,  $\succsim_{h'|\sigma}$ , inherits all of the structure imposed on  $\succsim_h$ ; it satisfies *all* of the previous axioms.

**Theorem 2.** Let  $\succsim$  satisfy HST. If  $\succsim$  satisfies ORD, STA, FLX, CON, PLN, and CSQ, then for each  $h, \sigma \in \Sigma(h)$ , and  $h \subseteq h'$ ,  $\succsim_{h'|\sigma}$  satisfies ORD, STA, FLX, CON, PLN, and CSQ.

This result allows us to obtain the branching, recursive representation. Even as the preferences change, their structure –and so, the functional form of the representation– remains intact. Additionally, Theorem 2 indicates that there is nothing special about the ex-ante preference,  $\succsim$ , other than being associated with a particular state space. If preferences were elicited at a later date, when some items had been consumed, the resulting axiomatic structure and representation would be identical, but for a smaller (conditional) state space.

Our last axiom imposes consistency on the way the DM's preferences over outcomes (i.e., over  $X$ ) can change as she learns. In particular we impose that the value of a singleton is constant from an ex-ante point of view. This implies that the value a decision maker places on an outcome  $a$  is the aggregated value she places on  $a$  at each state which she considers possible. We want this to hold, not only at period zero, but also for each interim preference, when looking forward. We also impose that there is a best and a worst outcome that are ex-ante identified. This assumption is simply a convenient way to calibrate states to one another.

**Definition.**  $\bar{a}$  and  $\underline{a}$  are universal best and worst elements, if  $\bar{\rho} = \times^\infty \bar{a}$  and  $\underline{\rho} = \times^\infty \underline{a}$  are such that, for all  $z, w \in Z$  and all feasible decision sequences given  $z$ :  $z_{-\sigma} \bar{\rho} \succcurlyeq z_{-\sigma} w \succcurlyeq z_{-\sigma} \underline{\rho}$ .

[A7: SINGLETON RECURSIVITY (RCV)]. There exist universal best and worst element,  $\bar{a}$  and  $\underline{a}$ , such that  $\times^\infty \{\bar{a}, \underline{a}\}$  is order-dense in  $Z$ . Moreover,

$$\begin{aligned} \rho \succcurlyeq_{h|\sigma} \tau &\iff \rho \succcurlyeq_{h'|\sigma} \tau \\ \nu \succcurlyeq \pi &\iff \nu \succcurlyeq_{h'|\sigma} \pi \end{aligned}$$

for all  $\sigma \in \Sigma(h'')$  such that  $h'' \subseteq h \cap h'$ ,  $\rho, \tau \in \times^\infty X$ , and  $\pi, \nu \in \times^\infty \{\bar{a}, \underline{a}\}$ .

Even though a decision maker expects to learn, a pre-determined stream of consumption offers no way to capitalize on the new information. Hence, the DM's value of a stream of singleton prizes is the aggregation of her value in each state she still considers possible. Similarly, since the DM's value over streams concerning  $\bar{a}$  and  $\underline{a}$  are ex-ante known, her preference regarding such streams does not change, even when conditioned on particular information.

**Remark 1.** In applications, when information is exogenous and we assume that the agent is learning as she explores the different options, introducing flexibility is not necessary. In such environments we assume that we have access to what the agent expects to learn and what action she will take after learning every bit of information (i.e., the information structure is exogenous, and the actions are statistically optimal). Had the agent needed to commit to an action stream (consumption, in our case), she would not be able to capitalize on information she expects to learn as time progresses, and she would maximize her ex-ante expected utility. Axiom RCV states exactly that. When the agent needs to commit to a consumption stream, she calculates her ex-ante utility. As modelers, in order to observe what the agent expects to learn, we need to offer her flexibility so she can condition her future consumption on the different pieces of information she expects to learn. This is a feature of most axiomatic learning models.

## 4 Learning Through Consumption Representation

In this section we present the main results of the paper. We start by providing the formal *Learning Through Consumption (LTC)* functional form, formulate the representation result and, finally, present our uniqueness result.

### 4.1 The Functional Form

The LTC representation was introduced as equation (LTC) in Section 3. We here give the formal definition.

**Definition.** The tuple,  $\mathcal{S} = \{S, \{\mathcal{P}(h)\}_{h \subseteq X}\}$ , is an **LTC information structure** if

- S1.  $S \subseteq \text{SO}r(X)$ .
- S2. For each  $h$ ,  $\mathcal{P}(h)$  is a partition of  $S$ .
- S3. If  $h' \supseteq h$  then  $\mathcal{P}(h')$  is a refinement of  $\mathcal{P}(h)$ .

S4.  $\mathcal{P}(\emptyset) = \{S\}$  and  $\mathcal{P}(X) = S$ .

Moreover,<sup>13</sup> for  $P \in \mathcal{P}(h)$  and  $h' \supset h$ , define  $\mathcal{P}(h'|P) = \{P' \in \mathcal{P}(h') : P' \cap P \neq \emptyset\}$ , and  $\mathcal{P} = \bigcup_{h \subseteq X} \mathcal{P}(h)$ .

Since ultimately states will be identified with rankings over  $X$ , the first restriction, S1 simply ensures there are no duplicate states. Of course, it further restricts that states are *strict* orderings, but, this assumption can be relaxed so that  $S$  permits weak orderings as described in Appendix B. Loosely speaking, (S2) dictates that the DM has in mind some subjective state space,  $S$ , and that learning is always with respect to  $S$ . Restriction (S3) dictates that the process of learning has perfect recall and that signals are never contradictory. Finally, (S4) restricts attention to the case where everything can be learned and nothing is known at the outset. This last restriction will be without loss of generality given our target representation, since it merely entails a relabeling of states.

Note, the definition of an information structure is more general than the intuition provided in Section 2.2, in the sense that following every exploration history  $h$ ,  $\mathcal{P}(h)$  does not need to correspond to the natural partition of  $S$  induced by all strict orderings over  $h$ . This is rectified by Theorem 4; its proof pins down the structure of the partitions and shows it follows exactly the intuition provided earlier.

Given an LTC information structure,  $\mathcal{S}$ , an LTC representation of  $\succsim$  is the pair  $(\mathcal{S}, \mathcal{U})$ , where  $\mathcal{U}$  is a set of functions with a desired recursive formulation and consistent with the information structure  $\mathcal{S}$ . Specifically:

**Definition.** Let  $\mathcal{U} = \{\{u_P\}_{P \in \mathcal{P}}, \{\phi_{h'|P}\}_{h' \subseteq X, P \in \mathcal{P}}, F\}$ , then we say  $(\mathcal{S}, \mathcal{U})$  is an **LTC representation** of  $\succsim$  if  $\mathcal{S}$  is an LTC information structure and

U1.  $u_P : X \rightarrow \mathbb{R}$  with  $u_P(\bar{a}) \geq u_P(a) \geq u_P(\underline{a})$  for all  $a \in X$ ,  $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ , strictly increasing and continuous in its second component, and  $\phi_{h'|P} : \mathbb{R}^{|\mathcal{P}(h'|P)|} \rightarrow \mathbb{R}$ , strictly increasing for all  $P \in \mathcal{P}(h)$  for all  $h \subseteq h'$ .

U2.  $U_{\emptyset|S} : Z \rightarrow \mathbb{R}$  represents  $\succsim$  and is defined recursively by

$$U_{h'|P}(z) = \phi_{h'|P} \left( \max_{i \in M(z)} F(u_Q(a_i), U_{h' \cup a_i|Q}(z_i)) \right)_{Q \in \mathcal{P}(h'|P)},$$

such that the image of  $\times^\infty \{\bar{a}, \underline{a}\}$  is dense in the image of  $Z$ .

U3.  $u_P = \phi_{h'|P}(u_Q)_{Q \in \mathcal{P}(h'|P)}$ .

U4. For any  $a, b \in h$  and  $P \in \mathcal{P}(h)$ ,  $u_P(a) > u_P(b) \iff (a, b) \in s$  for any  $s \in P$ .

Restriction (U1) simply defines the nature of the functions involved and (along with (U2)) states that  $F$  and  $\phi$  aggregate preferences (and are strictly increasing). The most involved restriction, (U2), dictates that this family of functions represents the preferences. This provides the particular recursive structure to the preferences. To see how the recursive definition of  $U_{h'|P}$  can represent the primitive, observe, if  $h' = h$  then  $\mathcal{P}(h'|P) = \mathcal{P}(h|P) = P$ , and the corresponding aggregator,  $\phi_{h'|P}$ , is the identity mapping. So for all  $h$  and  $P \in \mathcal{P}(h)$ ,  $U_{h|P}(z) = \max_{i \in M(z)} F(u_P(a_i), U_{h \cup a_i|P}(z_i))$ . From this observation, we have  $U_{\emptyset|S}(z) = \max_{i \in M(z)} F(u_S(a_i), U_{a_i|S}(z_i))$ , where  $U_{a_i|S}(z_i)$  is defined by (U2). (U3) ensures that the DM, given some residual uncertainty, evaluates consumption prizes as an aggregate of states she considers possible given her current information. Moreover, the method of aggregation coincides with the process by which she evaluates IHCPs. Lastly, (U4) dictates that the DM knows with certainty her ranking over prizes she has

<sup>13</sup>Elements  $P$  are cells of the partition  $\mathcal{P}(h)$  of  $S$ . Cells of the finest partition  $\mathcal{P}(X)$  are states in  $S$ .

already consumed. Note, since for all  $s \in P(h)$  either  $(a, b) \in s$  or  $(b, a) \in s$ , (U4) requires all  $s \in P(h)$  to agree on previously consumed outcomes.

## 4.2 Representation Theorem

With these definitions we can now state our representation theorem.

**Theorem 3.** *The following are equivalent:*

1. *The preference  $\succsim$  satisfies HST, ORD, STA, FLX, CON, PLN, CSQ, and RCV.*
2. *There exists an LTC representation for  $\succsim$ .*

## 4.3 The Uniqueness of the Information Structure

It is well known that the state space in Kreps [1979] is not unique, a hurdle overcome by Dekel et al. [2001] by expanding the domain to include objective risk. In this section we show that, although we consider only ordinal rankings over outcomes, we can identify the state space,  $\mathcal{S}$ , in a meaningful way.

**Theorem 4.** *Assume that  $(\mathcal{S}, \mathcal{U})$  and  $(\mathcal{S}', \mathcal{U}')$  are LTC representations of  $\succsim$ , then  $\mathcal{S} = \mathcal{S}'$ .*

Our consumption space is finite, hence it is not necessary that  $\mathcal{U}$  is unique in a cardinal sense. Nevertheless, Theorem 4 states that the underlying uncertainty, including the partition associated with each consumption history, can be uniquely recovered from  $\succsim$ . Since the notion of learning in this framework is entirely ordinal, the underlying uncertainty is naturally associated with the possible orderings over  $X$  that the DM considers possible ex-ante. It is a direct consequence of Theorems 3 and 4 that the unique set of final states is some subset of  $SO_r(X)$  and the partition induced by exploration history  $h$  is  $\Sigma(h)$ .

The intuition behind the identification is as follows. First, we resort to similar arguments as in Section 2. Applying Theorem 1 allows us to identify the state space  $S$  as the uncertainty the DM considers following the exploration of  $X$  in its entirety. With strategic planning, we identify, following every exploration history  $h$ , the partition of  $S$  induced by the the strict orderings over  $h$ . At this point the additional structure of IHCPs becomes necessary for the identification of the learning through consumption representation. PLN, in the presence of flexibility at every period, allows us to define conditional preferences on a piece of information (i.e.,  $\succsim_{h|\sigma}$ ), and express how the DM *expects to further learn conditional on learning that information*. This also provides the language to ensure the different identified state spaces are consistent with one another in a recursive fashion. Specifically, we apply RCV to make sure the *aggregate* behavior conditional on the different states  $s \in P \in \mathcal{P}(h)$  is identical to the behavior conditional on  $P$  (the more general condition being U3).

**Remark 2.** *In order to fully identify the information structure we need to observe preference for flexibility after every exploration history. This is feasible even when restricted to IHCPs that exhibit no flexibility after “sufficiently many” periods of time. Therefore, one can also consider finite horizon problems if willing to forego recursivity (and thus STA).*

## 5 Related Literature

The literature on subjective states began with Kreps [1979]. In Kreps’ model, the set of future preference profiles the DM considers possible is identified by examining her preferences over *menus* of consumption

prizes. This framework has since been extended by [Dekel et al. \[2001\]](#) (DLR) to menus of lotteries, where the unique set of cardinal utility functions can be identified. While these models are interpreted with a dynamic component, there is only a single period of consumption.

Recently there have been papers that embed the DLR setup in dynamic settings. [Krishna and Sadowski \[2014\]](#) provide an infinite horizon model where each period the DM’s utility is drawn from a subjective distribution, depending on the current state. The model is not one of learning, and in particular, it is not responsive; the information and period-by-period resolution of uncertainty is unrelated to the choices made by the decision maker. Their representation has a recursive structure, lending itself to examine the intertemporal tradeoff between future constraints and current period consumption. Our model allows for similar contemplation, with the added intertemporal consideration regarding the tradeoff between future information structures and current period consumption, as is standard in models of strategic experimentation.

[Higashi et al. \[2014\]](#) consider a dynamic extension of DLR with the different subjective states having different discount factors. They axiomatize a recursive representation where subjective states evolve over time according to history of past consumption. In their model, however, correlations across subjective states (or, conditional preferences of conditional preferences, etc) are excluded, which is a clear distinction from the methodological point presented here.

In a model of subjective learning, [Dillenberger et al. \[2014\]](#) examine a DM who has preferences over menu-time pairs. This allows for the identification, not only of the set potential preference profiles considered possible, but also of the way that the DM expects to learn over time. As in our model, at each period the DM in [Dillenberger et al. \[2014\]](#) considers a state space that is a refinement of the previous periods’ state spaces. However, in contrast to our model, the path of learning is not responsive, that is, does not depend on the choices of the DM. In addition, their model is a static one and intertemporal considerations are not studied.

The most closely related paper to responsive learning is [Cooke \[2015\]](#), considering a model in which the agent has to rank consumption-menu pairs.<sup>14</sup> Upon first stage consumption, the agent learns the cardinal utility of the consumed element (thus learning is absolute), which helps to inform the choice out of the menu in the second stage. This also implies that learning is partitional, similar to the model presented here. There are two conceptual differences from this paper. First, in the set up [Cooke](#) presents, there is a single period of learning. The conditioning event (that is, first period consumption) is explicitly modeled, rather than identified via choices *from* menus. Second, learning is cardinal, rather than through comparisons as studied here. We elaborate on these differences and their implications in [Section D](#); the advantage of our approach is that it facilitates the study of multiple periods of learning.

Models of rational inattention [[Sims, 2003](#)] share a common feature with models of responsive learning. In both theories, agents make choices of consumption and of information acquisition. Nevertheless, in the former the two choices are taken separately. In models of responsive learning, there is a single (joint) constraint; the act of consumption directly determines how the agent is going to be further informed about the underlying uncertainty. [Ergin and Sarver \[2010\]](#) and [Dillenberger et al. \[2015\]](#) provide a behavioral foundation to a class of (introspective) learning models that include rational inattention with partitional information structures.

Case Based Decision Theory (CBDT), first introduced by [Gilboa and Schmeidler \[1995\]](#), also investigates a decision maker who evaluates her choices based on her experience. Like this paper, CBDT includes multistage programming and learning from past actions (see chapters 5 and 7 of [Gilboa and Schmeidler \[2001\]](#)). Despite

---

<sup>14</sup>[Hyogo \[2007\]](#) raises a somewhat related question and provides an analysis in an environment similar to that of [Cooke \[2015\]](#), resorting to an objective state space.

these conceptual similarities, the formal frameworks differ significantly, as we stay within the confines of (standard) utility maximization. Our notion of learning is based *only* on past consumption, while CDBT considers hypothetical memories, where decision problems are compared according to an abstract notion of similarity. Moreover, our recursive structure allows for a more direct analysis of the tradeoff between the learning and consumption components of each choice.

## A Proofs

### A.1 Preliminary Results

**Lemma 1.** *The decision maker has a weak preference for delaying choices. That is, for all  $z, w, v \in Z$ ,  $\sigma$  feasible with respect to  $v$  and finite decision problems  $A$ ,*

$$v_{-\sigma} A\{z \cup w\} \succsim_h v_{-\sigma} \{Az \cup Aw\}.$$

*Proof.* This is an immediate consequence of the decision makers **FLX**. That is, we know that the left hand side is identically equal to  $v_{-\sigma} \{A\{z \cup w\} \cup A\{z \cup w\}\}$ , which by flexibility and transitivity

$$v_{-\sigma} \{A\{z \cup w\} \cup A\{z \cup w\}\} \succsim_h v_{-\sigma} \{A\{z\} \cup A\{z \cup w\}\} \succsim_h v_{-\sigma} \{A\{z\} \cup A\{w\}\},$$

which is identically equal to the right hand side. ■

**Lemma 2.** *For all  $\sigma \in \Sigma(h)$ , and  $W \subset Z$ , there exists a strategic plan,  $p^W : \Sigma(h) \rightarrow W$ , such that  $p^W(\sigma) = z$  if any only if*

$$\sigma z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho W \sim_h \bigcup_{\rho \in \Sigma(h)} \rho W.$$

*Proof.* First, assume such a strategic plan exists. So we have

$$\begin{aligned} \bigcup_{\rho \in \Sigma(h)} \rho W &\succsim_h \sigma z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho W \\ &\succsim_h \sigma z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho(p^W(\rho)) \\ &\sim_h \bigcup_{\rho \in \Sigma(h)} \rho W, \end{aligned}$$

where both weak preferences are consequences of **FLX** and the indifference relation by **PLN** and our assumption.

Now assume that

$$\sigma z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho W \sim_h \bigcup_{\rho \in \Sigma(h)} \rho W. \quad (\text{A.1.8})$$

By **PLN** we know there exists some  $p^W : \Sigma(h) \rightarrow W$  such that

$$\bigcup_{\rho \in \Sigma(h)} \rho p^W(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho W. \quad (\text{A.1.9})$$

Consider the collection  $V = \{W, w_i\}_{w_i \in W}$  (In an effort to be clear,  $V = \{\{\bigcup_{w \in W} w\}, w_1, w_2, \dots, w_n\}$ ). It is

immediate that  $\bigcup_{w \in W} w = \bigcup_{v \in V} v$  and so, utilizing this identity, we can rewrite (A.1.8) and (A.1.9):

$$\sigma z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho W \sim_h \bigcup_{\rho \in \Sigma(h)} \rho V, \quad (\text{A.1.10})$$

$$\bigcup_{\rho \in \Sigma(h)} \rho p^W(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho V. \quad (\text{A.1.11})$$

So, by (A.1.10) and (A.1.11) we satisfy the conditions for CSQ and so:

$$\sigma z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho p^W(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho V,$$

or, since  $\bigcup_{w \in W} w = \bigcup_{v \in V} v$ ,

$$\sigma z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho p^W(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho W.$$

■

**Lemma 3.** *Let  $h \subset h'$ . The collection of functions,  $(p_\sigma : \Sigma(h'|\sigma) \rightarrow W)_{\sigma \in \Sigma(h)}$  constitute a collection of strategic plans for the corresponding preference relations  $(\succsim_{h'|\sigma})_{\sigma \in \Sigma(h)}$  if and only if the map*

$$p : \Sigma(h') \longrightarrow W$$

$$p : \sigma \tau \mapsto p_\sigma(\tau)$$

is a strategic plan for  $\succsim_{h'}$ . Where  $\sigma \in \Sigma(h)$  and  $\tau \in \Sigma(h'|\sigma)$ .

*Proof.* Let  $(p_\sigma : \Sigma(h'|\sigma) \rightarrow W)_{\sigma \in \Sigma(h)}$  constitute a collection of strategic plans for the corresponding preference relations  $(\succsim_{h'|\sigma})_{\sigma \in \Sigma(h)}$ . Then we know that, for each  $\sigma \in \Sigma(h)$ :

$$\bigcup_{\tau \in \Sigma(h'|\sigma)} \tau p_\sigma(\tau) \sim_{h'|\sigma} \bigcup_{\tau \in \Sigma(h'|\sigma)} \tau W.$$

Using the definition of  $\sim_{h'|\sigma}$  this is equivalent to

$$(h' \setminus h) \bigcup_{\tau \in \Sigma(h'|\sigma)} \tau p_\sigma(\tau) \sim_{h|\sigma} (h' \setminus h) \bigcup_{\tau \in \Sigma(h'|\sigma)} \tau W. \quad (\text{A.1.12})$$

Hence, we know that there must exist two strategic plans (with respect the  $\succsim_h$ ) such that each side of (A.1.12) gets assigned to  $\sigma$  in one of the two plans. Then, utilizing CSQ to combine these strategic plans, we know

$$\bigcup_{\sigma \in \Sigma(h)} \sigma (h' \setminus h) \bigcup_{\tau \in \Sigma(h'|\sigma)} \tau p_\sigma(\tau) \sim_h \bigcup_{\sigma \in \Sigma(h)} \sigma (h' \setminus h) \bigcup_{\tau \in \Sigma(h'|\sigma)} \tau W. \quad (\text{A.1.13})$$

Since by FLX, the right-hand side of (A.1.13) is indifferent to

$$\bigcup_{\sigma \in \Sigma(h)} \sigma (h' \setminus h) \bigcup_{\tau \in \Sigma(h'|\sigma)} \{\tau W, \tau p_\sigma(\tau)\},$$

CON implies

$$\bigcup_{\sigma \in \Sigma(h)} \sigma \bigcup_{\tau \in \Sigma(h'|\sigma)} \tau p_\sigma(\tau) \sim_{h'} \bigcup_{\sigma \in \Sigma(h)} \sigma \bigcup_{\tau \in \Sigma(h'|\sigma)} \{\tau W, \tau p_\sigma(\tau)\}.$$

Again by FLX,

$$\bigcup_{\sigma \in \Sigma(h)} \sigma \bigcup_{\tau \in \Sigma(h'|\sigma)} \tau p_\sigma(\tau) \sim_{h'} \bigcup_{\sigma \in \Sigma(h)} \sigma \bigcup_{\tau \in \Sigma(h'|\sigma)} \tau W,$$



which is equivalent to

$$\bigcup_{\sigma \in \Sigma(h)} \bigcup_{\tau \in \Sigma(h'|\sigma)} \sigma \tau p_\sigma(\tau) \sim_{h'} \bigcup_{\sigma \in \Sigma(h)} \bigcup_{\tau \in \Sigma(h'|\sigma)} \sigma \tau W.$$

And so, the specified map,  $p$ , is indeed a strategic plan for  $\succcurlyeq_{h'}$ . Each step is a bi-directional implication so this proves the if and only if statement.  $\blacksquare$

## A.2 Proof of Theorem 2

Fix some  $h, h' \supset h$  and some  $\sigma \in \Sigma(h)$ .

(WEAK ORDER). That  $\succcurlyeq_{h'|\sigma}$  is complete is immediate from the definition. Now assume that  $z \succcurlyeq_{h'|\sigma} w$  and  $w \succcurlyeq_{h'|\sigma} v$ . Let  $\hat{V} = \{h'z, h'w\}$  and  $\bar{V} = \{h'w, h'v\}$ . By the definition of  $\succcurlyeq_{h'|\sigma}$  there exist strategic plans,  $p^{\hat{V}} : \Sigma(h) \rightarrow \hat{V}$ , and  $p^{\bar{V}} : \Sigma(h) \rightarrow \bar{V}$ , such that  $p^{\hat{V}}(\sigma) = h'z$  and  $p^{\bar{V}}(\sigma) = h'w$ .

Define  $V = \{h'z, h'w, h'v\}$  and let  $p^V : \Sigma(h) \rightarrow V$  be the plan ensured by PLN. We claim that there exists some strategic plan,  $q^V$ , thereover, such that  $q^V(\sigma) = h'z$ . There are three cases: Case 1:  $p^V(\sigma) = h'z$ . Then the claim holds. Case 2:  $p^V(\sigma) = h'w$ , then since  $p^{\hat{V}}(\sigma) = h'z$  we satisfy the conditions of CSQ:

$$\sigma h'z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho p^V(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho \{h'z \cup h'w \cup h'v\}, \quad (\text{A.2.1})$$

which defines the desired plan. Case 3:  $p^V(\sigma) = h'v$ , we can iterate the process used in case 2 to prove the claim.

Finally, PLN ensures that there exists a plan  $p^W$  over  $W = \{h'z, h'v\}$ . If  $p^W(\sigma) = h'z$  then it cannot be that  $h'v \succ_{h|\sigma} h'z$  and we are done. If  $p^W(\sigma) = h'v$ , then the fact that  $p^V(\sigma) = h'z$  and CSQ provide a new strategic plan  $q^W(\sigma) = h'z$ , so again it cannot be that  $h'v \succ_{h|\sigma} h'z$ .

(CONTINUITY). Let  $\{z_n\}_{n \in \mathbb{N}}$  be a convergent sequence in  $Z$ , with limit point  $z$ , such that  $z_n \succcurlyeq_{h'|\sigma} w$  for all  $n$  and some  $w \in Z$ . If it is the case that  $z \succcurlyeq_{h'|\sigma} z_n$  for some  $n \in \mathbb{N}$ , then by the transitivity of  $\succcurlyeq_{h'|\sigma}$  we have that  $z \succcurlyeq_{h'|\sigma} w$ . So assume that for all  $n$ ,  $z_n \succcurlyeq_{h'|\sigma} z$ . By Lemma 2 (and CSQ) this implies that:

$$\sigma h'z_n \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho \{h'z_n \cup h'z \cup h'w\} \sim_h \bigcup_{\rho \in \Sigma(h)} \rho \{h'z_n \cup h'z \cup h'w\}.$$

Taking the limit, the continuity of  $\succcurlyeq_h$  provides:

$$\sigma h'z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho \{h'z \cup h'w\} \sim_h \bigcup_{\rho \in \Sigma(h)} \rho \{h'z \cup h'w\},$$

implying that  $z \succcurlyeq_{h'|\sigma} w$ . Hence the contour sets of  $\succcurlyeq_{h'|\sigma}$  are closed.

(H-STATIONARITY). Let  $z \succcurlyeq_{h'|\sigma} w$ , and  $A$  be a finite decision problem which is contained in  $h' \cup \bigcap_{\sigma \in n(z) \cup n(w)} \sigma$  for some  $n$ . So, there exists a strategic plan,  $p^V : \Sigma(h) \rightarrow V$ , with  $V = \{h'z, h'w\}$ , such that

$$\hat{z} = \sigma h'z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho \{h'z, h'w\} \sim_h \bigcup_{\rho \in \Sigma(h)} \rho \{h'z, h'w\} = \hat{w}.$$

Let  $m = |\sigma| + |h'|$ . Consider  $H = h \cup \bigcap_{\sigma \in [n+m](\hat{z}) \cup [n+m](\hat{w})} \sigma$ . For any  $a \in \text{supp}(A)$  we claim  $a \in H$ .

There are three cases: (1)  $a \in h$ , in which case we are done since  $h \subset H$ . (2)  $a \in h' \setminus h$  in which case it is in the first  $m$  periods of  $\hat{z}$  and  $\hat{w}$  (by construction of the above plans you *have* to consume  $h'$ ). (3)  $a \in \bigcap_{\sigma \in n(z) \cup n(w)} \sigma$  in which case it is simply pushed back  $m$  periods (by construction of the above plans you *have* to consume either the first  $n$  periods of  $z$  or the first  $n$  periods of  $w$ ). Hence, by STA we can interject

A after  $n + m$  periods. Letting  $W = \{h'z_{-n}Az_n, h'w_{-n}Aw_n\}$  this gives

$$\sigma h'z_{-n}Az_n \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho W \sim_h \bigcup_{\rho \in \Sigma(h)} \rho W,$$

so, by Lemma 2,

$$z_{-n}Az_n \succsim_{h'|\sigma} w_{-n}Aw_n.$$

The converse holds from the bi-directional implication of each step.

Now, for some  $v$ , and  $\rho$  feasible for  $v$ , let  $B$  be a finite decision problem with  $\text{supp}(B) \subseteq h \cup \rho$ . Call  $V = \{h'v_{-\pi}B\{z \cup w\}, h'v_{-\sigma}\{Bz \cup Bw\}\}$ . There exists a strategic plan  $p^V : \Sigma(h) \in V$ . Assume that  $p^V(\sigma) = h'v_{-\pi}B\{z \cup w\}$ . So

$$\sigma h'v_{-\pi}B\{z \cup w\} \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho p^V(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho V,$$

but by (2) of STA we have

$$\sigma h'v_{-\pi}\{Bz \cup Bw\} \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho p^V(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho V,$$

and so,  $v_{-\pi}B\{z \cup w\} \succsim_{h'|\sigma} v_{-\sigma}\{Bz \cup Bw\}$ . Of course, if we assume  $p^V(\sigma) = h'v_{-\pi}\{Bz \cup Bw\}$  the same argument holds. We have  $v_{-\pi}B\{z \cup w\} \sim_{h'|\sigma} v_{-\sigma}\{Bz \cup Bw\}$  as desired.

(PREFERENCE FOR FLEXIBILITY). Assume, to the contrary, that there exists some  $v, z, w \in Z$ , and  $\pi$  feasible for  $v$  such that  $v_{-\pi}z \succ_{h'|\sigma} v_{-\pi}(z \cup w)$ . Let  $V = \{h'v_{-\pi}z, h'v_{-\pi}(z \cup w)\}$ . Lemma 2 therefore implies

$$\sigma h'v_{-\pi}z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho V \succ_h \sigma h'v_{-\pi}(z \cup w) \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho V,$$

an immediate contradiction to the DMs preference for FLX.

(CONSISTENT FLEXIBILITY). Assume that, for some  $A \subseteq \times_{i=1}^N h'$ , finite  $\pi$  such that  $h' \cup \pi = h''$  and  $f, g : A \rightarrow Z$  we have

$$\bigcup_{\tau \in A} \tau \pi f(\tau) \sim_{h'|\sigma} \bigcup_{\tau \in A} \tau \pi \{f(\tau) \cup g(\tau)\}. \quad (\text{A.2.2})$$

Define,  $V = \{h' \bigcup_{\tau \in A} \tau \pi f(\tau), h' \bigcup_{\tau \in A} \tau \pi \{f(\tau) \cup g(\tau)\}\}$ . Then, by the definition of  $\succsim_{h'|\sigma}$ , CSQ, and lemma 2 we know that there exists some strategic plan  $p^V : \Sigma(h) \rightarrow V$  such that

$$\sigma h' \bigcup_{\tau \in A} \tau \pi f(\tau) \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho p^V(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho V, \quad (\text{A.2.3})$$

$$\sigma h' \bigcup_{\tau \in A} \tau \pi \{f(\tau) \cup g(\tau)\} \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho p^V(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho V. \quad (\text{A.2.4})$$

Define the mapping  $\hat{f} : \Sigma(h)A \rightarrow Z$  as

$$\rho \tau \mapsto \begin{cases} f(\tau) & \text{if } \rho = \sigma \\ f(\tau) & \text{if } \rho \neq \sigma \text{ and } p^V(\rho) = \bigcup_{\tau \in A} \tau \pi f(\tau) \\ \{f(\tau) \cup g(\tau)\} & \text{if } \rho \neq \sigma \text{ and } p^V(\rho) = \bigcup_{\tau \in A} \tau \pi \{f(\tau) \cup g(\tau)\}, \end{cases}$$

and the mapping  $\hat{g} : \Sigma(h)A \rightarrow Z$  as

$$\rho\tau \mapsto \begin{cases} \{f(\tau) \cup g(\tau)\} & \text{if } \rho = \sigma \\ f(\tau) & \text{if } \rho \neq \sigma \text{ and } p^V(\rho) = \bigcup_{\tau \in A} \tau\pi f(\tau) \\ \{f(\tau) \cup g(\tau)\} & \text{if } \rho \neq \sigma \text{ and } p^V(\rho) = \bigcup_{\tau \in A} \tau\pi\{f(\tau) \cup g(\tau)\}. \end{cases}$$

Then, we can re-write the (A.2.3) and (A.2.4) as

$$\bigcup_{\rho \in \Sigma(h)} \rho h' \bigcup_{\tau \in A} \tau\pi \hat{f}(\rho\tau) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho h' \bigcup_{\tau \in A} \tau\pi\{\hat{f}(\rho\tau) \cup \hat{g}(\rho\tau)\}, \quad (\text{A.2.5})$$

$$\bigcup_{\rho \in \Sigma(h)} \rho h' \bigcup_{\tau \in A} \tau\pi \hat{g}(\rho\tau) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho h' \bigcup_{\tau \in A} \tau\pi\{\hat{f}(\rho\tau) \cup \hat{g}(\rho\tau)\}, \quad (\text{A.2.6})$$

so by CON (expanding  $h$  to  $h'$  by eliminating  $h'^{15}$ )

$$\begin{aligned} \bigcup_{\rho \in \Sigma(h)} \rho \bigcup_{\tau \in A} \tau\pi \hat{f}(\rho\tau) &\sim_{h'} \bigcup_{\rho \in \Sigma(h)} \rho \bigcup_{\tau \in A} \tau\pi\{\hat{f}(\rho\tau) \cup \hat{g}(\rho\tau)\}, \\ \bigcup_{\rho \in \Sigma(h)} \rho \bigcup_{\tau \in A} \tau\pi \hat{g}(\rho\tau) &\sim_{h'} \bigcup_{\rho \in \Sigma(h)} \rho \bigcup_{\tau \in A} \tau\pi\{\hat{f}(\rho\tau) \cup \hat{g}(\rho\tau)\}, \end{aligned}$$

again by CON (expanding  $h'$  to  $h''$  by eliminating  $\pi$ )

$$\bigcup_{\rho\tau \in \Sigma(h)A} \rho\tau \hat{f}(\rho\tau) \sim_{h''} \bigcup_{\rho\tau \in \Sigma(h)A} \rho\tau\{\hat{f}(\rho\tau) \cup \hat{g}(\rho\tau)\}, \quad (\text{A.2.7})$$

$$\bigcup_{\rho\tau \in \Sigma(h)A} \rho\tau \hat{g}(\rho\tau) \sim_{h''} \bigcup_{\rho\tau \in \Sigma(h)A} \rho\tau\{\hat{f}(\rho\tau) \cup \hat{g}(\rho\tau)\}. \quad (\text{A.2.8})$$

Let  $\hat{V} = \{\bigcup_{\tau \in A} \tau f(\tau), \bigcup_{\tau \in A} \tau\{f(\tau) \cup g(\tau)\}\}$  and Let  $\hat{p}^V : \Sigma(h)\setminus\sigma \rightarrow \hat{V}$  as

$$\rho \mapsto \begin{cases} \bigcup_{\tau \in A} \tau\{f(\tau) \cup g(\tau)\} & \text{if } p^V(\rho) = \bigcup_{\tau \in A} \tau\pi\{f(\tau) \cup g(\tau)\} \\ \bigcup_{\tau \in A} \tau f(\tau) & \text{if } p^V(\rho) = \bigcup_{\tau \in A} \tau\pi f(\tau). \end{cases}$$

Then, (A.2.7) and (A.2.8) are rewritten as

$$\begin{aligned} \sigma \bigcup_{\tau \in A} \tau f(\tau) \cup \bigcup_{\rho \in \Sigma(h)\setminus\sigma} \rho \hat{p}^V(\rho) &\sim_{h''} \bigcup_{\rho \in \Sigma(h)} \rho \hat{V}, \\ \sigma \bigcup_{\tau \in A} \tau\{f(\tau) \cup g(\tau)\} \cup \bigcup_{\rho \in \Sigma(h)\setminus\sigma} \rho \hat{p}^V(\rho) &\sim_{h''} \bigcup_{\rho \in \Sigma(h)} \rho \hat{V}. \end{aligned}$$

Again by CON,

$$\begin{aligned} \sigma h'' \bigcup_{\tau \in A} \tau f(\tau) \cup \bigcup_{\rho \in \Sigma(h)\setminus\sigma} \rho h'' \hat{p}^V(\rho) &\sim_h \bigcup_{\rho \in \Sigma(h)} \rho h'' \hat{V}, \\ \sigma h'' \bigcup_{\tau \in A} \tau\{f(\tau) \cup g(\tau)\} \cup \bigcup_{\rho \in \Sigma(h)\setminus\sigma} \rho h'' \hat{p}^V(\rho) &\sim_h \bigcup_{\rho \in \Sigma(h)} \rho h'' \hat{V}. \end{aligned}$$

By appealing to Lemma 1 we have

$$\sigma h'' \bigcup_{\tau \in A} \tau f(\tau) \cup \bigcup_{\rho \in \Sigma(h)\setminus\sigma} \rho h'' \hat{p}^V(\rho) \succcurlyeq_h \bigcup_{\rho \in \Sigma(h)} \rho \{h'' \bigcup_{\tau \in A} \tau f(\tau), h'' \bigcup_{\tau \in A} \tau\{f(\tau) \cup g(\tau)\}\}, \quad (\text{A.2.9})$$

$$\sigma h'' \bigcup_{\tau \in A} \tau\{f(\tau) \cup g(\tau)\} \cup \bigcup_{\rho \in \Sigma(h)\setminus\sigma} \rho h'' \hat{p}^V(\rho) \succcurlyeq_h \bigcup_{\rho \in \Sigma(h)} \rho \{h'' \bigcup_{\tau \in A} \tau f(\tau), h'' \bigcup_{\tau \in A} \tau\{f(\tau) \cup g(\tau)\}\}. \quad (\text{A.2.10})$$

The opposite weak preference preference to (A.2.9) and (A.2.10) are guaranteed by the DMs preference for

<sup>15</sup>This is permissible by defining  $\hat{f} : \rho \mapsto \bigcup_{\tau \in A} \tau\pi \hat{f}(\rho\tau)$ ,  $\hat{g}$  likewise, and  $\hat{h} : \rho \mapsto \bigcup_{\tau \in A} \tau\pi\{\hat{f}(\rho\tau) \cup \hat{g}(\rho\tau)\}$ . Then using CSQ, we can show that the antecedent for CF is satisfied (via  $\hat{f} \cup \hat{h}$  for (A.2.5) and  $\hat{g} \cup \hat{h}$  for (A.2.6)).

FLX. Hence (A.2.9) and (A.2.10) hold with indifference and thus define a pair of strategic plans indicating the conditional preference:

$$h'' \bigcup_{\tau \in A} \tau f(\tau) \sim_{h|\sigma} h'' \bigcup_{\tau \in A} \tau \{f(\tau) \cup g(\tau)\},$$

or

$$\bigcup_{\tau \in A} \tau f(\tau) \sim_{h''|\sigma} \bigcup_{\tau \in A} \tau \{f(\tau) \cup g(\tau)\},$$

as desired. The converse is also true by the same argument.

(STRATEGIC PLANNING). We need only show that there exists a strategic plan for any 2 element subset of  $Z$ : induction and the arguments used above can extend to any finite set. So take some  $z, w \in Z$ . We want to show that there exists some  $\hat{p}^{\{z,w\}} : \Sigma(h'|\sigma) \rightarrow \{z, w\}$  such that

$$\bigcup_{\tau \in \Sigma(h'|\sigma)} \tau \hat{p}^{\{z,w\}}(\tau) \sim_{h'|\sigma} \bigcup_{\tau \in \Sigma(h'|\sigma)} \tau \{z \cup w\}.$$

From strategic planning (as applied to  $\succcurlyeq_{h'}$ ), we know that there exists some  $p^{\{z,w\}} : \Sigma(h') \rightarrow \{z, w\}$  such that

$$\bigcup_{\pi \in \Sigma(h')} \pi p^{\{z,w\}}(\pi) \sim_{h'} \bigcup_{\pi \in \Sigma(h')} \pi \{z \cup w\}.$$

Lemma 3 completes the claim.

(CONSEQUENTIALISM). For some finite  $W, V \subset Z$  let  $\hat{p}^W : \Sigma(h'|\sigma) \rightarrow W$  be a strategic plan (over  $\succcurlyeq_{h'|\sigma}$ ). Assume that  $\hat{p}^W(\rho) \in W \cap V$ . So, by lemma 3, there exists a strategic plan,  $p^W : \Sigma(h') \rightarrow W$  such that  $p^W(\sigma\rho) = \hat{p}^W(\rho)$ . Let  $\hat{p}^V : \Sigma(h'|\sigma) \rightarrow V$  be some other strategic plan such that  $\hat{p}^V(\rho) \in W \cap V$ . Again we have an extension such that  $p^V(\sigma\rho) = \hat{p}^V(\rho)$ .

So by CSQ the function

$$\bar{p}^W(\pi) = \begin{cases} p^V(\pi) & \text{if } \pi = \sigma\rho \\ p^W(\pi) & \text{if } \pi \neq \sigma\rho \end{cases}$$

is also a strategic plan for  $W$ , according to  $\succcurlyeq_{h'}$ . Applying lemma 3 again provides the result.

### A.3 Proof of Theorem 3

**Step-1: Recursive Structure on Consumption Streams.** Let us begin by considering  $\hat{\succcurlyeq}$ , the restriction of  $\succcurlyeq$  to degenerate IHCPs that assign pre-determined streams of consumption. Since the set of such IHCPs is closed in  $Z$ , continuity is inherited by  $\hat{\succcurlyeq}$ . Let  $\hat{U} : \times_{i=1}^{\infty} X \rightarrow \mathbb{R}$  be the numerical representation of  $\hat{\succcurlyeq}$ . Further, note that RCV is equivalent on this domain to fully stationary preferences a la Koopmans [1960]. Thus, there exists a function  $u : X \rightarrow \mathbb{R}$  and a strictly increasing function, continuous in its second argument  $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  such that

$$\hat{U}(\sigma) = F[u(\sigma_1), \hat{U}(\{\sigma\}_{n=2}^{\infty})]. \quad (\text{A.3.1})$$

**Step-2: From  $\times^{\infty} X$  to  $Z$ .** Let  $\bar{U}$  be a continuous representation of  $\succcurlyeq$ . Since  $\bar{U}$  also represents  $\times^{\infty} X$ , there exists some continuous, strictly increasing,  $\bar{\psi} : \bar{U}[\times^{\infty} X] \rightarrow \mathbb{R}$  such that  $\bar{\psi}(\bar{U}(\sigma)) = \hat{U}(\sigma)$ . Note that  $\bar{U}[\times^{\infty} X]$  is compact by its closure in  $Z$  and the continuity of  $\bar{U}$ . So, according to Husseinov

[2010] there exists a continuous, strictly increasing  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  such that  $\bar{\psi}(x) = \psi(x)$  for  $x \in \bar{U}[\times^\infty X]$ . Define  $U = \psi \circ \bar{U}$ .  $U$  is a continuous representation of  $\succsim$  that coincides with (A.3.1) over elements of  $\times^\infty X$ .

For each  $x \in \mathbb{R}$  we can define the function  $F_x : \mathbb{R} \rightarrow \mathbb{R}$  by  $F_x(y) = F[x, y]$ . Notice that for any  $u(a)$ ,  $F_{u(a)}$  is continuous and strictly monotone over the interval  $[U(\underline{\rho}), U(\bar{\rho})]$ , where  $\underline{\rho} = \times^\infty \underline{a}$  and  $\bar{\rho} = \times^\infty \bar{a}$ . Hence,  $F_{u(a)}[U(\underline{\rho}), U(\bar{\rho})] = [U(a\underline{\rho}), U(a\bar{\rho})]$ , with identification via (A.3.1). Using these couriering functions, we can define  $U_a : Z \rightarrow \mathbb{R}$  for each  $a \in X$  as

$$U_a : z \mapsto F_{u(a)}^{-1}(U(az)).$$

which is well defined since  $U(az) \in [U(a\underline{\rho}), U(a\bar{\rho})]$  by definition of universal outcomes, and hence in the image of  $F_{u(a)}$ .

The functional  $U_a$  represents  $\succsim_a$  by construction. Indeed,

$$\begin{aligned} z \succsim_a w &\iff az \succsim aw \\ &\iff U(az) \geq U(aw) \\ &\iff F_{u(a)}^{-1}(U(az)) \geq F_{u(a)}^{-1}(U(aw)). \end{aligned}$$

Note that for consumption streams,  $U_a(\sigma) = U(\sigma)$  by (A.3.1). Now given any  $U_h$ , we can define  $U_{h \cup a}$  using the same map,

$$U_{h \cup a} : z \mapsto F_{u(a)}^{-1}(U_h(az)),$$

to inductively define all hypothetical preferences. Note, of course, that we have:

$$U_h(az) = F[u(a), U_{h \cup a}(z)]. \tag{A.3.2}$$

Notice that if  $a \in h$  then  $U_{h \cup a} = U_h$ . To see this assume that for some  $z$ ,  $U_h(z) > U_{h \cup a}(z)$ . Now by the density of streams of universal outcomes, there exists some  $\sigma$  such that  $U_h(z) > U_h(\sigma) = U_{h \cup a}(\sigma) > U_{h \cup a}(z)$ . But this is a direct contradiction of STA.

**Step-3: The existence of a Krepsian State Space.** Let  $\hat{\Sigma}(h) \subseteq \Sigma(h)$  denote the set such that  $\succsim_{h|\sigma}$  is a non-trivial preference. Theorem 2 states that there for each  $\sigma \in \hat{\Sigma}(h)$  the corresponding preference relation  $\succsim_{h|\sigma}$  is a continuous weak order, and therefore representable by a continuous value function  $U_{h|\sigma} : Z \rightarrow \mathbb{R}$ . We claim that these states form a Krepsian state space.

**Lemma 4.** *Let  $z = (a_i, z_i)_{i \in M(z)}$  be a choice problem and  $\sigma \in \Sigma(h)$ . There exists some  $i \in M(z)$  such that  $(a_i, z_i) \sim_{h|\sigma} z$ .*

*Proof.* We first claim that the lemma holds if  $z$  is a finite set. By identifying  $z$  as a collection of singleton menus  $W = \{(a_i, z_i) \mid i \in M(z)\}$ , the Strategic Planning axiom ensures that there exist  $(a_i^\sigma, z_i^\sigma) \in z$ ,  $\sigma \in \Sigma(h)$  such that

$$\bigcup_{\sigma \in \Sigma(h)} \sigma\{(a_i^\sigma, z_i^\sigma)\} \sim_h \bigcup_{\sigma \in \Sigma(h)} \sigma z.$$

By Lemma 2,

$$\sigma\{(a_i^\sigma, z_i^\sigma)\} \cup \bigcup_{\rho \in \Sigma(h), \rho \neq \sigma} \rho z \sim_h \bigcup_{\sigma \in \Sigma(h)} \sigma z,$$

which implies that there exists a strategic plan for  $U = \{z, \{(a_i^\sigma, z_i^\sigma)\}\}$  such that  $\{(a_i^\sigma, z_i^\sigma)\}$  is assigned to  $\sigma$ . By definition of  $\succsim_{h|\sigma}$ ,  $(a_i^\sigma, z_i^\sigma) \succsim_{h|\sigma} z$ . On the other hand, by Preference for flexibility,  $z \succsim_{h|\sigma} (a_i^\sigma, z_i^\sigma)$ . Hence, we have  $(a_i^\sigma, z_i^\sigma) \sim_{h|\sigma} z$ , as desired.

Take any  $z \in Z$ . Since  $Z$  is metrizable with the Hausdorff metric, there exists a sequence  $z^n \in Z$  such that  $z^n \rightarrow z$  and  $z^n$  is a finite subset of  $z$ . By the above claim, there exists  $(a^n, w^n) \in z^n$  such that  $z^n \sim_h (a^n, w^n)$ .

Since  $X \times Z$  is compact, without loss of generality, assume that  $(a^n, w^n)$  converges to some point  $(a, w)$ . By continuity,  $z \sim_h (a, w)$  with  $(a, w) \in z$ . This completes the proof.  $\star$

By the consequence of Lemma 4,

$$U_{h|\sigma}(z) = \max_{i \in M(z)} U_{h|\sigma}(a_i z_i). \quad (\text{A.3.3})$$

For any  $z \in Z$  define

$$f_h(z) = \{w \in Z \mid z \succsim_{h|\sigma} w, \forall \sigma \in \hat{\Sigma}(h)\}.$$

The set  $f_h(z)$  returns the intersections of the lower contour sets of each state induced preference. We now claim that  $w \in f_h(z)$  if and only if  $z \sim_h z \cup w$ . This is the characterization of the dominance relation in Kreps [1979]:

$$\begin{aligned} w \in f_h(z) & \iff \\ \bigcup_{\sigma \in \hat{\Sigma}(h)} \sigma z \sim_h \bigcup_{\sigma \in \hat{\Sigma}(h)} \sigma \{z \cup w\} & \iff \\ z \sim_h \{z \cup w\}, & \end{aligned}$$

where the first implication follows from the definition of  $\succsim_{h|\sigma}$  and CSQ, and the second implication from STA.

From here we can follow Kreps' proof. Now define  $\phi_h : \mathbb{R}^{\hat{\Sigma}(h)} \rightarrow \mathbb{R}$  as any strictly increasing function that extends the map  $(U_{h|\sigma}(z))_{\sigma \in \hat{\Sigma}(h)} \mapsto U_h(z)$ . This is well defined since  $(U_{h|\sigma}(z))_{\sigma \in \hat{\Sigma}(h)} = (U_{h|\sigma}(w))_{\sigma \in \hat{\Sigma}(h)}$  implies that  $w \in f_h(z)$  and  $z \in f_h(z)$  which in turn implies that  $z \sim_h z \cup w \sim_h w$ . Likewise, if  $(U_{h|\sigma}(z))_{\sigma \in \hat{\Sigma}(h)} \geq (U_{h|\sigma}(w))_{\sigma \in \hat{\Sigma}(h)}$  with some strict equality then  $w \in f_h(z)$  but  $z \notin f_h(w)$  implying  $z \sim_h z \cup w \succ_h w$ . Plugging (A.3.3) into the aggregator we have

$$U_h(z) = \phi_h \left( \max_{i \in M(z)} U_{h|\sigma}(a_i, z_i)_{\sigma \in \hat{\Sigma}(h)} \right). \quad (\text{A.3.4})$$

**Step-4: States within States.** Now Theorem 2 tells us that  $\succsim_{h|\sigma}$  obeys the same set of axioms as  $\succsim$ . Therefore, if the above claim holds, and  $U_{h|\sigma}$  represents some  $\succsim_{h|\sigma}$  then we can repeat the previous steps (1-3) to obtain the state dependent version of equation (A.3.2), which implies

$$U_{h|\sigma}(az) = F_\sigma[u_\sigma(a), U_{h \cup a|\sigma}(z)],$$

From RCV we know that it is WLOG that  $F_\sigma = F$  for all  $\sigma$ . Assume that this was not the case, so that for some  $\sigma$ ,  $\dot{U} : z \mapsto \max_{i \in M(z)} F(u_\sigma(a_i), \dot{U}(z_i))$  does not represent  $\succsim_{h|\sigma}$ . So there must be some  $z$  and  $w$  such that  $z \succsim_{h|\sigma} w$  but  $\dot{U}(w) > \dot{U}(z)$ . By STA we have that  $\sigma_1 z \succsim_{h|\sigma} \sigma_1 w$ . Also, by the strict increasingness of  $F$  we have that  $\dot{U}(\sigma_1 w) > \dot{U}(\sigma_1 z)$ . Therefore, by the continuity of  $\succsim_{h|\sigma}$  and by RCV we can find sequences of universal outcome,  $\pi$  and  $\tau$  such that  $\pi \succsim_{h|\sigma} z \succsim_{h|\sigma} w \succsim_{h|\sigma} \tau$  and  $\dot{U}(\sigma_1 \tau) > \dot{U}(\sigma_1 \pi)$ , and hence  $\dot{U}(\tau) > \dot{U}(\pi)$ .

But this contradicts the invariance of preferences regarding universal streams. Indeed, we choose some common normalization, over  $\bar{a}$  and  $\underline{a}$ , then it is clear that  $\dot{U} \equiv U$  on such a domain. But  $\pi \succsim_{h|\sigma} \tau \iff \pi \succ \tau$ , which is a contradiction.

This provides,

$$U_h(z) = \phi_h((U_{h|\sigma}(z))_{\sigma \in \Sigma(h)}) = \phi_h \left( \max_{i \in M(z)} F[u_\sigma(a_i), U_{h \cup a_i|\sigma}(z_i)]_{\sigma \in \Sigma(h)} \right).$$

Moreover, we can repeat this entire exercise starting with the conditional preferences, retaining any normalizations. Note we use the definition of a conditional strategic plan (i.e., over sequences in  $\Sigma(h'|\sigma)$ ). Finally,

$$U_{h'|\sigma}(z) = \phi_{h'|\sigma}((U_{h'|\tau}(z))_{\tau \in \Sigma(h'|\sigma)}) = \phi_{h'|\sigma}\left(\max_{i \in M(z)} F[u_\tau(a_i), U_{h' \cup a_i|\tau}(z_i)]_{\tau \in \Sigma(h'|\sigma)}\right). \quad (\text{A.3.5})$$

This is the representation we are after.

**Step-5: Properties of the State Space.** It is immediate that the tuple,  $\mathcal{S} = \{\hat{\Sigma}(X), \{\hat{\Sigma}(h)\}_{h \subseteq X}\}$  satisfies properties (S1) (by the identification of  $\hat{\Sigma}(X)$  and  $SO_r(X)$ ), (S2), (S3), and (S4), and so, is a LTC information structure. Let  $\Sigma = \bigcup_{h \subseteq X} \hat{\Sigma}(h)$ . Moreover, it can be directly verified from (A.3.5) that  $\mathcal{U} = \{\{u_\sigma\}_{\sigma \in \Sigma}, \{\phi_{h'|\sigma}\}_{h' \subseteq X, \sigma \in \Sigma}, F\}$  satisfies (U1) and (U2). (U3) is an immediate consequence of RCV (most easily seen with the normalization  $u_\sigma(\underline{a}) = 0$  and  $F(x, 0) = x$ ). Lastly, the following Lemma provides (U4). By the construction of  $\mathcal{S}$ , it is clear that all  $s \in P(h) \equiv \hat{\Sigma}(h)$  agree on the ranking over  $h$ . The aggregation property, (U3), therefore establishes property (U4).

So,  $\{\{\hat{\Sigma}(X), \{\hat{\Sigma}(h)\}_{h \subseteq X}\}, \{\{u_\sigma\}_{\sigma \in \Sigma}, \{\phi_{h'|\sigma}\}_{h' \subseteq X, \sigma \in \Sigma}, F\}\}$  constitutes an LTC representation of  $\succcurlyeq$  as desired.

## A.4 Proof of Theorem 4

Let  $(\mathcal{S} = \{S, \{\mathcal{P}(h)\}_{h \subseteq X}\}, \mathcal{U})$  and  $(\mathcal{S}' = \{S', \{\mathcal{P}'(h)\}_{h \subseteq X}\}, \mathcal{U}')$  represent  $\succcurlyeq$ . A direct application of Theorem 1 (to the restriction of  $\succcurlyeq_X$  to menus of sequences) shows  $S = S'$ .

It remains to show that  $\mathcal{P}(h) = \mathcal{P}'(h)$  for all  $h \subseteq X$ . Towards a contradiction, assume this was not the case: for some  $h \subseteq X$  and  $P \in \mathcal{P}(h)$ ,  $P \not\subseteq P'$  for all  $P' \in \mathcal{P}'(h)$ . As in the construction in Theorem 1, let  $\Gamma(X)$  denote the set of all pairs of distinct elements of  $X$ , enumerated as  $\gamma_1 \dots \gamma_{\binom{|X|}{2}}$ . For any  $s \in S$  define  $\sigma^s \in X^{\binom{|X|}{2}}$  by

$$\sigma_n^s = \operatorname{argmax}_{b \in \gamma_n} u_s(b) \quad (\text{A.4.1})$$

Define the mapping  $k : P \mapsto \{\sigma^s | s \in P\}$ . Now, consider the IHCPs

$$z = \bigcup_{P \in \mathcal{P}(h)} (X \setminus h)k(P)\underline{a},$$

$$w = \bigcup_{P' \in \mathcal{P}'(h)} (X \setminus h)k'(P')\underline{a}.$$

Using the representation, it is clear  $P \in \mathcal{P}(h) \implies U_{h|P}(z) > U_{h|P}(w)$ , and hence  $U_{\emptyset}(h\{z \cup w\}) > U_{\emptyset}(hw)$ . But for every  $P' \in \mathcal{P}'(h)$ , we have by construction that  $U'_{h|P'}(h\{z \cup w\}) = U'_{h|P'}(hw)$ , and therefore,  $U'_{\emptyset}(h\{z \cup w\}) = U'_{\emptyset}(hw)$ , a contradiction to their joint representation.

## B Strategic Plans over Weak Rankings

The analysis by which strategic planning separates the preference ordering over the history via the construction of  $\Sigma(h)$  is not restricted to strict orderings. Although the actual sequences that identify each state may become very long they are ensured to exist and be finite. Explicitly, we are assuming that each of the preference ordering satisfies ORD and STA. This assumption is substantiated by Theorem 2, which states that if a strategic plan exists over such a  $\Sigma(h)$  exists, the the resulting conditional preferences will inherit those axioms.

We will first examine the case with  $|h| = 2$  to gain intuition then show how this can be extended. For this section let  $m$  denote the universally worst outcome. If  $h = \{a, b\}$ , then there are 3 states:  $a \succ b$ ,  $a \sim b$ , and  $b \succ a$ .

The complication is creating a sequence that is optimal when the DM strictly prefers  $a$  to  $b$  but *not* optimal when the preference is weak. This of course cannot be simply a choice of  $a$  since indifference does not rule out this choice as optimal. However, by using a longer sequence and the universal worst outcome we can find the desired sequences. Consider,

1:	$a \succ b$	$aa, aa, \dots, aw$
2:	$a \sim b$	$ab, ab, \dots, ab$
3:	$b \succ a$	$bb, bb, \dots, bw$

Where “...” denote the repetition of the previous consumption entries. If the decision maker is indifferent, it is clear that 2 is the unique optimal choice as it is the only sequence to avoid consumption of  $m$ . Moreover, for sufficiently long repetitions, 1 becomes the unique choice for a DM who prefers  $a$  to  $b$  strictly, since a one time consumption of  $m$  is better than the repeated consumption of  $b$ . To see why, note that **ORD** (in particular, continuity with respect to the product topology) implies tail robustness. We say a preference is *tail robust* if for all  $\rho, \tau \in \times^\infty X$ ,  $\rho > \tau$  implies that there exists some  $n \in \mathbb{N}$  such that  $\rho_n \mathbf{m} > \tau$ . This is a straightforward application of continuity and the fact that  $\rho_n \mathbf{m} \rightarrow \rho$  as  $n \rightarrow \infty$ . Moreover, **STA** implies that if  $aaaa\dots > bbbb\dots$  then  $aaaa\dots > abab\dots > bbbb\dots$ . These two facts ensure that for some finite repetition, the above sequences will produce a strategic plan.

This intuition is extendable to the general case. To outline how, we show the case with  $h = \{a, b, c\}$ . There are 13 weak orderings:

1:	$a \succ b \succ c$	$aab, aab, \dots, wab$	$aab, aab, \dots, awb$	$aab, aab, \dots, aaw$
2:	$a \succ c \succ b$	$aac, aac, \dots, wac$	$aac, aac, \dots, awc$	$aac, aac, \dots, aaw$
3:	$b \succ a \succ c$	$bab, bab, \dots, wab$	$bab, bab, \dots, bwb$	$bab, bab, \dots, baw$
4:	$b \succ c \succ a$	$bc b, bc b, \dots, wcb$	$bc b, bc b, \dots, bw b$	$bc b, bc b, \dots, bc w$
5:	$c \succ a \succ b$	$acc, acc, \dots, wcc$	$acc, acc, \dots, awc$	$acc, acc, \dots, acw$
6:	$c \succ b \succ a$	$bcc, bcc, \dots, wcc$	$bcc, bcc, \dots, bw c$	$bcc, bcc, \dots, bc w$
7:	$a \succ b \sim c$	$aab, aab, \dots, wab$	$aab, aab, \dots, awb$	$aab, aac, \dots, aab$
8:	$b \succ a \sim c$	$bab, bab, \dots, wab$	$bab, bcb, \dots, bab$	$bab, bab, \dots, baw$
9:	$c \succ a \sim b$	$acc, bcc, \dots, acc$	$acc, acc, \dots, awc$	$acc, acc, \dots, acw$
10:	$a \sim b \succ c$	$aab, bab, \dots, aab$	$aab, aab, \dots, awb$	$aab, aab, \dots, aaw$
11:	$a \sim c \succ b$	$aac, aac, \dots, wac$	$aac, acc, \dots, aac$	$aac, aac, \dots, aaw$
12:	$b \sim c \succ a$	$bc b, bc b, \dots, wcb$	$bc b, bc b, \dots, bw b$	$bc b, bcc, \dots, bc b$
13:	$a \sim b \sim c$	$aab, bab, \dots, aab$	$aac, acc, \dots, aac$	$aac, aab, \dots, aac$

One only needs to break up each pairwise distinct aspect of different preference relations into “blocks” that imitate the  $|h| = 2$  case. I.e., the first block separates strict and weak indifference between  $a$  and  $b$  the second between  $a$  and  $c$  and finally between  $a$  and  $c$ . Since DMs preferences (over IHCPs) are stationary (since we consume only elements in the exploration history and universal outcomes) this process separates all the weak preference relations from one another.



## C The Dominance Relation Behind Strategic Planning

In this section we provide equivalent statements for **PLN**, which postulates the existence of a strategic plan given any history. Similar planning type axioms, which posit the existence of an optimal way to reduce flexibility, have played similar roles (constructing subjective dynamic information structures) in the previous work such as Dillenberger et al. [2014] and the supplementary appendix to Krishna and Sadowski [2014]. This section, in addition to providing alternative axioms in our framework, sheds light on the connection between planning axioms and dominance axioms such as those used in Kreps [1979] and implied by the independence axiom in Dekel et al. [2001].

In this section we provide equivalent statements for **PLN**, which postulates the existence of a strategic plan given any history. This section, in addition to providing alternative axioms in our framework, sheds light on the connection between planning axioms and dominance axioms such as those used in Kreps [1979] and (implied by the independence axiom) in Dekel et al. [2001].

### C.1 The Dominance Relation

**Definition.** Set  $z, w \in Z$ . Then we say that  $z$   $\sigma$ -dominates  $w$ , denoted  $z \succcurlyeq_\sigma w$ , if and only if,

$$\sigma z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho(z \cup w) \succcurlyeq_h \sigma w \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho(z \cup w).$$

We say that  $z$   $\sigma$ -dominates  $w$  when the assignment is preferred in the absence of any cross state hedging. Hedging concerns are removed by ensuring that  $z$  and  $w$  can be chosen in any other “state.”

**Definition.** Set  $p \in \text{As}(W, h)$ . Then we say that  $p$  is in the **envelope of  $W$**  if  $p(\sigma)$   $\sigma$ -dominates  $w$  for all  $w \in W$  and  $\sigma \in \Sigma(h)$ .

An assignment is in the envelope of  $W$  if each assignment is  $\sigma$ -dominance maximal element of the collection; it is the maximal assignment according to dominance. Without further restrictions, the envelope of a collection need not exist. The following two restrictions will imply that (i) the envelope of any *finite*  $W$  exists, and (ii) an assignment in the envelope is a strategic plan for  $W$ . Moreover, since  $\sigma$ -dominance does not depend on the assignment, these restrictions will also guarantee that **CSQ** is satisfied.

For these claims to hold, we need to first ensure that  $\sigma$ -dominance is a weak order. Since the completeness of  $\succcurlyeq_\sigma$  is inherited by the completeness of  $\succcurlyeq$ , we need only to impose transitivity.

[A8:  $\sigma$ -TRANSITIVITY ( $\sigma$ -TRV)]. For all  $\sigma \in \Sigma(h)$ ,  $\succcurlyeq_\sigma$  is transitive.

While  $\sigma$ -TRV is written in terms of the auxiliary relation  $\succcurlyeq_\sigma$ , it is straight forward to write as a restriction on our primitive  $\succcurlyeq$ . Axiom  $\sigma$ -TRV, has the following immediate consequence:

**Lemma 5.** If  $\succcurlyeq$  is a weak order that satisfies  $\sigma$ -TRV, then  $\succcurlyeq_\sigma$  is a weak order over  $Z$ .

The next axiom is a modularity condition, ensuring the maximal element according to the dominance relation (i.e., maximal for all  $\sigma$ ) is a strategic plan.

[A9: MODULARITY (MOD)]. For all finite  $W$  and  $p \in \text{As}(W, h)$ ,  $p$  is in the envelope of  $W$  if and only if

$$\bigcup_{\rho \in \Sigma(h)} \rho(p(\rho)) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho W.$$

Modularity characterizes preferences for flexibility entirely in terms of dominance: flexibility is of no benefit unless it can improve the dominance of its assignments. This is essentially the content behind **PLN** (i.e., that flexibility can be muted) and **CSQ** (that plans are evaluated without hedging concerns). The next result states that in Theorem 3, our main representation result, **PLN** and **CSQ** could be replaced with  $\sigma$ -TRV and **MOD**.

**Theorem 5.** *If the preference  $\succsim$  satisfies **HST**, **ORD**, **STA**, **FLX**, and **CON**, then the following are equivalent:*

1.  $\succsim$  satisfies **PLN** and **CSQ**;
2.  $\succsim$  satisfies  $\sigma$ -TRV and **MOD**.

## C.2 Proof of Theorem 5

*Proof.* ( $2 \Rightarrow$  **STRATEGIC PLANNING**) Fix  $h$  and finite  $W$ . By Lemma 5 the  $\sigma$ -dominance relation induces a weak ordering,  $\succsim_\sigma$ , over  $Z$  for each  $\sigma \in \Sigma(h)$ . So, by the finiteness of  $W$  there is a maximal element (in  $W$ ) for each  $\sigma$ : call this  $w_\sigma$ . Define the assignment:  $\bar{p} : \sigma \mapsto w_\sigma$ .

We claim that  $\bar{p}$  is a strategic plan. By construction  $\bar{p}$  is in the envelope of  $W$ . Therefore, by applying Axiom **MOD** we have

$$\bigcup_{\rho \in \Sigma(h)} \rho p(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho W.$$

( $2 \Rightarrow$  **CONSEQUENTIALISM**) Let  $W, V \subseteq Z$ , and strategic plans  $p^W$  and  $p^V$ , such that  $p^W(\sigma), p^V(\sigma) \in V \cap W$  for some  $\sigma \in \Sigma(h)$ . Define  $w = p^W(\sigma)$  and  $v = p^V(\sigma)$ . Since

$$\bigcup_{\rho \in \Sigma(h)} \rho p^W(\rho) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho W,$$

Axiom **MOD** implies that  $p^W$  is in the envelope of  $W$ . Likewise,  $p^V$  is in the envelope of  $V$ . So, by definition  $v \succsim_\sigma z$  for all  $z \in V$ , hence  $v \succsim_\sigma w$ . Lastly, this implies that  $\hat{p}^W$ , defined by

$$\hat{p}^W(\pi) = \begin{cases} v & \text{if } \pi = \sigma \\ p^W(\pi) & \text{if } \pi \neq \sigma, \end{cases}$$

is in the envelope of  $W$ . Applying Axiom **MOD** again:

$$\sigma p^V(\rho) \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho(p^W(\sigma)) \sim_h \bigcup_{\rho \in \Sigma(h)} \rho W,$$

so **CSQ** holds.

( $1 \Rightarrow$   $\sigma$ -**TRANSITIVITY**). Claim:  $z \succsim_\sigma w$  if and only if  $z \succsim_{h|\sigma} w$ . Indeed, letting  $p \in As(\{z, w\}, h)$  be a strategic plan we have:

$$z \succsim_{h|\sigma} w \iff \sigma z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho p(\rho) \succsim_h \sigma w \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho p(\rho) \quad (\text{C.2.1})$$

$$\iff \sigma z \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho(z \cup w) \succsim_h \sigma w \cup \bigcup_{\rho \in \Sigma(h) \setminus \sigma} \rho(z \cup w) \quad (\text{C.2.2})$$

$$\iff z \succsim_\sigma w, \quad (\text{C.2.3})$$

where the equivalence in (C.2.1) invokes **CSQ**, and between (C.2.1) and (C.2.2) invokes Lemma 2. Given the claim, the transitivity of  $\succsim_{h|\sigma}$  (by Theorem 2) directly implies the transitivity of  $\succsim_\sigma$ .

( $1 \Rightarrow$  **MODULARITY**). Fix  $p \in As(W, h)$  with  $p$  in the envelope of  $W$ . By Strategic planning there exists strategic plan  $p^W$ . By the definition of the envelope of  $W$ ,  $p(\sigma) \succsim_\sigma w$  for all  $w \in W$ , which by the claim

implies  $p(\sigma) \succsim_{h|\sigma} w$  for all  $w \in W$ , in particular,  $p(\sigma) \succsim_{h|\sigma} p^W(\sigma)$ . By CSQ we have that  $p$  is a SP: the indifference condition of Axiom MOD holds. The opposite direction is near identical and thus omitted. ■

## D An Additive Representation

We conclude with a brief discussion regarding the possibility of introducing additive aggregators to the LTC representation. We believe it is possible to adapt the use of strategic planning and obtain an additive representation result for a set up including lotteries (as in Dekel et al. [2001]).<sup>16</sup> A more particular model with objective risk may be more tractable for some applications. However, while some axioms may be rewritten or weakened, it seems that the main axioms will be formulated in a similar way as in the current model. In that respect, it does not add much to the message of the paper and the identification techniques. Moreover, we argue there is an interpretational issue underlying the framework needed for such an exercise.

As in Dekel et al. [2001], Krishna and Sadowski [2014], in order to obtain an additive representation, we must consider a richer structure of choice objects and allow for lotteries. In this case, learning ordinally through comparisons between elements in the history, is not “compatible” with learning a cardinal utility function. That is, if learning takes place only through the *consumption of outcomes*, it is not possible that two different states would correspond to two different utility functions over lotteries that rank consumption elements in the same way. Note that this does not mean one can not identify the von Neumann–Morgenstern state dependent utility, but rather that no two states could share the same ordinal ranking. As such, the DM cannot be certain of her ordinal ranking over outcomes, but be unsure of her risk preferences. We find this somewhat unnatural.

A possible way to avoid this interpretational quagmire is to allow for consumption of, and thus comparisons between, lotteries. The DM considers consumption of lotteries ex-ante and she believes she will be able to compare such lotteries after experiencing them. In this case, learning different cardinal utilities (that agree on the rankings of extreme points) is possible. However, the information structure will always pertain to a finite set of elements (in a world of a continuum). More importantly, comparison of lotteries seems as unnatural as the problem we are trying to rectify. Since histories stand for experience, it seems more natural comparing between outcomes of lotteries, than comparing between lotteries themselves.

The origin of these interpretational difficulties is our reliance on strategic planning to identify conditional preferences. Strategic planning states that we can reduce everything the decision maker has learned to a *finite* series of pairwise choices. But, to identify a cardinal preference, an infinite number of pairwise choices is required. Notice that this issue does not arise in Cooke [2015]. Strategic planning is not necessary in that environment because, with two periods (where flexibility exists only in the second period), the DM’s choice *out* of a menu is never explicitly modeled –therefore, neither is the resolution of the learning process. It is only when moving from identifying the set of preferences the DM might learn (i.e.,  $\mathcal{P}(h)$ ) to identifying the preferences *conditional* on having learned a particular piece of information (i.e.,  $\mathcal{P}(h|P)$ ) that strategic planning is necessary.

---

<sup>16</sup>The proof would go roughly in the following line. We can consider a richer environment than that adopted in this paper, similar to that in Gul and Pesendorfer [2004]. With similar techniques we can obtain a result like in Theorem 3. For every exploration history  $h$ , every IHCP  $z$  is associated with  $U_h(z)$  and the vector  $(U_{h|\sigma}(z))_{\sigma \in \Sigma(h)}$ . So an additive representation exists if there is a vector  $(p_\sigma)_{\sigma \in \Sigma(h)}$  such that  $U_h(z) = \sum_{\sigma \in \Sigma(h)} p(\sigma) U_{h|\sigma}(z)$ . An infinite dimensional version of Farkas Lemma could be applied and, given an independence axiom, similar arguments as in Seo [2009] guarantee that such a representation exists. One still has to show that the weights aggregate properly across histories, but this holds due to the aggregation property (U3) of the LTC functional form, which is implied by Rcv.

## References

- Dirk Bergemann and Juuso Välimäki. Bandit problems. Technical report, Cowles Foundation for Research in Economics, Yale University, 2006.
- Kevin Cooke. Preference discovery and experimentation. 2015.
- Eddie Dekel, Barton L. Lipman, and Aldo Rustichini. Representing preferences with a unique subjective state space. *Econometrica*, 69(4), 2001.
- David Dillenberger, Juan Sebastian Lleras, Philipp Sadowski, and Norio Takeoka. A theory of subjective learning. *Journal of Economic Theory*, 153, 2014.
- David Dillenberger, R Vijay Krishna, and Philipp Sadowski. Dynamic rational inattention. 2015.
- Haluk Ergin and Todd Sarver. A unique costly contemplation representation. *Econometrica*, 78(4):1285–1339, 2010.
- Itzhak Gilboa and David Schmeidler. Case-based decision theory. *The Quarterly Journal of Economics*, pages 605–639, 1995.
- Itzhak Gilboa and David Schmeidler. *A theory of case-based decisions*. Cambridge University Press, 2001.
- John C Gittins and David M Jones. A dynamic allocation index for the discounted multiarmed bandit problem. *Biometrika*, 66(3):561–565, 1979.
- Faruk Gul and Wolfgang Pesendorfer. Self-control and the theory of consumption. *Econometrica*, 72, 2004.
- Faruk Gul and Wolfgang Pesendorfer. The revealed preference theory of changing tastes. *The Review of Economic Studies*, 72(2):429–448, 2005.
- Youichiro Higashi, Kazuya Hyogo, and Norio Takeoka. Stochastic endogenous time preference. *Journal of Mathematical Economics*, 51:77–92, 2014.
- Farhad Husseinov. Monotonic extension. Technical report, Bilkent University, 2010.
- Kazuya Hyogo. A subjective model of experimentation. *Journal of Economic Theory*, 133(1):316–330, 2007.
- Asen Kochov. Time and no lotteries: An axiomatization of maxmin expected utility. *Econometrica*, 83(1):239–262, 2015.
- Tjalling C Koopmans. Stationary ordinal utility and impatience. *Econometrica: Journal of the Econometric Society*, pages 287–309, 1960.
- David M. Kreps. A representation theorem for 'preference for flexibility'. *Econometrica*, 47:565–576, 1979.
- R Vijay Krishna and Philipp Sadowski. Dynamic preference for flexibility. *Econometrica*, 82(2):655–703, 2014.
- R Vijay Krishna and Philipp Sadowski. Unobservable persistent taste shocks. 2015.
- Evan Piermont and Roe Teper. Plans of action. 2016.
- Herbert Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58, 1952.
- Kyoungwon Seo. Ambiguity and second-order belief. *Econometrica*, 77:1575–1605, 2009.
- Christopher A Sims. Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690, 2003.

Marciano Siniscalchi. Dynamic choice under ambiguity. *Theoretical Economics*, 6(3):379–421, 2011.

Norio Takeoka. Subjective probability over a subjective decision tree. *Journal of Economic Theory*, 136(1): 536–571, 2007.

Martin L Weitzman. Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, pages 641–654, 1979.