

# Randomly Evolving Tastes and Delayed Commitment\*

R. VIJAY KRISHNA†

PHILIPP SADOWSKI‡

JUNE 9, 2016

## Abstract

We consider a decision maker with randomly evolving tastes who faces dynamic decision situations that involve intertemporal tradeoffs, such as those in consumption savings problems. We axiomatize a recursive representation of choice that features uncertain consumption utilities, which evolve according to a subjective Markov process. The parameters of the representation, which are the subjective Markov process governing the evolution of utilities, and the discount factor, are uniquely identified from behavior. We relate the correlation of tastes over time and the desire to delay commitment to future consumption.

## 1. Introduction

While taste is often modeled as a stable trait of the individual decision maker, tastes do change over time in many instances. For example, risk aversion tends to change over time (see Bekaert, Engstrom, and Grenadier (2010) and the references therein) and people's ranking over consumption outcomes may vary as well.

---

(\*) We would like to thank Haluk Ergin, Wolfgang Pesendorfer, Todd Sarver, and Norio Takeoka for helpful comments and suggestions, and Vivek Bhattacharya, Matt Horne, and Justin Valasek for valuable research assistance.

(†) Florida State University <rvk3570@gmail.com>

(‡) Duke University <p.sadowski@duke.edu>

Consider the dynamic behavior of a forward looking decision maker, who is aware that his tastes may change, and who embraces the future changes in tastes, in the sense that he evaluates future consumption based on his expectation over future consumption tastes. On the one hand, if tastes evolve randomly, then he prefers not to commit to consumption choice in advance. On the other hand, if tastes are correlated between subsequent periods, then the reluctance to commit will be reduced as the time of consumption draws nearer. That is, the decision maker prefers to delay commitment.<sup>1</sup>

Specifically, we consider a decision maker who perceives a particular type of Markov process that governs the evolution of his tastes, that is, his current taste is a sufficient statistic for his current believe over future tastes. We then analyze preferences over Infinite Horizon Consumption Problems as introduced in Gul and Pesendorfer (2004) (henceforth GP) where, in every period, choice is between lotteries over current consumption and a continuation choice problem for the next period.

As argued above, our decision maker will want to delay commitment in general, but due to the Markovian structure of the process that governs the evolution of tastes, he is willing to commit to a continuation problem for the next period, *contingent* on the current taste. While this taste is not observable by the analyst, we argue that strategic rationality should also hold contingent on current consumption *choice* from a large enough menu. We formalize this notion in an axiom called *Choice Contingent Continuation Strategic Rationality (Choice Contingent CSR)*. The Markov process that governs the evolution of tastes over time in our representation is uniquely identified from first period preferences, as is the only other preference parameter, the discount factor.

If tastes are correlated over time, knowledge of the taste at one point will reduce an individual's uncertainty about future consumption tastes. In that case, the individual's willingness to commit to a consumption choice should increase between one period and the next, more so the more correlated tastes are over time. For example, an investor will be more willing to commit to a more or less risky portfolio (say by accepting a penalty for reallocating his funds) given his

---

(1) Following Kreps (1979) and Dekel, Lipman, and Rustichini (2001), models of preference for flexibility have been used to capture a decision maker with the desire to accommodate future changes in taste in an environment with only one instance of consumption choice, leaving no room for changes in the willingness to commit over time.

current risk aversion, if his risk aversion is strongly correlated over time. This is independent of ex-ante uncertainty about the level of risk aversion. Theorem 2 provides comparative statics that formalize this intuition.

Krishna and Sadowski (2014) (henceforth KS) model a decision maker in a dynamic environment who chooses over acts on an objective space of states of the world, and who has stable but noisy state-contingent tastes. A common special case of their model and ours features a stable and state independent underlying taste that is perturbed by iid noise (or transient taste shocks). Our main axiom relaxes their notion of unconditional *Continuation Strategic Rationality*, and some of our more standard axioms are common to both models.

The remainder of this paper is organized as follows. Section 2 specifies our environment, the axioms and the representation result. Section 3 has the comparative statics and Section 4 concludes by discussing related decision theoretic literature and directions for future research.

## 2. A Model of Randomly Evolving Tastes

This section provides our representation result. Section 2.1 describes the environment, Section 2.2 has the behavioral axioms, Section 2.3 contains the representation result and Section 2.4 provides intuition for the proof of the representation result.

### 2.1. Environment

For a compact metric space  $Y$ , let  $\mathcal{P}(Y)$  denote the space of probability measures endowed with the topology of weak convergence, so that  $\mathcal{P}(Y)$  is compact and metrizable. Let  $\mathcal{F}(Y)$  denote the space of closed subsets of a compact metric space  $Y$ , endowed with the Hausdorff metric, which makes  $\mathcal{F}(Y)$  a compact metric space.

Let  $K$  be a finite set of consumption prizes with typical member  $k$ . We follow GP in defining an infinite horizon consumption problem (IHCP) as a collection of lotteries that yield a prize in the present period and a new infinite horizon problem starting in the next period. Let  $Z$  be the collection of all IHCPs.<sup>2</sup> GP show that  $Z$  is a compact metric space, and that each  $z \in Z$  can be identified with a compact set

---

(2) See GP for the recursive construction of  $Z$ .

of probability measures over  $K \times Z$ . In particular, it can be shown that  $Z$  is linearly homeomorphic to the space of all closed subsets of  $\mathcal{P}(K \times Z)$ . We shall denote this linear homeomorphism by  $Z \simeq \mathcal{F}(\mathcal{P}(K \times Z))$ . Typical elements  $x, y, z \in Z$  are interpreted as menus of lotteries over consumption and continuation problems, while  $p, q \in \mathcal{P}(K \times Z)$  are typical lotteries, with  $p_k$  and  $p_z$  denoting the marginal distributions of  $p$  on  $K$  and  $Z$ .

We explicitly model choice between consumption problems from an ex-ante perspective, before consumption begins. That is, we analyze a binary relation  $\succsim \subset Z \times Z$ , which we refer to as a *preference*. We let  $\succ$  and  $\sim$  denote, respectively, the asymmetric and symmetric parts of  $\succsim$ . The recursive domain of IHCPs is rich; for instance, it can accommodate temporal lotteries as in Kreps and Porteus (1978). It is also amenable to analysis by stochastic dynamic programming.

We will also consider the space of menus of consumption lotteries,  $\mathcal{F}(\mathcal{P}(K))$ , with typical members being  $a, b, c$ . By the recursive nature of  $Z$ , continuation problems are members of  $Z$ . Let  $A, B, C$  denote typical elements of the collection of menus of continuation lotteries,  $\mathcal{F}(\mathcal{P}(Z))$ . To ease notational burden, we will often write  $\mathcal{F}$  for  $\mathcal{F}(\mathcal{P}(K \times Z))$ ,  $\mathcal{F}_K$  for  $\mathcal{F}(\mathcal{P}(K))$ , and  $\mathcal{F}_Z$  for  $\mathcal{F}(\mathcal{P}(Z))$ . When there is no risk of confusion, we identify prizes and continuation problems with degenerate lotteries and lotteries with singleton menus. For example, we denote the lottery over continuation problems that yields  $z$  with certainty by  $z$ , and the lottery that yields current consumption  $k$  and continuation problem  $x$  with certainty by  $(k, x)$ .

## 2.2. Axioms

Axioms 1–5 are fairly standard and are discussed in more detail in KS.

**AXIOM 1** (Non-triviality).  $\succsim$  is non-trivial, in the sense that there exist  $x, y \in Z$  such that  $x \succ y$ .

**AXIOM 2** (Continuous Order).  $\succsim$  satisfies the following:

- (a)  $\succsim$  is complete and transitive.
- (b)  $\succsim$  is continuous, in the sense that  $\{y : y \succsim x\}$  and  $\{y : x \succsim y\}$  are closed.

We take the convex sum of sets to be the Minkowski sum, namely  $\lambda x + (1 - \lambda)y := \{\lambda p + (1 - \lambda)q : p \in x, q \in y\}$  whenever  $\lambda \in [0, 1]$ . Notice that if  $x, y \in \mathcal{F}$ ,

then  $\lambda x + (1 - \lambda)y$  is also closed, and hence is in  $\mathcal{F}$ . The following axiom is von Neumann-Morgenstern's Independence axiom.

**AXIOM 3** (Independence).  $x \succ y$  implies  $\lambda x + (1 - \lambda)z \succ \lambda y + (1 - \lambda)z$  for all  $\lambda \in (0, 1]$  and  $z \in Z$ .<sup>3</sup>

A standard decision maker is one whose preferences are *strategically rational* in the sense that  $x \succsim y$  implies  $x \cup y \sim x$ . A standard decision maker who satisfies Axiom 2 chooses as if he evaluates each set by its best element. There exists, then, a continuous function  $w : \mathcal{P}(K \times Z) \rightarrow \mathbb{R}$  that is linear, such that the functional  $x \mapsto \max_{p \in x} w(p)$  represents  $\succsim$ .<sup>4</sup>

We are interested in a decision maker (henceforth DM) who values flexibility.

**AXIOM 4** (Monotonicity).  $x \cup y \succsim x$  for all  $x, y \in Z$ .

This is the central axiom in Kreps (1979). It says that additional alternatives are always weakly beneficial.

Theorem 4 in KS shows that Axioms 1–4 are necessary and sufficient to afford  $\succsim$  a *finitely additive EU representation*. In particular, there exists a subjective state space,  $\mathfrak{U}_{K \times Z}$ , which is a collection of all the (twice-normalized) vN-M utility functions on  $K \times Z$ , along with the Borel algebra  $\mathcal{A}_{\mathfrak{U}_{K \times Z}}$ , and a charge  $\mu$  on  $\mathfrak{U}_{K \times Z}$  that induces the preference functional

$$V(x) := \int_{\mathfrak{U}_{K \times Z}} \max_{p \in x} u(p) \, d\mu(u)$$

The state space  $\mathfrak{U}_{K \times Z}$  consists of all continuous functions on  $K \times Z$  that are identified up to positive affine transformation. In particular, all the utility functions in  $\mathfrak{U}_{K \times Z}$  have (i) the same utility for some  $x^* \in Z$ , and (ii) the same (supremum) norm. The first requirement corresponds to normalizing the constant term to 0, and the second requirement amounts to normalizing the scaling factor to 1.<sup>5</sup>

---

(3) A lottery  $p \in \mathcal{P}(K \times Z)$  is a *singleton* menu. A weaker version of Independence is *Singleton Independence*, which says that Independence holds for all singleton menus. GP show that Singleton Independence along with Stationarity (Axiom 6) and Indifference to Timing (Axiom 7) imply Independence (Axiom 3) assumed here.

(4) See footnote 5 of GP for a formal argument.

(5) There exists a unique element  $x^* \in Z$  such that  $x^* \simeq (p_k^*, x^*) \in \mathcal{F}$ , where  $x^*$  consists of the uniform lottery over  $K$ , namely  $p_k^* \in \mathcal{P}(K)$ , in each period. Analogous to the definition in DLR, we formally have  $\mathfrak{U}_{K \times Z} := \{u \in C(K \times Z) : \|u\|_\infty = 1, \sum_{k \in K} u(k, x^*) = 0\}$ , where  $C(K \times Z)$  is the Banach space of continuous functions on  $K \times Z$  with the supremum norm.

Each probability measure  $p$  over  $K \times Z$  induces marginal distributions  $p_k$  and  $p_z$  over  $K$  and  $Z$  respectively. The next axiom says that DM does not care about correlations between outcomes in  $K$  and  $Z$ , but only cares about the marginal distributions induced by the lotteries in the menu. In particular, if two lotteries induce the same marginal distributions over  $K$  and  $Z$ , then DM does not value the flexibility of having both lotteries available for choice.

**AXIOM 5** (Separability). If, for  $p, q \in \mathcal{P}(K \times Z)$ , the marginal distributions satisfy  $p_k = q_k$  and  $p_z = q_z$ , then  $\{p, q\} \sim \{p\}$ .

Versions of the next two axioms appear in GP, who provide a more detailed discussion. We are interested in stationary preferences, where the ranking of continuation problems does not depend on time. The recursive nature of the domain allows us to capture this notion via the following axiom, which says that if  $x \succsim y$ , then  $x$  is also better than  $y$  as a continuation problem after consumption of  $k$ . Recall that  $(k, x)$  denotes the degenerate lottery that gives  $(k, x) \in K \times Z$  with probability one.

**AXIOM 6** (Stationarity).  $\{(k, x)\} \succsim \{(k, y)\}$  if, and only if,  $x \succsim y$ .

In what follows, we will find particular use for menus with a product structure.

**Definition 2.1.** For  $p \in \mathcal{P}(K \times Z)$ , let  $p_k \in \mathcal{P}(K)$  and  $p_z \in \mathcal{P}(Z)$  be the corresponding marginals, and let  $(p_k, p_z)$  be the product lottery. For  $c \in \mathcal{F}_K$  and  $A \in \mathcal{F}_Z$ , we write  $(c, A) \in Z$  to denote the *rectangular menu*  $\{(p_k, p_z) : p_k \in c, p_z \in A\}$ .

If  $\mathbf{x} \subset Z$  is closed, then  $(k, \mathbf{x}) := \{(k, x) : x \in \mathbf{x}\}$ . For closed  $\mathbf{x}, \mathbf{y} \subset Z$ , let  $\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}$  be the menu of continuation problems  $\{\lambda x + (1 - \lambda) y : x \in \mathbf{x}, y \in \mathbf{y}\}$ . Notice that although  $\mathbf{x} \in \mathcal{F}_Z$ , it consists only of degenerate lotteries.

While our domain is sufficiently rich to permit different attitudes towards temporal lotteries and the timing of the resolution of uncertainty, we shall keep matters as close as possible to the standard model. That is, we shall abstract from particular patterns of preference for the timing of uncertainty, and consider the axiom.

**AXIOM 7** (Indifference to Timing).  $\lambda(k, \mathbf{x}) + (1 - \lambda)(k, \mathbf{y}) \sim (k, \lambda \mathbf{x} + (1 - \lambda) \mathbf{y})$  for all  $\lambda \in [0, 1]$ ,  $k \in K$  and  $\mathbf{x}, \mathbf{y} \subset Z$ .

The axiom states that DM is indifferent between (i) receiving lottery  $\lambda(k, \mathbf{x}) + (1-\lambda)(k, \mathbf{y})$ , which yields consumption  $k$  and determines whether the continuation problem will be chosen from  $\mathbf{x}$  or  $\mathbf{y}$ , (early resolution) and (ii) receiving with certainty consumption  $k$  and choosing a continuation menu from  $\lambda\mathbf{x} + (1-\lambda)\mathbf{y}$  (late resolution). The version of Indifference to Timing state above is stronger than that considered in GP or KS. This is because our present setting allows DM to have a preference for flexibility with respect to continuation problems, while GP and KS do not.

Separability (Axiom 5) allows us to consider an induced marginal preference relation  $\succsim_K \subset \mathcal{F}_K \times \mathcal{F}_K$ .

**Definition 2.2.** Fix  $A \in \mathcal{F}_Z$ . Let  $a \succsim_K b$  if, and only if,  $(a, A) \succ (b, A)$ .

Lemma B.1 in Appendix B.1 shows that for a separable preference that has a finitely additive EU representation,  $\succsim_K$  is independent of the choice of  $A \in \mathcal{F}_Z$ . In order to elicit continuation preferences contingent on any particular consumption ranking, we aim to identify two consumption menus  $a$  and  $b$  in  $\mathcal{F}_K$  with  $a \cup b \succ_K b$ , such that the best alternative is in  $a$  only under that consumption ranking. This is only possible if the collection of relevant consumption rankings is finite.

**AXIOM 8** (Finiteness). For all  $a \in \mathcal{F}_K$ , there is a finite set  $b \subset a$  with  $b \sim_K a$ .

Intuitively, if every set  $a$  has a finite subset  $b$  that is as good as  $a$  itself, then only a finite collection of consumption rankings can be relevant. The formal statement of this result is provided by Riella (2013, Theorem 2), who establishes that Axiom 8 is the appropriate version of the finiteness assumption in Dekel, Lipman, and Rustichini (2009) (their Axiom 11) when Monotonicity (Axiom 4) is assumed (but see also Lemma B.2 in Appendix B.1 for a direct proof).

The next axiom is our main behavioral assumption. The goal is to require, in terms of behavior, that DM be strategically rational with respect to continuation problems *contingent* on his consumption ranking,  $\succsim_K$ . Given a particular element of a finite collection of consumption rankings, it is straight forward to construct two consumption menus  $a$  and  $b$  in  $\mathcal{F}_K$ , such that the most preferred alternative from  $a \cup b$  is in  $a$  only under that consumption ranking. Hence, it would be sufficient to ask DM, for all  $a, b$  with  $a \cup b \succ_K b$ , whether he is strategically rational contingent on his preferred alternative from  $a \cup b$  being in  $a$ .

However, this requirement would be too strong, as for some  $a$  and  $b$ , the best alternative is in  $a$  for more than one consumption ranking.<sup>6</sup> In order to avoid such situations, we allow the addition of any collection of alternatives,  $c$ , to  $b$ , so long as  $a \cup b \cup c \succ_K b \cup c$ . If the best alternative from  $a \cup b$  is in  $a$  for only one ranking, this must remain true for any such  $c$ . If the best alternative is in  $a$  for multiple rankings, then there is  $c$  such that the best alternative from  $a \cup b \cup c$  is in  $a$  for only one ranking. Summing up, we would like to require the following:

If  $a \cup b \succ_K b$  then there is  $c \in \mathcal{F}_K$ , such that (i)  $a \cup b \cup c \succ_K b \cup c$ , and (ii) DM is strategically rational with respect to continuation problems, contingent on his preferred consumption choice from  $a \cup b \cup c$  being in  $a$ .

To formalize (ii), consider two continuation menus  $A$  and  $B$ . Let  $C$  be a subset of  $A \cup B$ . Then, the most preferred lottery from  $(a, C) \cup (b \cup c, A \cup B)$  is in  $(a, C)$  only if the most preferred consumption choice is in  $a$ . The following definition simplifies notation.

**Definition 2.3.** Given a menu  $z$ , let  $\overset{\circ}{\succ}_z$  denote the induced ranking over additional alternatives:  $x \overset{\circ}{\succ}_z y$  if, and only if,  $x \cup z \succ y \cup z$ .

**AXIOM 9** (Choice Contingent CSR). If  $a \cup b \succ_K b$  then there is  $c \in \mathcal{F}_K$ , such that<sup>7</sup>

- [i]  $a \cup b \cup c \succ_K b \cup c$
- [ii]  $(a, A) \overset{\circ}{\succ}_{(b \cup c, A \cup B)} (a, B)$  implies  $(a, A) \overset{\circ}{\succ}_{(b \cup c, A \cup B)} (a, A \cup B)$

In the setting where DM considers only finitely many consumption rankings relevant, the axiom is falsifiable in spite of the existential qualifier. To see this, note that  $\succ_K$  determines the finite set of relevant consumption rankings. As noted in the discussion preceding the axiom, the construction of  $a$  and  $b$  such that  $a$  outperforms  $b$  for only one given consumption ranking is straightforward. For such  $a$  and  $b$  we may assume, without loss of generality, that the  $c$  whose existence

---

(6) In particular, Continuation Strategic Rationality would be implied for  $a$  and  $b$  such that the best alternative is always in  $a$ .

(7) The induced relation  $\overset{\circ}{\succ}_{(b \cup c, A \cup B)}$  is only an expositional device. Condition (ii) can also be written in terms of  $\overset{\circ}{\succ}$ :  $(a, A) \cup (b \cup c, A \cup B) \overset{\circ}{\succ} (a, B) \cup (b \cup c, A \cup B)$  implies  $(a, A) \cup (b \cup c, A \cup B) \sim (a, A \cup B) \cup (b \cup c, A \cup B)$ .



is guaranteed by the axiom is empty. This observation is formally established in Lemma B.3.

Finally, we want to ensure that ex-ante preferences are a useful description of choice on the recursive domain of IHCPs, in the sense that no alternative that is relevant from the ex-ante perspective can become permanently irrelevant in the future. To avoid conditioning on entire consumption paths, we impose the stronger requirement that the collection of relevant alternatives is not choice contingent.

**AXIOM 10** (Persistent Preference for Flexibility). For all  $a, b \in \mathcal{F}_K$  such that  $a \cup b \succ_K b$ ,

$$x \cup y \succ x \text{ implies } (a, \{x \cup y\}) \succ_{(b, \{x \cup y\})}^{\circ} (a, \{x\}).$$

The first qualifier considers two consumption menus  $a$  and  $b$ , where  $a$  is not dominated by  $b$ . Again, the axiom only has implications for the case where the preferred consumption choice is in  $a$  rather than  $b$ . The axiom says that, if  $x \cup y \succ x$ , then this must also be true contingent on the preferred consumption choice being in  $a$ . We emphasize that this requirement and Choice Contingent CSR (Axiom 9) are *not* mutually exclusive: contingent on next period's preferred consumption choice being in  $a$ , the singleton  $\{x \cup y\}$  provides DM with additional alternatives for choice two periods from now, which he may value (Axiom 10), even though he is strategically rational with respect to the union  $\{x\} \cup \{y\}$  (Axiom 9), which would force him to choose one of the two smaller continuation problems in the next period.

### 2.3. Preference for Flexibility with Ranking Persistent Utilities

Choice Contingent CSR allows for correlation of consumption utilities. To see this, notice that the choice of a lottery in  $a$  over any lottery in  $b \cup c$  only carries information about DM's *current* ranking of immediate consumption and, contingent on this information, preferences are required to satisfy strategic rationality with respect to continuation problems. Hence, today's consumption ranking must be a sufficient statistic for today's beliefs over future consumption utilities. In what follows,  $\mathcal{U} := \{u \in \mathbb{R}^K : \sum_i u_i = 0\}$  denotes the space of all consumption utilities normalized up to an additive constant. Also, we shall say that a probability

measure  $\mu$  on  $\mathcal{U}$  is *nice* if  $\mu u := \int u \, d\mu(u) \in \mathcal{U}$ . Intuitively, a probability measure (or subjective belief) is nice if the expected utility from every prize  $k \in K$  is finite.

**Definition 2.4.** A *ranking persistent* Markov process  $(\mathcal{U}_M, M)$  consists of a state space  $\mathcal{U}_M \subset \mathcal{U}$  that is a finitely generated cone,<sup>8</sup> and a Markov kernel<sup>9</sup>  $M$  from  $\mathcal{U}_M$  to itself, such that  $M$  is:

- (a) *ranking contingent*:  $M(u, \cdot) = M(\lambda u, \cdot)$  for all  $\lambda > 0$ , and
- (b) *persistent*:  $M(u, \{\lambda u' : \lambda > 0\}) > 0$  for all  $u, u' \in \mathcal{U}_M$ .

The ranking persistent Markov process  $(\mathcal{U}_M, M)$  is *nice* if  $M(u, \cdot)$  is a nice probability measure for each  $u \in \mathcal{U}_M$ .

We identify the Markov process  $(\mathcal{U}_M, M)$  by its Markov kernel  $M$ , when the state space  $\mathcal{U}_M$  is understood. The following lemma establishes that the class of ranking persistent Markov kernels is a desirable subclass of all Markov kernels on  $\mathcal{U}$ , as ranking persistence guarantees the existence of a unique invariant (and hence ergodic) measure of the Markov process.<sup>10</sup>

**Lemma 2.5.** If  $(\mathcal{U}_M, M)$  is a ranking persistent Markov process, then an invariant measure  $\mu_0$  of  $M$  exists and is unique, where  $\mu_0(du) = \int_{\mathcal{U}} M(u', du) \mu_0(du')$ .

A proof is in appendix A. As before, we represent integrals with respect to measures as extensions by linearity and continuity, which allows us to write  $V(\cdot, \mu_0) := \int_{\mathcal{U}} V(\cdot, u) \, d\mu_0(u)$ . Similarly,  $V(p_z, u)$  denotes the linear extension (by continuity) of  $V(z, u)$  from  $Z$  to  $\mathcal{P}(Z)$ .

**Definition 2.6.** Let  $\delta \in (0, 1)$  and let  $(\mathcal{U}_M, M)$  be a nice, ranking persistent Markov process. A preference  $\succsim$  has an *evolving tastes* representation  $((\mathcal{U}_M, M), \delta)$ , if

- 
- (8) A set  $D \subset \mathcal{U}$  is a *cone* if for all  $\lambda > 0$ ,  $u \in D$  implies  $\lambda u \in D$ . The cone  $D$  is *generated* by a set  $D_0 \subset \mathcal{U}$  if  $D := \bigcup_{\lambda > 0} \lambda D_0$ . It is *finitely generated* if it is generated by a finite set.
  - (9) Let  $(X, \mathcal{X})$  be a measurable space. Then,  $M : X \times \mathcal{X} \rightarrow [0, 1]$  is a *Markov kernel* from  $(X, \mathcal{X})$  to itself if (i) for each  $x \in X$ ,  $M(x, \cdot)$  is a probability measure on  $(X, \mathcal{X})$ ; and (ii) for each  $D \in \mathcal{X}$ ,  $M(\cdot, D)$  is a measurable function defined on  $X$ . The Markov kernel represents the transition probabilities for a Markov process with state space  $X$ .
  - (10) This is intuitive, because the induced Markov process on the relevant rankings is fully connected. For the purpose of the identification results it would be sufficient to consider irreducible Markov processes on the relevant rankings. Note that this generalization is small, in the sense that the class of fully connected Markov processes is dense in the class of irreducible processes. The small gain in generality does not seem to warrant imposing a weaker, but harder to falsify assumption that only requires the existence of some finite consumption path, contingent on which persistence is satisfied.

$V(\cdot, \mu_0)$  represents  $\succsim$ , where  $\mu_0$  is the invariant measure of  $M$ , and  $V$  is defined recursively as

$$[2.1] \quad V(x, u) = \int_{\mathcal{U}} \max_{p \in x} [u'(p_k) + \delta V(p_z, u')] M(u, du')$$

Because  $\mu_0$  is the invariant measure of the Markov process,  $V(x, \mu_0)$  takes the intuitive form

$$V(x, \mu_0) = \int_{\mathcal{U}} \max_{p \in x} [u(p_k) + \delta V(p_z, u)] d\mu_0(u)$$

To see this, consider some menu  $x \in Z$ . Then,

$$\begin{aligned} \int_{\mathcal{U}} V(x, u) d\mu_0(u) &= \int_{\mathcal{U}} \left[ \int_{\mathcal{U}} \max_{p \in x} [u'(p_k) + \delta V(p_z, u')] M(u, du') \right] d\mu_0(u) \\ &= \int_{\mathcal{U}} \left[ \max_{p \in x} [u'(p_k) + \delta V(p_z, u')] \underbrace{\int_{\mathcal{U}} \mu_0(du) M(u, du')}_{=\mu_0(du')} \right] \\ &= \int_{\mathcal{U}} \max_{p \in x} [u'(p_k) + \delta V(p_z, u')] d\mu_0(u) \end{aligned}$$

where the first equality uses equation [2.1], the second equality uses Fubini's theorem to reverse the order of integration, and  $\mu_0(du') = \int_{\mathcal{U}} M(u, du') \mu_0(du)$  because  $\mu_0$  is the invariant distribution of the Markov process. The ranking persistence of  $M$  implies that for all  $u \in \mathcal{U}$ ,  $V(x, u) = V(x, \lambda u)$  for all  $\lambda > 0$ .

**Proposition 2.7.** Each evolving tastes representation  $((\mathcal{U}_M, M), \delta)$  induces a unique continuous function  $V \in C(Z \times \mathcal{U}_M)$  that satisfies equation [2.1] above.

The proof is in Appendix B.4. We shall say that two Markov processes  $(\mathcal{U}_M, M)$  and  $(\mathcal{U}_{M'}, M')$  are *identical up to scaling* if  $\mathcal{U}_M = \mathcal{U}_{M'}$  and there exists  $\lambda > 0$  such that  $M(u, D) = M'(u, \lambda D)$  for all measurable  $D \subset \mathcal{U}$ .

**Theorem 1.** *The binary relation  $\succsim$  satisfies Axioms 1–10 if, and only if, it has an evolving tastes representation  $((\mathcal{U}_M, M), \delta)$ . Moreover,  $(\mathcal{U}_M, M)$  is unique up to scaling, and  $\delta$  is unique.*

The proof of the theorem is in Appendix B. The next section provides a sketch of the most instructive steps in the proof of the theorem, and also demonstrates how the recursive structure of the evolving tastes representation implies uniqueness of  $M$  up to scaling, even though the continuation value varies with the consumption utility.

## 2.4. Proof Intuition for the Evolving Tastes Representation

Theorem 3 of KS establishes the existence of a finitely additive EU representation. KS (appendix C) show that by Separability, each utility  $u \in \mathfrak{U}_{K \times Z}$  can be written as  $u(p) = u(p_k; u) + v(p_z; u)$ , where  $u \in \mathcal{U}$ , and  $v \in \mathfrak{U}_Z := \{v \in C(Z) : \|v\|_\infty = 1, v(x^*) = 0\}$  (recall that  $x^* \simeq (p_k^*, x^*) \in Z$ ). KS (Appendix A.2) also show that if a preference with a finitely additive EU representation is *strategically rational* — ie,  $x \succsim y$  implies  $x \sim x \cup y$  — it can be viewed as having only one relevant utility  $u \in \mathfrak{U}_{K \times Z}$ . In particular, the charge representing his subjective beliefs is a probability measure concentrated at a single point.

Finiteness implies that there are only finitely many relevant consumption rankings. Let these relevant consumption rankings be  $u_1, \dots, u_n \in \mathcal{U}$ , and define  $[u] := \{\lambda u : \lambda > 0\}$ . Then, set  $\mathcal{U}_M := \bigcup_i [u_i]$ , which will serve as the state space for the Markov process that we will construct as part of the representation. We argued, when discussing Choice Contingent CSR (Axiom 9), that we can establish strategic rationality with respect to continuation problems for any one of those consumption rankings. Therefore the consumption ranking must be a sufficient statistic for beliefs. That is, for each ranking  $[u_i]$  there exists a unique  $v(\cdot, [u_i]) : \mathcal{F} \rightarrow \mathbb{R}$  that evaluates continuation problems, so that there are only finitely many valuations for continuation problems. This observation is key to showing that the representation is jointly identified because this allows us to show that the subjective state space can be taken to be finite dimensional, which in turn allows us to use DLR's joint identification result.

Persistent Preference for Flexibility (Axiom 10) implies that the functions  $v(\cdot, [u_i])$  are monotone with respect to set inclusion and, in particular, are locally non-satiated. To show that the functions  $v(\cdot, [u_i])$  are linear, let  $v_i(z) := v(z, [u_i])$ , and consider the set  $O := \bigcap_i v_i^{-1}(\text{int } v_i(Z))$ . It is easy to see that  $O$  is open, and because the functions  $v_i$  are locally non-satiated,  $O$  is also dense in  $Z$ . Moreover, each  $v_i$  is uniformly continuous (because  $Z$  is compact). Therefore, to show that  $v_i$  is linear on  $Z$ , it suffices to show that  $v_i$  is linear on  $O$ .

Suppose now, there exist  $x, y \in O$  and  $\lambda \in (0, 1)$  such that  $v_1(\lambda x + (1-\lambda)y) \neq \lambda v_1(x) + (1-\lambda)v_1(y)$ . There exist sets  $\mathbf{x} := \{x_1, \dots, x_n\}$  and  $\mathbf{y} := \{y_1, \dots, y_n\}$  where (i)  $x = x_1, y = y_1$ , (ii)  $v_i(x_i) > v_i(x_j)$  for all  $i \neq j$  and similarly for  $\mathbf{y}$ , and (iii)  $\mathbf{x}, \mathbf{y} \subset O$ . It is easy to see that for each  $\lambda \in [0, 1]$ ,  $\lambda(k, \mathbf{x}) + (1-\lambda)(k, \mathbf{y})$  is the unique lottery in the menu  $\lambda(k, \mathbf{x}) + (1-\lambda)(k, \mathbf{y})$  that maximizes  $u + \delta v_i$ . Indifference to

Timing (Axiom 7) allows us to conclude that,  $\lambda x_i + (1 - \lambda)y_i \in Z$  is the unique maximizer of  $v_i$  from the set  $\lambda \mathbf{x} + (1 - \lambda)\mathbf{y} \subset Z$  for each  $i$ , and for  $i = 1$  in particular, which is a contradiction. This establishes that  $v_i$  is linear on  $O$ .

Thus, each  $v_i$  has a finitely additive EU representation. By Stationarity (Axiom 6), we show that  $\succsim$  has a recursive representation

$$V(x, \mu_0) = \int_{\mathcal{U}} \max_{p \in x} [u(p_k) + \delta(u)V(p_z, u)] \mu_0(du)$$

where the discount factor may only depend on the consumption ranking, that is,  $\delta(u) = \delta(\lambda u)$  for all  $\lambda > 0$ , and  $(\mathcal{U}_M, M)$  is a ranking contingent Markov process such that

$$V(x, u) = \int_{\mathcal{U}} \max_{p \in x} [u'(p_k) + \delta(u')V(p_z, u')] M(u, du')$$

for each  $u \in \mathcal{U}_M$ .

We need to establish that the value function  $V$  can be renormalized to make the discount factor independent of  $u$ . Suppose, for simplicity, that the support of  $\mu_0$  is  $\mathcal{U}^* = \{u_1, \dots, u_n\}$  where no  $u_i$  and  $u_j$  are collinear. Suppose also that there exists a value function

$$\hat{V}(x, \hat{\mu}_0) = \sum_{\hat{u} \in \hat{\mathcal{U}}^*} \max_{p \in x} [\hat{u}(p_k) + \hat{\delta} \hat{V}(p_z, \hat{u})] d\hat{\mu}_0(\hat{u})$$

that represents  $\succsim$  and features a constant discount factor. The uniqueness result in DLR implies that  $\hat{\mathcal{U}}^*$  and  $\mathcal{U}^*$  must correspond to the same collection of vN-M rankings. That is, there is a reordering of  $\hat{\mathcal{U}}^*$  such that  $\xi(u_i) := u_i/\hat{u}_i$  is well defined for  $i = 1, \dots, m$ .

With  $\mu_i(\cdot) := M(u_i, \cdot)$ , we write  $\langle \mu_i, \xi \rangle$  to denote  $\sum_j \mu_i(u_j)\xi(u_j)$ . Then  $\hat{\mu}_i(\hat{u}) = \frac{\mu_i(u)\xi(u)}{\langle \mu_i, \xi \rangle}$  must hold for all  $\hat{\mu}_i$ ,  $i = 0, 1, \dots, m$  and clearly

$$\sum_{\hat{u} \in \hat{\mathcal{U}}^*} \max_{p \in x} \left[ \hat{u}(p_k) + \frac{\delta(u)}{\xi(u)} V(p_z, u) \right] \hat{\mu}_0(\hat{u})$$

represents  $\succsim$ , as it is a renormalization of  $V(x, \mu_0)$ . Therefore,  $\hat{\delta} \hat{V}(\cdot, \hat{u}) = \frac{\delta(u)}{\xi(u)} V(\cdot, u)$  must hold for all  $u \in \mathcal{U}^*$ . At the same time  $\hat{V}(\cdot, \hat{u}_i) = \frac{V(\cdot, u_i)}{\langle \mu_i, \xi \rangle}$ . Hence,  $\hat{\delta} = \frac{\delta(u_i)}{\xi(u_i)} \langle \mu_i, \xi \rangle$ . To establish that the value function  $\hat{V}$  exists, we have to show that there is  $\xi$  such that  $\frac{\delta(u_i)}{\xi(u_i)} \langle \mu_i, \xi \rangle$  is constant for all  $u_i \in \mathcal{U}^*$ .

Letting  $A$  denote the Markov transition matrix for the Markov process  $M$ , the condition amounts to finding  $\xi$  such that  $\kappa\xi = \Delta A\xi$  for some  $\kappa > 0$ , where  $\Delta$  is the diagonal matrix<sup>11</sup>

$$\begin{bmatrix} \delta(u_1) & 0 & \dots & 0 \\ 0 & \delta(u_1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \delta(u_m) \end{bmatrix}$$

That is, we need a strictly positive eigenvector  $\xi$  of the matrix  $\Delta A$ . Persistent Preference for Flexibility (Axiom 10) implies that  $A$  (and hence  $\Delta A$ ) is strictly positive. The Perron Theorem (Theorem 3 in appendix B.3) then implies that there is such an eigenvector  $\xi$  that is strictly positive, and hence there is a representation  $\hat{V}$  as above. It also implies that the eigenvector  $\xi$  is unique up to scaling and the corresponding eigenvalue  $\hat{\delta}$  is unique. This implies that beliefs in the representation  $\hat{V}$ , ie, the collection  $\{\hat{\mu}_i : i = 0, \dots, m\}$ , must be unique up to a common scaling. The proof generalizes this argument to apply to the case where  $\mathcal{U}^*$  may contain collinear elements (and hence may not be finite).

To see directly why uniqueness of beliefs up to scaling must hold, consider the example of two evolving tastes representations of the same preferences,  $V$  and  $\hat{V}$  as above, where  $\mathcal{U}^*$  does not contain collinear elements and  $\delta$  and  $\hat{\delta}$  are constant. Suppose, without loss of generality, that  $1 > \delta \geq \hat{\delta}$ . We argued above that  $\xi(u_i) = \frac{\delta}{\hat{\delta}} \langle \mu_i, \xi \rangle$  must hold. The Markov process  $M$  is ranking persistent, and hence the support of  $\mu_i$  is  $\mathcal{U}^*$  for all  $i \in \{0, \dots, m\}$ . Suppose  $\xi(u_i) \neq \xi(u_j)$  for some  $i, j \in \{0, \dots, m\}$ . Pick the  $i$  that minimizes  $\xi(u_i)$  and observe that  $\xi(u_i) < \langle \mu_i, \xi \rangle \leq \frac{\delta}{\hat{\delta}} \langle \mu_i, \xi \rangle$ , a contradiction. Hence,  $\xi(u_i) = \xi(u_j)$  for all  $i, j \in \{0, \dots, m\}$ , which just says that  $\{\mu_i : i = 0, \dots, m\}$  and  $\{\hat{\mu}_i : i = 0, \dots, m\}$  must be identical up to a common scaling.

To better understand why Persistent Preference for Flexibility (Axiom 10) is necessary, we now provide an example of a preference that satisfies all axioms except Axiom 10, and show that it does not have an evolving tastes representation. Suppose  $\succsim$  can be represented by the value function  $V$  above where  $\mathcal{U}^* = \{u_1, u_2\}$  is the support of  $\mu_0$ ,  $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ , and  $\delta(u_1) \neq \delta(u_2)$ . The interpretation is that

(11) It is easy to see that with  $\Delta$  and  $A$  defined as above,  $\Delta A$  is the square matrix where the row of  $A$  that corresponds to  $u$  is multiplied by  $\delta(u)$ .

today DM is uncertain about tomorrow's consumption utility, but once he learns his utility, he does not expect it to ever change again. In that case, there is no representation with a constant discount factor, because  $\frac{\delta(u_i)}{\xi(u_i)}(\mu_i, \xi) = \delta(u_i)$  for all positive  $\xi$ .

### 3. Correlated Tastes and Delayed Commitment

Tastes in our model are correlated over time. Thus, DM's knowledge of his taste at one point in time might reduce his uncertainty about future tastes. This implies that DM's willingness to commit to future consumption upon learning his current taste will depend on the *degree* to which his tastes are correlated over time. We now formalize this intuition as a basis for comparing decision makers.<sup>12</sup>

Given our model, we calculate the discounted expected future taste contingent on the initial taste realization, averaged over time and over the realizations of the Markov transitions. The distribution over initial tastes thus generates a distribution over discounted expected future tastes. The more informative initial taste realizations are about discounted expected future tastes, the more disperse this distribution will be. This is how we propose to measure the correlation of tastes over time.

Let  $Z_S$  be the collection of time separable choice problems, in which choice does not affect the distribution over continuation problems, and hence consumption choice does not involve any commitment. Thus,  $Z_S = \mathcal{F}_K^\infty$ .

Consider DM and  $DM^\dagger$ , with preferences  $\succsim$  and  $\succsim^\dagger$ , and evolving tastes representations  $((\mathcal{U}_M, M), \delta)$  and  $((\mathcal{U}_{M^\dagger}, M^\dagger), \delta^\dagger)$ , where  $\mu_0$  and  $\mu_0^\dagger$  are the ergodic distributions of  $M$  and  $M^\dagger$ , respectively. In order to compare DM and  $DM^\dagger$ , we assume that both agree on the ranking of time-separable problems which, by construction, do not offer any opportunity to commit after learning the initial taste.

**Definition 3.1.** DM and  $DM^\dagger$  *agree on separable problems* if for all  $z, z' \in Z_S$ ,  $z \succsim z'$  if and only if  $z \succsim^\dagger z'$ .

---

(12) It is also going to be true that a decision maker with more uncertain tastes will have stronger preference for flexibility, which is in complete analogy to Theorem 2 in KS. In their model, taste shocks are transient, which precludes the type of comparison considered here.

**Proposition 3.2.** DM and  $DM^\dagger$  agree on separable problems if and only if (i)  $\delta = \delta^\dagger$ , and (ii) there is  $\zeta > 0$  such that for all  $u \in \mathcal{U}$ ,

$$\int_{\{\lambda u: \lambda \geq 0\}} \lambda \, d\mu_0(\lambda u) = \zeta \int_{\{\lambda u: \lambda \geq 0\}} \lambda \, d\mu_0^\dagger(\lambda u)$$

A proof is in Appendix C. The proposition establishes that both representations discount the future at the same rate, and (up to a positive scalar) they also agree on the ex ante weight they assign to each vN-M ranking, where the weight is the aggregate of the probability assigned to all utilities that correspond to that ranking, weighted by their scaling.

This allows us to normalize the representations so that  $\zeta = 1$  in Proposition 3.2. Given this normalization, we now introduce our measure of the correlation of tastes over time.

For the evolving tastes representation  $((\mathcal{U}_M, M), \delta)$  with ergodic distribution  $\mu_0$ , for each  $p_k \in \mathcal{P}(K)$ , let  $U(\cdot, p_k) : \mathcal{U}_M \rightarrow \mathbb{R}$  be defined as

$$[3.1] \quad U(u, p_k) := u(p_k) + \delta \int_{\mathcal{U}_M} U(u', p_k) M(u, du')$$

This lets us define the function  $\tilde{U}(\cdot) : \mathcal{U}_M \rightarrow \mathcal{U}$  as

$$[3.2] \quad \tilde{U}(u) := (U(u, k))_{k \in K}$$

Let  $\tilde{\mu}$  be the probability measure on  $\mathcal{U}$  that is defined as

$$[3.3] \quad \tilde{\mu}(E) := \mu_0(\{u : \tilde{U}(u) \in E\})$$

for all measurable  $E \subset \mathcal{U}$ .

In words,  $\tilde{U}(u)$  is the discounted expected future taste (in present value terms) contingent on taste  $u$ . The measure  $\tilde{\mu}$  is the measure over these discounted expected tastes induced by the ergodic distribution over initial tastes.

**Definition 3.3.**  $M$  is  $\delta$ -more correlated than  $M'$  if  $\tilde{\mu}$  is a dilation of  $\tilde{\mu}^\dagger$ .<sup>13</sup>

Let  $Z_{\text{iid}}$  be menus of iid lotteries over  $K$ , where an iid lottery is a sequence of lotteries of the form  $\mathbf{p}_k := (p_k, p_k, \dots)$  where  $p_k \in \mathcal{P}(K)$ . Then,  $Z_{\text{iid}} = \mathcal{F}(\{\mathbf{p}_k :$

(13) Let  $Q$  be a Markov kernel from  $\mathcal{U}$  to itself. Then,  $Q$  is a *dilation* if it preserves expectations, ie,  $\int_{\mathcal{U}} u' Q(u, du') = u$  for all  $u \in \mathcal{U}$ . For probability measures  $\mu$  and  $\mu^\dagger$  on  $\mathcal{U}$ ,  $\mu$  is a dilation of  $\mu^\dagger$  if there is a dilation  $Q$  such that  $\mu(du') = \int Q(u, du') \mu^\dagger(du)$ .



$p_k \in \mathcal{P}(K)\}$ ), and consists of sets of consumption lotteries where the chosen lottery will be consumed in every future period, and so represents perfect commitment. We can now relate our notion of correlation over time to our behavioral comparison.

**Definition 3.4.** DM is *greater preference for delayed commitment* than  $DM^\dagger$  if  $z \succsim^\dagger y$  implies  $z \succsim y$  for all  $z \in Z_{\text{iid}}$  and  $y \in Z_S$ .

**Theorem 2.** *Suppose DM and  $DM^\dagger$  agree on separable problems (and hence  $\delta = \delta^\dagger$ ). Then DM has greater preference for delayed commitment than  $DM^\dagger$  if and only if  $M$  is  $\delta$ -more correlated than  $M'$ .*

A proof is in Appendix C. We argued at the outset of this section that it is plausible that a decision maker whose tastes are more correlated over time is more open to commit to a consumption choice after initially learning his taste. The theorem spells out the correct notion of correlation of tastes.

#### 4. Related Literature

Rather than passively learning about the taste, as in our model, a decision maker might actively contemplate his taste, where such contemplation is costly. Ergin and Sarver (2010) provide a model of costly contemplation in the context with only one instance of consumption choice, as first analyzed by Dekel, Lipman, and Rustichini (2001). A related model in a dynamic context is provided by Dillenberger, Krishna, and Sadowski (2016), where the decision maker tries to learn about the verifiable state of the world, rather than his own taste.

Piermont, Takeoka, and Teper (2015) consider a forward looking decision maker with a stable but unknown taste, who expects to learn through exploration. That is, rather than experiencing changing tastes, the decision maker becomes more and more informed about his stable taste as he experiences different consumption alternatives.

Tastes could also change in response to consumption, as in the context of habit forming preferences, where the forward looking DM understands how his consumption choice today will affect preferences in the future. A dynamic model of habit formation along the lines of the model presented in this paper is the subject of a separate research project.

A forward looking DM who expects his tastes to change over time is affected by his attitude towards those changes. In this paper we consider one extreme, where DM embraces changes and tries to accommodate them. In the other extreme he may resent changing tastes; he acknowledges that his choice in the future will be governed by the changed tastes, but in the current period he evaluates this future choice based on the current taste. As a consequence, he wants to commit to a future choice now. This is the familiar intuition behind models of time inconsistent preferences with sophistication, as for example in O’Donoghue and Rabin (2000). Based on the model by Strotz (1955), in which preferences change over time in a deterministic fashion, Gul and Pesendorfer (2005) analyze how such a decision maker chooses between different dynamic decision problems that are similar to ours.

Of course, it is plausible that a decision maker would find it desirable to retain the flexibility to react to some aspects of his changing tastes (as in our model), while at the same time seeking to commit not to accommodate other changes (as in GP). For example, he might want to accommodate changes in risk aversion while committing not to accommodate changes in his discount factor. We leave an investigation of the interaction of the two attitudes towards changing tastes as another topic for future research.

## Appendices

### A. Proof of Lemma 2.5

By definition of the Markov process  $(\mathcal{U}_M, M)$ , there exist  $u_1, u_2, \dots, u_n \in \mathcal{U}$  such that  $\mathcal{U}_M := \text{cone}(\{u_1, \dots, u_n\})$ . Denoting  $[u_i] = \{\lambda u_i : \lambda > 0\}$ , we see an induced Markov chain on the rankings, with state space  $\{[u_i] : i = 1, \dots, n\}$ , and with transition probabilities  $M([u_i], [u_j])$ , where we have abused notation, because  $M(u_i, \cdot) = M(\lambda u_i, \cdot)$  for all  $\lambda > 0$ . This is a Markov chain on a finite state space, and because all transition probabilities are strictly positive, there is a unique invariant distribution  $v_0([u_i])$ .

Let  $\nu$  be any measure on  $\mathcal{U}_M$  such that  $\nu([u_i]) = v_0([u_i])$  for  $i = 1, \dots, n$ .

Define the measure  $\mu_0$  on  $\mathcal{U}_M$  as follows: for any measurable  $D \subset \mathcal{U}_M$ ,

$$\begin{aligned}\mu_0(D) &:= \sum_{i=1}^n \left[ \int_{[u_i]} M(u, D) v(du) \right] \\ &= \sum_i M(u_i, D) \left[ \int_{[u_i]} v(du) \right] \\ &= \sum_i M(u_i, D) v([u_i])\end{aligned}$$

It is easy to see that  $\mu_0([u_i]) = v_0([u_i])$  for  $i = 1, \dots, n$ . Moreover,  $\mu_0(D)$  is independent of the choice of the measure  $v$ , as long as  $v$  satisfies  $v([u_i]) = v_0([u_i])$  for  $i = 1, \dots, n$ . Therefore,  $\mu_0$  is the unique invariant measure of the Markov process  $(\mathcal{U}_M, M)$ .

## B. Proof of Theorem 1

Some steps of our proof parallel the proof of the representation theorem in KS. We refer to formal statements in KS by adding the letters KS to their numbering. Thus, Lemma nKS refers to Lemma n in KS, and similarly for Propositions and Theorems.

### B.1. Separable Representation: Existence and Identification

The relevant domain here is  $Z \simeq \mathcal{F}(\mathcal{P}(K \times Z))$ . Lemma 20KS establishes that if  $\succsim$  has a finitely additive EU representation (Theorem 3KS) and satisfies Separability (Axiom 5), then it has a separable representation of the form

$$[\mathbf{B.1}] \quad W(x) = \int_{\{u \in \mathcal{U}_{K \times Z} : u(p) = u(p_k) + v(p_z)\}} \max_{p \in x} [u(p_k) + v(p_z)] d\mu(u)$$

**Lemma B.1.** Let  $\succsim$  have a separable representation as in [B.1]. If  $\succsim_K$  is as in definition 2.2, then  $\succsim_K$  is independent of the choice of  $A \in \mathcal{F}_Z$ .

*Proof.* Let  $W$  be as in [B.1]. Then, for any  $a, b \in \mathcal{F}_K$ , and  $A, B \in \mathcal{F}_Z$ , we have

$$\begin{aligned}W((a, A)) &= \int_{\{u \in \mathcal{U}_{K \times Z} : u(p) = u(p_k) + v(p_z)\}} \max_{p_k \in a; p_z \in A} [u(p_k) + v(p_z)] d\mu(u) \\ &=: \varphi_K(a) + \varphi_Z(A)\end{aligned}$$

where  $\varphi_K : \mathcal{F}_K \rightarrow \mathbb{R}$  and  $\varphi_Z : \mathcal{F}_Z \rightarrow \mathbb{R}$ . By construction,  $\varphi_K$  represents  $\succsim_K$ , which is independent of the choice of  $A$ , which completes the proof.  $\square$

By a variation of Proposition 20KS, we can transform the state space in [B.1] to be  $\mathcal{U}_K \times [0, 1] \times \mathcal{U}_Z$ . In what follows, let  $\nu$  be the marginal of  $\mu$  on  $\mathcal{U}_K$ .

**Lemma B.2.** Let  $\succsim$  have a separable representation as in (B.1). If  $\succsim_K$  satisfies Finiteness (Axiom 8), then the support of the marginal of  $\mu$  on  $\mathcal{U}_K$  is finite.<sup>14</sup>

*Proof.* Let  $\varphi_K$  represent  $\succsim_K$ . By lemma B.1, it follows that  $\varphi_K$  is well defined. As  $\varphi_K$  is continuous, linear, and monotone, it has an additive EU representation of the form  $\varphi_K(a) = \int_{\mathcal{U}_K} \max_{p_k \in a} u(p_k) d\nu(u)$  where  $\nu$  is countably additive and is the marginal of  $\mu$  on  $\mathcal{U}_K$ .

To see that  $\nu$  must have finite support, suppose not. Let  $a$  be the ball of radius  $\varepsilon$  around  $p_k^*$ , the uniform lottery over  $K$ . It is easy to see that because  $\nu$  has infinite support, there can be no finite  $b \subset a$  such that  $\varphi_K(b) = \varphi_K(a)$ , which contradicts Finiteness (Axiom 8).  $\square$

**Lemma B.3.** Suppose  $\succsim$  has a separable representation as in [B.1], and also satisfies Finiteness (Axiom 8). Fix  $r^* \in \mathcal{U}_K$  such that  $\nu(r^*) > 0$ . Let  $a, b \in \mathcal{F}_K$  be such that  $\max_{\alpha \in a} r^*(\alpha) > \max_{\alpha \in b} r^*(\alpha)$  and  $\max_{\alpha \in a} r(\alpha) < \max_{\alpha \in b} r(\alpha)$  for all  $r \neq r^*$ . For such  $a, b \in \mathcal{F}_K$ , we may take  $c = b$  in Choice Contingent CSR (Axiom 9).

*Proof.* By hypothesis,  $\succsim$  has a representation of the form

$$W(x) = \int_{\mathcal{U}_K \times [0,1] \times \mathcal{U}_Z} \max_{p \in x} [\gamma r(p_k) + (1 - \gamma)v(p_z, r)] d\mu(r, \gamma, v)$$

where the marginal of  $\mu$  on  $\mathcal{U}_K$  has finite support. The representation implies that if  $a \cup b \cup c \succ_K b \cup c$ , then  $a$  dominates  $b \cup c$  in state  $r^*$ , while  $c$  possibly dominates  $b$  in relevant states  $r \neq r^*$ . However, in those states,  $b$  already dominates  $a$ , and thus

$$(a, A) \overset{\circ}{\succsim}_{(b \cup c, A \cup B)} (a, B) \text{ implies } (a, A) \overset{\circ}{\sim}_{(b \cup c, A \cup B)} (a, A \cup B)$$

if, and only if,

$$(a, A) \overset{\circ}{\succsim}_{(b, A \cup B)} (a, B) \text{ implies } (a, A) \overset{\circ}{\sim}_{(b, A \cup B)} (a, A \cup B)$$

Thus, under the conditions stated, we may take  $c = b$  in Axiom 9.  $\square$

(14) Riella (2013) offers a different proof of this claim.

**Proposition B.4.** Let  $\succsim$  have a separable representation as in equation [B.1], and suppose it also satisfies Finiteness (Axiom 8) and Choice Contingent CSR (Axiom 9). In that case, for all  $u$  in the carrier of  $\mu$ , the induced marginal charge  $\mu(\cdot|u)$  on  $\mathfrak{U}_Z$  has singleton support. Moreover, if  $\lambda > 0$  and  $u, \lambda u$  are in the carrier of  $\mu$ , then  $\mu(\cdot|u) = \mu(\cdot|\lambda u)$ .

*Proof.* As before, we may regard  $\mu$  as a charge on  $\mathfrak{U}_K \times [0, 1] \times \mathfrak{U}_Z$ . We will show that for each  $r^* \in \mathfrak{U}_K$ , the induced marginal charge  $\mu(\cdot|r^*)$  has singleton support in  $[0, 1] \times \mathfrak{U}_Z$ . Proposition 20KS implies that the separable representation can be written as

$$W(x) = \int_{\mathfrak{U}_K \times [0, 1] \times \mathfrak{U}_Z} \max_{p \in x} \left[ r(p_k) + \frac{1-\gamma}{\gamma} v(p_z) \right] \gamma \, d\mu(r, \gamma, v)$$

Fix  $r^*$  in the finite support of the marginal of  $\mu$  on  $\mathfrak{U}_K$ . Let  $\varepsilon > 0$  be such that  $b := \{\alpha \in \mathfrak{U}_K : \|\alpha - p_k^*\|_2 \leq \varepsilon\}$  is contained in the interior of  $\mathcal{P}(K)$ .<sup>15</sup> Then, there exists  $\alpha \in \mathcal{P}(K)$  that is superior to all alternatives in  $b$  in consumption taste  $r^*$ , while for all other relevant consumption utilities, some alternative from  $b$  is preferred to  $\alpha$ . Letting  $a := \{\alpha\}$ , we see that  $a \cup b \succ_K b$ . Choice Contingent CSR (Axiom 9) now implies that there exists  $c \in \mathcal{F}_K$  such that (i)  $a \cup b \cup c \succ_K b \cup c$  and (ii)  $(a, A) \overset{\circ}{\succsim} (b \cup c, A \cup B)(A, B)$  implies  $(a, A) \overset{\circ}{\sim} (b \cup c, A \cup B)(A, B)$ . By the construction of  $a$  and  $b$ , and by Lemma B.3, we may assume, without loss of generality, that Choice Contingent Strategic Rationality (Axiom 9) is satisfied for  $c = b$ .

Property (i) of the axiom is trivially satisfied. To check (ii), consider arbitrary  $A, B \in \mathcal{F}_Z$  and define  $x := (a, A) \cup (b, A \cup B)$  and  $y := (a, B) \cup (b, A \cup B)$ . We may assume, without loss of generality, that  $x \succsim y$ . Then,  $W(x) = r^*(\alpha) + \psi(A) + \kappa \geq r^*(\alpha) + \psi(B) + \kappa = W(y)$ , where the  $\kappa$  is the utility from consumption in all states other than  $r^*$ , and is the same in both menus  $x$  and  $y$ , and where for any menu  $A \in \mathcal{F}_Z$ ,

$$\psi(A) := \int_{[0, 1] \times \mathfrak{U}_Z} \frac{1-\gamma}{\gamma} \max_{p_z \in A} v(p_z) \gamma \, d\mu(\gamma, v|r^*)$$

Notice that  $W(x) \geq W(y)$  if, and only if,

$$\text{[B.2]} \quad \psi(A) \geq \psi(B)$$

---

(15) Recall that  $p_k^*$  is the uniform lottery over  $K$ .

Consider the menu  $y' := (a \cup b, A \cup B)$ . Since the menus  $A, B \in \mathcal{F}_Z$  are such that  $x \succsim y$ , Choice Contingent CSR (Axiom 9) implies that  $x \sim y'$ , which requires  $W(x) = r^*(\alpha) + \psi(A) + \kappa = r^*(\alpha) + \psi(A \cup B) + \kappa = W(y')$  and is equivalent to

$$\text{[B.3]} \quad \psi(A) = \psi(A \cup B)$$

The function  $\psi$  is a utility function on  $\mathcal{F}_Z$ . It is easily seen that  $\psi$  is linear, and monotone, and induces a preference on  $\mathcal{F}_Z$  that is continuous, satisfies Independence (Axiom 3), and Monotonicity (Axiom 4). Displays [B.2] and [B.3] now imply that the preference induced by  $\psi$  is *strategically rational*, and so for each  $r^*$ , the marginal of  $\mu$  on  $\mathcal{U}_Z$  has singleton support (see Proposition 15KS), which completes the proof.  $\square$

Proposition B.4 implies that there exists a continuous  $v : Z \times \mathcal{U} \rightarrow \mathbb{R}$  such that for all  $(u, v') \in \mathcal{U} \times \mathcal{U}_Z$ ,  $v'(z) = \kappa(u)v(z, u)$  (up to adding a constant),  $\mu$ -almost surely. Moreover, for  $u, \lambda u$  in the carrier of  $\mu$ , where  $\lambda > 0$ , it follows that  $v(\cdot, u) = v(\cdot, \lambda u)$ .

The proposition further implies that  $W$  is  $Z$ -simple in the sense that the marginal of  $\mu$  on  $\mathcal{U}_Z$  has finite support. It follows from Proposition 23KS that the representation is jointly identified and  $\mu$  is a regular probability measure on  $\mathcal{U}_K \times [0, 1] \times \mathcal{U}_Z$ . We can rewrite the representation, transforming  $\mathcal{U}_K \times [0, 1] \times \mathcal{U}_Z$  to become  $\mathcal{U} \times \mathcal{U}_Z$ . Proposition B.4 implies that each consumption state  $u \in \mathcal{U}$  corresponds to a unique continuation utility function  $v(\cdot, u)$ . Hence, we may assume that the state space is  $\mathcal{U}$  and let  $\mu_0$  be the corresponding measure on  $\mathcal{U}$ .

Let  $\mathcal{U}^* := \text{supp}(\mu_0)$  be the support of  $\mu_0$ . By Proposition B.4, there exists  $\{u_1, \dots, u_n\} \subset \mathcal{U}$  such that  $\mathcal{U}^* \subset \mathcal{U}_M = \bigcup_{\lambda > 0} \lambda \{u_1, \dots, u_n\}$ . As before, we write  $[u] := \{\lambda u : \lambda > 0\}$  for all  $u \in \mathcal{U}_M$ . Intuitively,  $[u]$  is an equivalence class of consumption utilities, all of which induce the same continuation utility. Thus, we may write  $W(x)$  as

$$\text{[B.4]} \quad W(x) = \int_{\mathcal{U}} \max_{p \in x} [u(p_k) + v(p_z, [u])] \, d\mu_0(u)$$

**Proposition B.5.** In the representation in [B.4], the marginal of  $\mu_0$  on each  $[u_i]$  is identified uniquely up to scaling.

*Proof.* Define

$$W_i(x) = \int_{[u_i]} \max_{p \in x} [u(p_k) + v(p_z, [u_i])] \, d\mu_0(u)$$

so that  $W(x) = \sum_i W_i(x)$ . Arguments analogous to those in Proposition 25KS establish that the marginal of  $\mu_0$  on each  $[u_i]$  is identified up to scaling, which proves the proposition.  $\square$

## B.2. Recursive, Uniformly Ranking Persistent Representation

To show that a recursive representation exists, we need to first show that the functions  $v(\cdot, [u_i]) : Z \rightarrow \mathbb{R}$  in [B.4] are linear. Towards this end, consider the mapping  $v : Z \rightarrow \mathbb{R}^n$ , defined as follows:  $v_i(x) := v(x, [u_i])$ . Let  $z^* \simeq (\mathcal{P}(K), z^*)$  be the IHCP that gives all the lotteries in each period. By definition of  $S_Z$ , we have  $v_i(z^*) = 1$  and  $v_i(x^*) = 0$  for all  $i$  (where  $x^* \simeq (p_k^*, x^*)$  and  $p_k^*$  is the uniform lottery over  $K$  so that  $x^*$  gives this lottery in each period).

Let  $O_i := v_i^{-1}(\text{int } v_i(Z))$ , and let  $O := \bigcap_i O_i$ . Since each  $v_i$  is continuous, it follows that each  $O_i$  and hence  $O$ , is open. Let  $D_i := Z \setminus \text{cl } O_i$ . To show that each  $v(\cdot, [u_i])$  is linear on  $Z$ , we need to show that it is locally non-satiated, which amounts to  $D_i = \emptyset$ . To see this, we need the following lemma.

**Lemma B.6.** For every  $x \in Z$  and for every open neighbourhood  $N \ni x$ , there exist  $y, z \in N$  such that (i)  $y \subset z$ , and (ii)  $z \succ y$ .

*Proof.* Fix  $x \in Z$  and let  $N \ni x$  be open. Then, there exists  $y \in N$  such that  $y$  is not a  $\succsim$ -maximal or  $\succsim$ -minimal element in  $Z$ . Recall that  $z^* \simeq (\mathcal{P}(K), z^*)$ , and by definition,  $z^* \supset y$ . Moreover,  $z^* \succ y$ . Therefore, there exists  $\lambda \in (0, 1)$  such that (i)  $\lambda z^* + (1 - \lambda)y \subset N$  (by the definition of the Hausdorff metric), and (ii)  $\lambda z^* + (1 - \lambda)y \succ y$  (by Independence, Axiom 3). Setting  $z := \lambda z^* + (1 - \lambda)y$  completes the proof.  $\square$

**Lemma B.7.** Let  $\succsim$  have a representation as in [B.4], and suppose  $\succsim$  satisfies Persistent Preference for Flexibility (Axiom 10), Then,  $D_i = \emptyset$  for each  $i = 1, \dots, n$ .

*Proof.* Notice that by definition,  $D_i$  is the union of two disjoint, connected sets that are open. The function  $v_i$  achieves its maximum and minimum values on these components, and is therefore constant on each of the components of  $D_i$ .

Suppose  $D_i \neq \emptyset$ , ie, suppose  $x \in D_i$ , and let  $N$  be an open neighbourhood of  $x$  such that  $x \in N \subset D_i$  and  $v_i$  is constant on  $N$ . Then, by Lemma B.6, there exist  $y, z \in N$  where  $y \subset z$  and  $z \succ y$ . By Persistent Preference for Flexibility

(Axiom 10), we must then have  $v_i(z) > v_i(y)$ , which contradicts the fact that  $v_i$  must be constant on  $N$ , completing the proof.  $\square$

If each  $D_i$  is empty, then each  $O_i$  is open and dense in  $Z$ , which implies that  $O = \bigcap_i O_i$  is also dense in  $Z$ . Therefore, to show that each  $v_i$  is linear in  $Z$ , it suffices to show that  $v_i$  is linear on  $O$ . We establish this next.

**Definition B.8.** A pair of sets  $(\mathbf{x}, \mathbf{y}) \in Z^2$  is *amenable* if  $\mathbf{x} := \{x_i \in Z : i = 1, \dots, n\}$  and  $\mathbf{y} := \{y_i \in Z : i = 1, \dots, n\}$  and if the following hold:

- $v_i(x_i) > v_i(x_j)$  for all  $j \neq i$ , and
- $v_i(y_i) > v_i(y_j)$  for all  $j \neq i$

Because  $O$  is open, given  $x, y \in O$ , we can construct an amenable pair  $(\mathbf{x}, \mathbf{y})$  such that  $x = x_i \in \mathbf{x}$  and  $y = y_i \in \mathbf{y}$ . In particular, this demonstrates that amenable sets exist.

For each  $\lambda \in [0, 1]$ , define  $\mathbf{z}_\lambda := \lambda \mathbf{x} + (1 - \lambda) \mathbf{y}$ . As  $\mathbf{x}$  and  $\mathbf{y}$  consist of degenerate lotteries,  $\mathbf{z}_\lambda$  also consists of degenerate lotteries. For each  $i = 1, \dots, n$ , define  $\Psi_i : [0, 1] \rightarrow \mathbf{x} \times \mathbf{y}$  as follows:

$$\Psi_i(\lambda) := \left\{ (x, y) : x \in \mathbf{x}, y \in \mathbf{y}, \lambda x + (1 - \lambda)y \in \arg \max_{z \in \mathbf{z}_\lambda} v_i(z) \right\}$$

Let  $\Psi_i(\lambda) = (\Psi_{i,\mathbf{x}}, \Psi_{i,\mathbf{y}})$ . We shall establish some properties of  $\Psi_i$  for amenable pairs  $(\mathbf{x}, \mathbf{y})$ .

**Proposition B.9.** Let  $\succsim$  have a representation as in [B.4], and suppose  $\succsim$  satisfies Indifference to Timing (Axiom 7). Then, the correspondence  $\Psi$  has the following properties:

- (a)  $\Psi_i$  is ‘onto’. That is, for each  $\lambda \in [0, 1]$  and  $x \in \mathbf{x}$ , there exists  $j$  such that  $x_i \in \Psi_{j,\mathbf{x}}(\lambda)$ , with a similar claim for  $y_i \in \mathbf{y}$ .
- (b)  $\Psi_i$  is a function.
- (c)  $\Psi_i$  is continuous.
- (d)  $\Psi_i$  is constant, ie, is independent of  $\lambda$ .

*Proof.* In the proof, we make repeated use of the fact that by Indifference to Timing (Axiom 7),  $\lambda W((k, \mathbf{x})) + (1 - \lambda)W((k, \mathbf{y})) = W((k, \mathbf{z}_\lambda))$  for all  $\lambda \in [0, 1]$ .

- (a) Suppose not, so that  $x_i \notin \Psi_{j,\mathbf{x}}$  for any  $j = 1, \dots, n$ . Then, by perturbing  $x_i$ , we obtain a contradiction to the equality  $\lambda W((k, \mathbf{x})) + (1 - \lambda)W((k, \mathbf{y})) = W((k, \mathbf{z}_\lambda))$  and the fact that  $(\mathbf{x}, \mathbf{y})$  is an amenable pair.



- (b) Suppose not, so for some  $\lambda \in (0, 1)$ ,  $\Psi_i(\lambda) = \{(x_j, y_j) : j = 1, \dots, m\}$ . Let  $\lambda^* := \inf\{\lambda : \Psi_i(\lambda) \text{ is not a singleton}\}$ . It is easy to see that  $\lambda^* > 0$ . It is also easy to see that we may choose, without loss of generality,  $\mathbf{x}$  and  $\mathbf{y}$  such that there exists a unique  $i$  where  $\Psi_i(\lambda^*)$  is not a singleton. Since  $\Psi$  is onto (as established above), we can perturb one of the elements of  $\Psi_i(\lambda^*)$  without affecting  $W((k, \mathbf{z}_{\lambda^*}))$ , but affecting  $\lambda^* W((k, \mathbf{x})) + (1 - \lambda^*) W((k, \mathbf{y}))$ , which is a contradiction.
- (c) This is a simple consequence of the Theorem of the Maximum.
- (d) This follows because  $\Psi_i$  is continuous,  $[0, 1]$  is connected, and  $\mathbf{x}$  and  $\mathbf{y}$  are finite sets.  $\square$

**Proposition B.10.** Let  $\succsim$  have a representation as in [B.4], and suppose  $\succsim$  satisfies Indifference to Timing (Axiom 7) and Persistent Preference for Flexibility (Axiom 10). Then, each  $v_i$  is linear on  $Z$ .

*Proof.* Recall that  $Z$  is compact and  $v_i$  is continuous on  $Z$ , which implies that  $v_i$  is, in fact, uniformly continuous on  $Z$ . Lemma B.7 says that because  $\succsim$  satisfies Persistent Preference for Flexibility (Axiom 10),  $O$  is dense in  $Z$ . Therefore, it suffices to show that  $v_i$  is linear on  $O$ , because by lemma 3.8 of Aliprantis and Border (1999),  $v_i$  has a unique continuous extension to  $Z$  (which must also be linear).

Suppose now that there exist  $x, y \in O$  such that  $v_i(\lambda x + (1 - \lambda)y) \neq \lambda v_i(x) + (1 - \lambda)v_i(y)$ . By Indifference to Timing (Axiom 7), we have

$$[\star] \quad W((k, \mathbf{z}_\lambda)) = \lambda W((k, \mathbf{x})) + (1 - \lambda) W((k, \mathbf{y}))$$

for an amenable pair  $(\mathbf{x}, \mathbf{y})$  with  $x = x_i \in \mathbf{x}$  and  $y = y_i \in \mathbf{y}$ . By the properties of the function  $\Psi$ , we see that

$$\begin{aligned} W((k, \mathbf{z}_\lambda)) &= \kappa + v_i(\lambda x + (1 - \lambda)y) \\ \lambda W((k, \mathbf{x})) + (1 - \lambda) W((k, \mathbf{y})) &= \iota + \lambda v_i(x) + (1 - \lambda)v_i(y) \end{aligned}$$

where  $\kappa$  and  $\iota$  do not depend on  $x$  and  $y$  (locally). It is easy to see that we can now perturb  $x$  (say) such that  $[\star]$  no longer holds, which proves our claim.  $\square$

**Proposition B.11.** Let  $\succsim$  have a representation as in [B.4] where  $v(\cdot, [u])$  is linear, and suppose  $\succsim$  satisfies Stationarity (Axiom 6) and Persistence (Axiom 10). Then,

there is a value function

$$[\mathbf{B.5}] \quad V(x, u) = \int_{\mathcal{U}} \max_{p \in x} [u'(p_k) + \delta(u')V(p_z, u')] M(u, du')$$

for all  $u \in \mathcal{U}_M$ , where  $(\mathcal{U}_M, M)$  is a ranking contingent Markov process where  $\text{supp}(\mu_0) \subset \mathcal{U}_M$ , such that

$$[\mathbf{B.6}] \quad V(x, \mu_0) := \int_{\mathcal{U}} \max_{p \in x} [u(p_k) + \delta(u)V(p_z, u)] d\mu_0(u)$$

is a representation of  $\succsim$ .

*Proof.* For rectangular menus of the form  $\{(k, x)\}$ , [\[B.4\]](#) becomes

$$\begin{aligned} W(\{(k, x)\}) &= \int_{\mathcal{U}_M} [u(k) + v(x, [u])] d\mu_0(u) \\ &= \sum_{[u] \subset \mathcal{U}_M} v(x, [u]) \mu_0([u]) + \int_{\mathcal{U}_M} u(k) d\mu_0(u) \end{aligned}$$

where  $v(\cdot, [u])$  is continuous and linear. The last equality follows because  $\mathcal{U}_M = \bigcup_{i=1}^n [u_i]$ . Notice that the second term does not depend on  $x$ . By [Theorem 3](#), each  $v(\cdot, [u])$  induces a preference  $\succsim_{[u]}$  on  $\mathcal{F}(\mathcal{P}(K \times Z))$  that is continuous and satisfies Independence ([Axiom 3](#)). We claim:

**Claim B.12.** For each  $u \in \mathcal{U}_M$ ,  $\succsim_{[u]}$  also satisfies Monotonicity ([Axiom 4](#)).

*Proof of Claim.* Let  $a := \{\alpha \in \mathcal{P}(K) : \|\alpha - p_k^*\|_2 \leq \varepsilon\}$  for some sufficiently small  $\varepsilon > 0$ . Let  $k' \in K$  be such that  $u_1(\lambda p_k^* + (1 - \lambda)k') > \max_{\alpha \in a} u_1(\alpha)$  for all  $\lambda \in (0, 1)$  such that  $\lambda p_k^* + (1 - \lambda)k' \notin a$ . Define  $b_\lambda := \{\lambda p_k^* + (1 - \lambda)k'\} \cup a$  and  $\lambda^* := \min\{\lambda > 0 : \lambda p_k^* + (1 - \lambda)k' \in a\}$ , so that  $b_{\lambda^*} = a$ . By construction,  $a \cup b_\lambda \succ_K a$ , for all  $\lambda > \lambda^*$ .

Suppose, contrary to the claim, that  $x \succ_{[u_i]} x \cup y$  for  $i = 1, \dots, \ell$  but  $x \cup y \succ_{[u_i]} x$  for  $i = \ell + 1, \dots, n$  for some  $\ell \in \{1, \dots, n\}$ . By appropriately perturbing  $y$  if necessary, we may take  $y$  such that  $x \cup y \succ x$  (because  $\succsim$  is continuous). Observe that  $(a, \{x, x \cup y\}) \succ (a, \{x \cup y\})$  because  $W((a, \{x, x \cup y\})) - W((a, \{x \cup y\})) = \sum_{i=1}^{\ell} [v(x, [u_i]) - v(x \cup y, [u_i])] \mu([u_i]) > 0$ , where we have used the fact that  $v(x, [u_i]) > v(x \cup y, [u_i])$  for  $i = 1, \dots, \ell$ .

Since  $a \cup b_\lambda \succ_K a$  and  $x \cup y \succ x$ , [Persistent Preference for Flexibility](#) ([Axiom 10](#)) implies that  $(a \cup b_\lambda, \{x \cup y\}) \succ (a, \{x\}) \cup (b_\lambda, \{x \cup y\})$ . In the limit as  $\lambda \rightarrow \lambda^*$  so

that  $b_\lambda \rightarrow a$ , we must have  $(a, \{x \cup y\}) \succsim (a, \{x, x \cup y\})$ , because  $\succsim$  is continuous, which contradicts our earlier observation that  $(a, \{x, x \cup y\}) \succ (a, \{x \cup y\})$ , thereby proving the claim.  $\blacktriangle$

Theorem 3 then implies that because  $v(\cdot, [u])$  represents  $\succsim_{[u]}$ , and because  $v(\cdot, [u])$  is continuous, linear, and monotone with respect to set inclusion,  $v(\cdot, [u])$  can be written as

$$v(x, [u]) = \delta([u]) \int_{\mathfrak{U}_{K \times Z}} \max_{p \in x} u(p) \, d\mu_{[u]}(u)$$

where  $\mu_{[u]}$  is a probability charge on  $\mathfrak{U}_{K \times Z}$ , and  $\delta([u])$  is the scaling factor that allows us to take  $\mu_{[u]}$  to be a probability charge.

Let  $W'(x) := W(\{(k, x)\})$  for some fixed  $k \in K$ . By Stationarity (Axiom 6),  $W'(x) \geq W'(y)$  if, and only if,  $W(x) \geq W(y)$ . Indifference to Timing (Axiom 7) implies that  $W'$  is an affine transformation of  $W$ . Therefore, we have

$$\begin{aligned} W(x) &= \int_{\mathfrak{U}_M} \max_{p \in x} [u(p_k) + v(p_z, [u])] \, d\mu_0(u) \\ &\propto W'(x) = W(\{(k, x)\}) \\ &= \sum_{[u] \subset \mathfrak{U}_M} v(x, [u]) \mu_0([u]) + \text{constant} \\ &= \sum_{[u] \subset \mathfrak{U}_M} \left[ \delta([u]) \int_{\mathfrak{U}_{K \times Z}} \max_{p \in x} u(p) \, d\mu_{[u]}(u) \right] \mu_0([u]) + \text{constant} \end{aligned}$$

Define the charge  $\sigma^*$  on  $\mathfrak{U}_{K \times Z}$  as  $\sigma^*(du) := \sum_{[u] \in \mathfrak{U}_M} \delta([u]) \mu_0([u]) \mu_{[u]}(du)$ . Then, the last line in the display above can be written as

$$W(x) \propto \int_{\mathfrak{U}_{K \times Z}} \max_{p \in x} u(p) \, d\sigma^*(u)$$

where  $\sigma^*$  on  $\mathfrak{U}_{K \times Z}$  is not necessarily a probability charge.

Following the arguments in Proposition 28KS, and using the fact that  $[u]$  is a sufficient statistic for  $v \in \mathfrak{U}_Z$ , we may take  $\mu_0$  to be defined on  $\mathfrak{U}_{K \times Z}$ . By virtue of Proposition 23KS, which says that the state space and measure are jointly identified, it must be that  $\mu_0 \propto \sigma^*$ .

This implies the carrier of  $\sigma^*$  coincides with the carrier of  $\mu_0$  and is finite. But  $\sigma^*$  is the positive linear combination of charges  $\mu_{[u]}$  on  $\mathfrak{U}_{K \times Z}$ , and therefore,

the carrier of each  $\mu_{[u]}$  (where  $u \in \mathcal{U}_M$ ) must be a subset of the carrier of  $\sigma^*$ , and hence of the carrier of  $\mu_0$ , which is a subset of  $\mathcal{U}_M$ .

This allows us to write

$$v(x, [u]) = \delta([u]) \int_{\mathcal{U}_M} \max_{p \in x} [u'(p_k) + v(p_z, [u'])] d\mu_{[u]}(u')$$

for each  $u \in \mathcal{U}_M$ . Define the Markov kernel  $M$  on  $\mathcal{U}_M$  as follows:  $M(u, du') := \mu_{[u]}(du')$  for all  $u \in \mathcal{U}_M$ . It is clear that  $M$  is ranking contingent. Also define  $\delta(u) := \delta([u])$ , and  $V(x, u) := v(x, [u])/\delta(u)$  to find the recursive value function [B.5]. Finally, plugging into equation [B.4], and defining  $V(x, \mu_0) := W(x)$ , we see that  $V(x, \mu_0)$  as in [B.6] represents  $\succsim$ , as desired.  $\square$

We now show that the Markov kernel  $M$  is uniformly ranking persistent.

**Proposition B.13.** Let  $\succsim$  have a recursive representation as in [B.6], and suppose  $\succsim$  satisfies **Persistent Preference for Flexibility (Axiom 10)**. Then,  $M(u, [u']) > 0$  for all  $u, u' \in \mathcal{U}_M$ .

*Proof.* Suppose  $M(u', [u]) = 0$  for some  $u, u' \in \mathcal{U}_M$ . Construct menus  $x, y$  such that

$$\begin{aligned} & \int_{[u]} \max_{p \in y} [\hat{u}(p_k) + \delta(\hat{u})V(p_z, \hat{u})] d\mu_0(\hat{u}) \\ & > \int_{[u]} \max_{p \in x} [\hat{u}(p_k) + \delta(\hat{u})V(p_z, \hat{u})] d\mu_0(\hat{u}) \end{aligned}$$

but

$$\begin{aligned} & \int_{\mathcal{U}_M \setminus [u]} \max_{p \in x} [\hat{u}(p_k) + \delta(\hat{u})V(p_z, \mu_{\hat{u}})] d\mu_0(\hat{u}) \\ & > \int_{\mathcal{U}_M \setminus [u]} \max_{p \in y} [\hat{u}(p_k) + \delta(\hat{u})V(p_z, \mu_{\hat{u}})] d\mu_0(\hat{u}) \end{aligned}$$

This immediately implies  $x \cup y \succ x$ .

Similarly, construct consumption menus  $a$  and  $b$ , such that  $\max_{\alpha \in a} u'(\alpha) > \max_{\alpha \in b} u'(\alpha)$  and  $\max_{\alpha \in b} \hat{u}(\alpha) \geq \max_{\alpha \in a} \hat{u}(\alpha)$  for all  $\hat{u} \in \mathcal{U}_M \setminus [u']$ . Then, the best element from  $(a \cup b, \{x \cup y\})$  is in  $(a, \{x \cup y\})$  only in states in  $[u']$ . At the same time, contingent on being in a consumption state in  $[u']$ , the best choice of a continuation menu is surely  $\{x\}$ , since  $M(u', [u]) = 0$  by assumption. Hence,  $(a, \{x\}) \overset{\circ}{\sim} (b, \{x \cup y\})$ , which contradicts **Persistent Preference for Flexibility (Axiom 10)**, which requires  $(a, \{x \cup y\}) \overset{\circ}{\succ} (b, \{x \cup y\})$ .  $\square$

### B.3. Representation with Constant Discount Factor: Existence and Uniqueness

Propositions B.11 and B.13 establish that  $\succsim$  has a recursive representation as in [B.5] and [B.6], with a ranking persistent Markov process. Our goal is to show that there exists a unique equivalent representation with a *constant* discount factor. We now describe, in brief, how the approach parallels the construction of a representation with a constant discount factor that leads to the DPF representation in KS.

In the DPF representation, the states are  $S = \{1, \dots, n\}$  and uncertain utilities are described by measures  $\mu_s$  that depend on the state  $s \in S$ , while the Markov process on states  $S$  is given by the transition probabilities  $\Pi(s', s)$ .

In the evolving tastes representation, the uncertain utilities are described by the ranking contingent Markov kernel  $M(u, \cdot)$ , wherein  $M(u, \cdot) = M(\lambda u, \cdot)$ . Moreover, the probability of the ranking  $[u_i]$  is  $M(u, [u_i])$ . Consider the following correspondence:

DPF		Evolving Tastes
$S = \{1, \dots, n\}$	States for Markov Process	$\{[u_i] : i = 1, \dots, n\}$
$\Pi(s, s')$	Probabilities of States/Rankings	$M(u, [u_i])$
$\mu_s(D), D \subset \mathcal{U}$	Probabilities of utilities given states	$M(u, E)/M(u, [u_i]), E \subset [u_i]$
$\pi_0$ (of $\Pi$ )	Invariant Measures	$\mu_0$ (of $M$ )

In the DPF representation, at each instant in time, DM's uncertainty about utilities in the next period can be decomposed into uncertainty about the state  $s \in S$  and the conditional probabilities over utilities in  $\mathcal{U}$  given by  $\mu_s$ . Formally, this determines a probability measure over  $S \times \mathcal{U}$ .

In the evolving tastes representation, uncertainty about utilities in  $\mathcal{U}_M$  can be decomposed into uncertainty about the ranking  $[u_i]$ , given by  $M(u, [u_i])$ , and uncertainty about the intensity of the utility, conditional on the ranking, given by  $M(u, E)/M(u, [u_i])$  where  $E \subset [u_i]$  is measurable. Thus, the formal structure of the two representations is essentially the same.

In order to obtain a representation with a constant discount factor, one can apply the same transformations that we used in the case of the DPF representation as in Proposition 30KS. The Perron Theorem implies that such a representation

exists and is unique. In particular, the distribution of intensities given the ranking  $[u_i]$  is uniquely identified. To identify the stationary distribution  $\mu_0$  of  $M$ , it suffices then to calculate the stationary distribution of the induced Markov process on rankings.

Once we have a representation with a constant discount factor, all that remains is to show that the discount factor is less than one. This is done in a manner parallel to Proposition 32KS for the DPF representation.

Define the measure  $\mu_u(\cdot) := M(u, \cdot)$ . Let  $\xi \in \mathbb{R}_{++}^n$  be strictly positive, divide both sides of equation [B.6] by  $\langle \mu_0, \xi \rangle := \sum_{i=1}^n \mu_0([u_i])\xi([u_i])$ , and rewrite equation [B.6] as

$$\hat{V}(x, \mu_0) := \frac{V(x, \mu_0)}{\langle \mu_0, \xi \rangle} = \sum_{i=1}^n \int_{[u_i]} \max_{p \in x} \left[ \frac{u(p_k)}{\xi([u_i])} + \frac{\delta([u_i])}{\xi([u_i])} V(p_z, u) \right] d\mu_0(u)$$

where  $\mu_0([u_j]) := \mu_0([u_j])\xi([u_j]) / \langle \mu_0, \xi \rangle$  for  $j = 1, \dots, n$ , and  $\mu_0([\lambda u_j, \lambda' u_j]) := \mu_0\left(\left[\frac{\lambda u_j}{\xi([u_j])}, \frac{\lambda' u_j}{\xi([u_j])}\right]\right)$  so that  $\mu_0$  is clearly a probability measure. Notice that  $\hat{V}(x, \mu_0)$  also represents  $\succsim$ . Now, for each  $u \in \mathcal{U}_M$ , consider the following renormalizations:

- Setting  $\hat{u}(\cdot) := u(\cdot)/\xi([u])$ , we obtain  $\mu_{\hat{u}}(D) := \mu_u(D/\xi([u]))$  for all measurable  $D \subset [u']$  and  $u' \in \mathcal{U}_M$
- $\mu_{\hat{u}}([u_j]) := \mu_u([u_j])\xi([u_j]) / \langle \mu_u, \xi \rangle$
- $\hat{\delta}([u]) := \delta([u]) \langle \mu_u, \xi \rangle / \xi([u])$
- $\hat{V}(\cdot, \mu_{\hat{u}}) := V(\cdot, \mu_u) / \langle \mu_u, \xi \rangle$

Then, equation [B.5] can be rewritten as

$$[\text{B.7}] \quad \hat{V}(x, \hat{u}') = \int_{\mathcal{U}} \max_{p \in x} [\hat{u}(p_k) + \hat{\delta}(\hat{u}) \hat{V}(p_z, \hat{u})] d\mu_{\hat{u}'}(\hat{u})$$

where  $\hat{V}(x, \mu_0)$  represents  $\succsim$ . Before proceeding, it is worthwhile to discuss why these transformations are the only transformation that we need consider. By a variation of Proposition 25KS, we see that any transformation of a measure  $\mu_{\hat{u}}$  must be a scaling on the set  $[u]$  for each  $[u] \subset \mathcal{U}_M$ . Therefore, we must have  $\mu_{\hat{u}}([\lambda u_j, \lambda' u_j]) := \mu\left(\left[\frac{\lambda u_j}{\xi([u_j])}, \frac{\lambda' u_j}{\xi([u_j])}\right]\right)$  for all  $\lambda' > \lambda > 0$ ,  $j = 1, \dots, n$ . In particular, the variation of Proposition 25KS says that  $\mu_u([u_i])$  and  $\delta(u_i)V(\cdot, u_i)$  must be identified up to the same scaling, so that we must also transform  $\delta(u_i)V(\cdot, u_i)$  so that  $\hat{\delta}(\hat{u}_i)\hat{V}(\cdot, \hat{u}_i) := \delta(u_i)V(\cdot, u_i)/\xi([u_i])$ . But such a transformation means that  $\delta(u)$  and  $V(\cdot, u)$  must be transformed as above. We are now in a position to state and prove the following:

**Proposition B.14.** In the transformation of [B.5] to [B.7] above, there exists  $\xi \in \mathbb{R}_{++}^n$  such that  $\hat{\delta}$  is independent of  $[u]$ . Moreover,  $\xi$  is unique up to scaling, so that  $\hat{\delta}$  is unique, and therefore the representation in equation [B.7] is unique up to a common scaling of the measures  $\mu_0 \cup (\mu_{\hat{u}})$ .

*Proof.* A representation with a constant discount factor will obtain immediately if we can establish that there exists a vector  $\xi \gg \mathbf{0}$  and a number  $\hat{\delta} > 0$  such that  $\hat{\delta}\xi([u_j]) = \delta([u_j]) \langle \mu_u, \xi \rangle$  for all  $j = 1, \dots, n$ .

Let us now consider this problem from a slightly different point of view. Define a matrix  $A$  where the element  $a_{ij} := \mu_{u_i}([u_j])$ . The Markov kernel  $M$  is ranking persistent, which implies that  $A$  is a positive stochastic matrix. Define the diagonal matrix  $\Delta$  as follows:

$$\Delta := \begin{bmatrix} \delta([u_1]) & 0 & \dots & 0 \\ 0 & \delta([u_2]) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \delta([u_n]) \end{bmatrix}$$

where, as before,  $\mathcal{U}_M = \bigcup_{i=1}^n [u_i]$ .

In matrix notation, our problem amounts to finding a  $\xi \gg \mathbf{0}$  and  $\hat{\delta} > 0$  such that  $\hat{\delta}\xi = \Delta A\xi$ . But this amounts to showing that (i)  $\xi$  is an eigenvector of the matrix  $\Delta A$ , and (ii)  $\hat{\delta}$  is the corresponding eigenvalue. The matrix  $\Delta A$  is positive because  $A$  is positive. Therefore, by the Perron theorem, such a  $\xi$  and  $\hat{\delta}$  exist, and are unique, which proves the proposition.  $\square$

The Perron Theorem is standard and can be found, for instance, as Theorem 1 in chapter 16 of Lax (2007).

**Theorem 3 (Perron).** *Every positive matrix  $A$  has a **dominant eigenvalue** denoted by  $\hat{\delta}$  which has the following properties:*

- (a)  $\hat{\delta} > 0$  and the associated eigenvector  $\xi \gg \mathbf{0}$ .
- (b)  $\hat{\delta}$  is a simple eigenvalue, and hence has algebraic and geometric multiplicity one.
- (c)  $A$  has no other eigenvector with nonnegative entries.

We now show that  $\hat{\mu}_0$  is the stationary distribution of the Markov process  $M$ . To ease notational burden, we drop the ‘hats’ from the recursive representation in [B.7], and also use the Markov kernel  $M$  instead of the measures  $\mu_u$ .

**Proposition B.15.** In the representation [B.7],  $\mu_0$  is the unique measure that satisfies  $\mu_0(du) = \int_{\mathcal{U}} \mu_{u'}(du) \mu_0(du')$ , so that  $\mu_0$  is the unique invariant distribution of the Markov process  $M$ .

*Proof.* Fix  $k \in K$ , and define  $W(x, \mu_0) := V(\{k, x\}, \mu_0)$ . By Stationarity (Axiom 6),  $W(\cdot, \mu_0)$  represents  $\succsim$ , and by Indifference to Timing (Axiom 7),

$$W(x, \mu_0) = V(\{k, x\}, \mu_0) = \int_{\mathcal{U}} [u(k) + \delta V(x, u)] d\mu_0(u)$$

But  $V(x, u) = \int_{\mathcal{U}} \max_{p \in x} [u'(p_k) + \delta V(p_z, u')] M(u, du')$ , so that

$$\begin{aligned} W(x, \mu_0) &= \int_{\mathcal{U}} u(k) d\mu_0(u) + \\ &\quad \delta \int_{\mathcal{U}} \left[ \int_{\mathcal{U}} \max_{p \in x} [u'(p_k) + \delta V(p_z, u')] M(u, du') \right] d\mu_0(u) \\ &= \int_{\mathcal{U}} u(k) d\mu_0(u) + \\ &\quad \delta \int_{\mathcal{U}} \left[ \max_{p \in x} [u'(p_k) + \delta V(p_z, u')] \int_{\mathcal{U}} M(u, du') \right] \mu_0(du) \end{aligned}$$

Since both  $W(x, \mu_0)$  and  $V(x, \mu_0)$  represent  $\succsim$ , our uniqueness statement says that they must induce the same measure over  $\mathcal{U}$ . This implies that for each  $u \in \mathcal{U}$ ,

$$\mu_0(du') = \int_{\mathcal{U}} M(u, du') \mu_0(du)$$

that is,  $\mu_0$  is the stationary distribution of the Markov process  $M$ . By the Perron Theorem,  $\mu_0([u_i])$  must be unique, which implies that the measure on  $[u_i]$  given by  $\int_D \mu_0(du)$  for measurable  $D \subset [u_i]$  is also unique. Hence,  $\mu_0$  is the unique stationary, and hence ergodic, measure of the Markov process.  $\square$

An immediate corollary is that  $\mu_0$  is nice and  $(\mathcal{U}_M, M)$  is a nice ranking persistent Markov process.

**Corollary B.16.** In the representation [B.7], if  $\mu_0$  is the invariant measure of the Markov process  $(\mathcal{U}_M, M)$ , then  $\mu_0$  is nice and  $(\mathcal{U}_M, M)$  is a nice ranking contingent Markov process.

*Proof.* For the menu  $(k, x) \in Z$ ,  $V((k, x)) = \mu_0 u_k + \delta V(x, \mu_0)$ , where we have used the fact that  $\mu_0$  is the invariant measure of  $M$ . As  $V(x)$  and  $V((k, x))$  are both finite (by definition of  $V$ ), it follows that  $\mu_0 u_k \in \mathbb{R}$  for all  $k$ , which proves that  $\mu_0$  is nice. A similar proof establishes that  $M(u, \cdot)$  is nice for all  $u$ .  $\square$



We shall now establish that the discount factor  $\delta$  is less than 1.

**Proposition B.17.** If  $\succsim$  has a recursive representation of the form in (B.7) with constant  $\delta$ , then  $\delta \in (0, 1)$ .

*Proof.* We have already shown that  $\delta > 0$ , therefore it suffices to show that  $\delta < 1$ .

Let  $a$  be the  $\varepsilon$  neighbourhood of  $p_k^*$  for some sufficiently small  $\varepsilon > 0$ . Then,  $\int_{\mathcal{U}} \max_{\alpha \in a} u(\alpha) d\mu_0(u) > 0$  because  $\mu_0$  is nice. Construct the menu  $y^* \simeq (a, \{y^*\})$ . Let  $\eta := \int_{\mathcal{U}} \max_{\alpha \in a} u(\alpha) d\mu_0(u) > 0$ . Then,  $V(y^*, \mu_0) = \eta + \delta V(y^*, \mu_0)$ , which implies that  $V(y^*, \mu_0) = \eta \sum_{\tau \geq 0} \delta^\tau$ . The finiteness of  $V(y^*, \mu_0)$  implies  $\delta < 1$ .  $\square$

#### B.4. Proof of Proposition 2.7

Proposition 2.7 states that each evolving tastes representation  $((\mathcal{U}_M, M), \delta)$  induces a unique value function  $V \in C(Z \times \mathcal{U}_M)$  that satisfies equation (4.1) in the text.

Let  $W \in C(Z \times \mathcal{U}_M)$ , and consider the function  $\Phi W(x, u)$ , given by

$$\Phi W(x, u) := \int_{\mathcal{U}} \max_{p \in x} [u'(p_k) + \delta W(p_z, u')] M(u, du')$$

for all  $u \in \mathcal{U}_M$  and  $x \in Z$ . It is easy to see that  $\Phi$  is monotone, ie,  $W \leq W'$  implies  $\Phi W \leq \Phi W'$ , and satisfies discounting, ie,  $\Phi(W + \rho) \leq \Phi W + \delta \rho$  when  $\rho \geq 0$ . Moreover,  $\Phi W(x, u) = \Phi W(x, \lambda u)$  for all  $\lambda > 0$ . If we assume that  $\Phi W \in C(Z \times \mathcal{U}_M)$  for all  $W \in C(Z \times \mathcal{U}_M)$ , it follows that  $\Phi$  is a contraction mapping (with modulus  $\delta$ ), and has a unique fixed point which establishes the proposition.

For each  $x \in Z$ ,  $u \in \mathcal{U}_M$ , and  $W \in C(Z \times \mathcal{U}_M)$ , define

$$\varphi(x, u) = \max_{p \in x} [u(p_k) + \delta W(p_x, u)]$$

Then,

$$\begin{aligned} |\varphi(x, u)| &\leq \max_{p \in x} |u(p_k) + \delta W(p_x, u)| \\ &\leq \|u\|_2 \max_{p \in x} \left| \frac{u(p_k)}{\|u\|_2} \right| + \max_{p \in x} \delta |W(p_x, u)| \\ &\leq \|u\|_2 M_1 + M_{2,u} \end{aligned}$$

where  $M_1 := \max_{x \in Z} \max_{p \in x} \left| \frac{u(p_k)}{\|u\|_2} \right|$ ,  $M_{2,u} > 0$ , and the bounds follow from the definition of  $u \in \mathcal{U}$ , the compactness of  $Z$ , the continuity of  $W$ , the fact that  $M$  is

a nice ranking persistent Markov process, and because  $W(\cdot, u) = W(\cdot, \lambda u)$  for all  $\lambda > 0$ .

As  $W(\cdot, u)$  is continuous, the function  $u(p_k) + \delta W(p_z, u) \in C(K \times Z)$  is a continuous, linear functional on  $\mathcal{P}(K \times Z)$ , when the latter is endowed with the topology of weak convergence (which is metrizable). Therefore, by the Maximum Theorem, for each  $u \in \mathcal{U}_M$ ,  $\varphi(x, u)$  is continuous in  $x$ .

We will now show that if  $(x_n) \in Z^\infty$  is a sequence that converges to  $x \in Z$ , then  $\Phi W(x_n, \cdot) \rightarrow \Phi W(x, \cdot)$  whenever  $W \in C(Z \times \mathcal{U}_M)$ . (Since  $\mathcal{U}_M$  is a finitely generated cone, and because  $W(\cdot, u) = W(\cdot, \lambda u)$  for all  $\lambda > 0$ , any sequence of  $u$ 's in  $\mathcal{U}_M$  can be replaced by a sequence on the equivalence classes  $\{[u_i] : i = 1, \dots, n\}$ , which is finite. This sequence must eventually be constant, so we may assume, without loss of generality, that this constant is  $u \in \mathcal{U}_M$ .)

Consider any sequence  $(x_n)$  that converges to  $x$ . By the bounds established above,  $|\varphi(x_n, u)| \leq \|u\|_2 M_1 + M_{2,u}$ , and  $\|u\|_2 M_1 + M_{2,u}$  is  $M(u', \cdot)$ -integrable because  $M(u', \cdot)$  is nice. Moreover,

$$\begin{aligned} \lim_{n \rightarrow \infty} \Phi W(x_n, u) &= \lim_{n \rightarrow \infty} \int_{\mathcal{U}} \varphi(x_n, u') M(u, du') \\ &= \int_{\mathcal{U}} \lim_{n \rightarrow \infty} \varphi(x_n, u') M(u, du') \\ &= \int_{\mathcal{U}} \varphi(x, u') M(u, du') \\ &= \Phi W(x, u) \end{aligned}$$

As  $x$  and  $(x_n)$  are arbitrary, we conclude that  $\Phi W \in C(Z \times \mathcal{U}_M)$  whenever  $W \in C(Z \times \mathcal{U}_M)$ . The equalities above rely on the Dominated Convergence Theorem to interchange the order of limits and integration, and the continuity of  $\varphi(\cdot, u')$  for each  $u'$  to establish the pointwise limit. This completes the proof.

### B.5. Putting it all together

We are now in a position to complete the proof of Theorem 1.

*Proof of Theorem 1.* It is straightforward to show that the representation satisfies all the axioms. With regards to Choice Contingent CSR (Axiom 9), we can adapt the ideas from Lemma B.3 to construct consumption menus  $a$  and  $b$  such that  $a$  outperforms  $b$  in only one taste.

Consider then, a preference  $\succsim$  that satisfies Non-triviality (Axiom 1), Order (Axiom 2), Independence (Axiom 3) and Monotonicity (Axiom 4). By Theorem 3KS,  $\succsim$  has a finitely additive, EU representation. As  $\succsim$  also satisfies Separability (Axiom 5), it follows from Lemma 19KS and Proposition 20KS that any such finitely additive EU representation also has a finitely additive separable representation by the functional  $V(x) = \int_{\mathfrak{U}_K \times [0,1] \times \mathfrak{U}_Z} \max_{p \in x} [\lambda r(p_k) + (1 - \lambda)v(p_z)] d\mu(r, \lambda, v)$ .

By rewriting  $\lambda r(p_k) + (1 - \lambda)v(p_z)$  as  $u(p_k) + v(p_z)$ , where  $u = \frac{\lambda}{1-\lambda}r$ , and by making an appropriate transformation to the charge  $\mu$  (using the change of state space Lemma 16KS), we may regard the state space as  $\mathcal{U} \times \mathfrak{U}_Z$ . Finiteness (Axiom 8) says that  $\mathcal{U}^*$ , the carrier of the marginal charge on  $\mathcal{U}$ , has finitely many non-collinear components. Proposition B.4 says that if the preference also satisfies Choice Contingent CSR (Axiom 9), then for each  $u \in \mathcal{U}^*$ , the induced marginal charge  $\mu(\cdot|u)$  on  $\mathfrak{U}_Z$  has singleton support. In other words, for each  $u \in \mathcal{U}^*$ , there corresponds a unique continuation value function  $v \in \mathfrak{U}_Z$ , which is denoted by  $v(\cdot, [u])$ , which allows us to take  $\mu$  as a measure instead of as a charge.

Proposition B.10 says that because  $\succsim$  satisfies Indifference to Timing (Axiom 7),  $v(\cdot, [u])$  is also linear. Persistent Preference for Flexibility (Axiom 10) implies that  $v(\cdot, [u])$  is monotone. All that remains is to show that there exists a recursive representation with a constant discount factor, that the discount factor is less than 1, and that  $\mu_0$  is the unique stationary distribution of the Markov process. These are established in Propositions B.15–B.17.  $\square$

### C. Proofs from Section 3

*Proof of Proposition 3.2.* The ‘if’ part is straightforward. To see the ‘only if’ part, fix  $a \in \mathcal{F}(\mathcal{P}(K))$  and let  $z_a \in Z_S$  be the separable menu that provides  $a$  in each period. Because  $\succsim$  and  $\succsim^\dagger$  agree on separable problems, it follows that  $V$  and  $V^\dagger$  are equivalent up to scaling on  $Z_S$ . That is,

$$\begin{aligned} V(z_a, \mu_0) &= \frac{1}{1 - \delta} \int_{\mathfrak{U}_M} \max_{p \in a} u(p_k) d\mu_0(u) \\ &= \zeta \cdot \frac{1}{1 - \delta^\dagger} \int_{\mathfrak{U}_M} \max_{p \in a} u(p_k) d\mu_0^\dagger(u) \\ &= \zeta V^\dagger(z_a, \mu_0^\dagger) \end{aligned}$$

where we have used the fact that  $\mu_0$  and  $\mu_0^\dagger$  are ergodic distributions. Up to a scaling, this is the representation in Dekel, Lipman, and Rustichini (2001), which is also defined on  $\mathcal{F}(\mathcal{P}(K))$ . Using their identification result, it follows that

$$[\mathbf{C.1}] \quad \int_{\{\lambda u: \lambda \geq 0\}} \lambda \, d\mu_0(\lambda u) = \zeta \int_{\{\lambda u: \lambda \geq 0\}} \lambda \, d\mu_0^\dagger(\lambda u)$$

for each  $u \in \mathcal{U}_M$ . Recall that  $[u] = \{u' \in \mathcal{U}_M : u' \propto u\}$  and that there exists finitely many  $u_1, \dots, u_n$  such that  $\mathcal{U}_M = \bigcup_{i=1}^n [u_i]$ . Summing over all these equivalence classes and using [C.1], which implies  $\mathcal{U}_M = \mathcal{U}_M^\dagger$ , we find that

$$\begin{aligned} \mu_0 u &:= \int_{\mathcal{U}_M} u \, d\mu_0(u) = \sum_{i=1}^n \int_{[u_i]} \lambda u \, d\mu_0(\lambda u) \\ &= \zeta \sum_{i=1}^n \int_{[u_i]} \lambda u \, d\mu_0^\dagger(\lambda u) \\ &= \zeta \int_{\mathcal{U}_M} u \, d\mu_0^\dagger(u) =: \zeta \mu_0^\dagger u \end{aligned}$$

Now, pick any prize  $k$  for which  $\mu_0 u(k) \neq 0$ . Then, for  $z_{\{k\}} \in Z_S$ , which gives the prize  $k$  in each period, we have

$$\begin{aligned} V(z_{\{k\}}, \mu_0) &= \frac{1}{1-\delta} \int_{\mathcal{U}_M} u(k) \, d\mu_0(u) \\ &= \zeta \cdot \frac{1}{1-\delta^\dagger} \int_{\mathcal{U}_M} u(k) \, d\mu_0^\dagger(u) \\ &= \zeta V^\dagger(z_{\{k\}}, \mu_0^\dagger) \end{aligned}$$

But we have already established that  $\mu_0 u = \zeta \mu_0^\dagger(u)$ , which implies that  $\delta = \delta^\dagger$ , as claimed.  $\square$

*Proof of Theorem 2.* Observe first, that with  $\zeta = 1$  in Proposition 3.2,  $V(z, \mu_0) = V^\dagger(z, \mu_0^\dagger)$  for all  $z \in Z_S$ . Second, observe that for  $z \in Z_{\text{iid}}$  we have

$$V(z, \mu_0) = \int_{\mathcal{U}_M} \max_{\mathbf{p}_k \in Z} [u(p_k) + \delta V(\mathbf{p}_k, u)] \, d\mu_0(u)$$

Notice that  $u(p_k) + \delta V(\mathbf{p}_k, u)$  can be expanded as

$$u(p_k) + \delta \int u^{(1)}(p_k) M(u, du^{(1)}) + \delta^2 \int u^{(2)}(p_k) M^2(u, du^{(2)}) + \dots$$

where  $M^n(u, \cdot)$  is the  $n$ -step probability measure over  $\mathcal{U}_M$  induced by  $M$ . But notice now that  $U(u, p_k) = u(k) + \delta \int U(u^{(1)}, p_k) M(u, du^{(1)})$  can be expanded as

$$u(p_k) + \delta \int u^{(1)}(p_k) M(u, du^{(1)}) + \delta^2 \int u^{(2)}(p_k) M^2(u, du^{(2)}) + \dots$$

where  $U(u, p_k)$  is defined in [3.1]. Therefore,

$$\begin{aligned} V(z, \mu_0) &= \int_{\mathcal{U}_M} \max_{p_k \in z} U(u, p_k) d\mu_0(u) \\ &= \int_{\mathcal{U}_M} \max_{p_k \in z} \langle \tilde{U}(u), p_k \rangle d\mu_0(u) \\ &= \int_{\tilde{U}(\mathcal{U}_M)} \max_{p_k \in z} \langle \tilde{u}, p_k \rangle d\tilde{\mu}(\tilde{u}) \\ &= \int_{\tilde{U}(\mathcal{U}_M)} \max_{p_k \in z} \tilde{u}(p_k) d\tilde{\mu}(\tilde{u}) \end{aligned}$$

where  $\langle \cdot, \cdot \rangle$  is a standard inner product and  $\tilde{U}$  and  $\tilde{\mu}$  are defined in [3.2] and [3.3] respectively.

But  $\max_{p_k \in z} \tilde{u}(p_k)$  is convex in  $\tilde{u}$ . Therefore, Theorem 7.2.1.7 in Torgersen (1991) (which generalizes the theorem of Blackwell (1953) to measures with unbounded support) implies that  $V(z, \mu_0) \geq V^\dagger(z, \mu_0^\dagger)$  if, and only if,  $\tilde{\mu}$  is a dilation of  $\tilde{\mu}^\dagger$ , as claimed.  $\square$

## References

- Aliprantis, Charalambos, and Kim C. Border. 1999. *Infinite Dimensional Analysis: A Hitchhiker's Guide, 2/e*. Springer, New York. (Cit. on p. 25).
- Bekaert, Geert, Eric Engstrom, and Steven R. Grenadier. 2010. "Stock and Bond Returns with Moody Investors". *Journal of Empirical Finance* 17 (5): 867–894. (Cit. on p. 1).
- Blackwell, David. 1953. "Equivalent Comparison of Experiments". *Annals of Mathematical Statistics* 24:265–272. (Cit. on p. 37).
- Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini. 2001. "Representing Preferences with a Unique Subjective State Space". *Econometrica* 69 (4): 891–934. (Cit. on pp. 2, 17, 36).

- . 2009. “Temptation Driven Preferences”. *Review of Economic Studies* 76 (3): 937–971. (Cit. on p. 7).
- Dillenberger, David, R. Vijay Krishna, and Philipp Sadowski. 2016. *Subjective Dynamic Information Constraints*. Tech. rep. Duke University. (Cit. on p. 17).
- Ergin, Haluk, and Todd Sarver. 2010. “A Unique Costly Contemplation Representation”. *Econometrica* 78 (4): 1285–1339. (Cit. on p. 17).
- Gul, Faruk, and Wolfgang Pesendorfer. 2004. “Self-Control and the Theory of Consumption”. *Econometrica* 72 (1): 119–158. (Cit. on p. 2).
- . 2005. “The Revealed Preference Theory of Changing Tastes”. *Review of Economic Studies* 72:429–448. (Cit. on p. 18).
- Kreps, David M. 1979. “A Representation Theorem for ‘Preference for Flexibility’”. *Econometrica* 47 (3): 565–577. (Cit. on pp. 2, 5).
- Kreps, David M., and Evan L Porteus. 1978. “Temporal Resolution of Uncertainty and Dynamic Choice Theory”. *Econometrica* 46 (1): 185–200. (Cit. on p. 4).
- Krishna, R. Vijay, and Philipp Sadowski. 2014. “Dynamic Preference for Flexibility”. *Econometrica* 82 (2): 655–703. (Cit. on p. 3).
- Lax, Peter D. 2007. *Linear Algebra and its Applications*. Second. New Jersey: Wiley. (Cit. on p. 31).
- O’Donoghue, Ted, and Matthew Rabin. 2000. “The Economics of Immediate Gratification”. *Journal of Behavioral Decision Making* 13 (233–250). (Cit. on p. 18).
- Piermont, Evan, Norio Takeoka, and Roe Teper. 2015. *Learning the Krepsian State: Exploration Through Consumption*. Tech. rep. University of Pittsburgh. (Cit. on p. 17).
- Riella, Gil. 2013. “Preference for Flexibility and Dynamic Consistency”. *Journal of Economic Theory* 148 (6): 2467–2482. (Cit. on pp. 7, 20).
- Strotz, Robert H. 1955. “Myopia and Inconsistency in Dynamic Utility Maximization”. *Review of Economic Studies* 23 (3): 165–180. (Cit. on p. 18).
- Torgersen, Erik. 1991. *Comparison of Statistical Experiments*. Cambridge University Press. (Cit. on p. 37).