For Online Publication
Supplementary Appendix to

Subjective Information Choice Processes

DAVID DILLENBERGER∗ R. VIJAY KRISHNA† PHILIPP SADOWSKI‡

JANUARY 2017

All references to definitions and results in this Supplement refer to Dillenberger, Krishna, and Sadowski (2017, henceforth DKS) unless otherwise specified. The supplement is organized as follows. Section 1 below provides details of the Information Cost representation in Section 6 as well as a proof of Theorem 4 in DKS. Section 2 reviews some relevant notions from convex analysis, which are used in the sequel. Section 3 establishes the static representation in [**D.1**] of DKS. Section 4 extends the existence of the Recursive Anscombe-Aumann representation, which is established in Krishna and Sadowski (2014) for finite prize spaces, to our domain with a compact set of prizes, as discussed in Appendix D.1 of DKS. Finally, Section 5 provides a metric on the space of partitions as described in Appendix A.3 of DKS.

1. Information Cost Representation

1.1. Domain for Unbounded Utilities

As before, S is a finite set of *objective* states. Let C be a sigma-compact metric space, representing consumption. Thus, there exist compact metric spaces $C^{(n)}$ for $n = 0, 1, \dots$ such that $C := \bigcup_{n \geq 0} C^{(n)}$. As noted in Appendix A.2 of DKS, for each $C^{(n)}$, there exists a compact metric space $X^{(n)}$ such that $X^{(n)} \simeq \mathcal{K}\big(\mathcal{F}(\Delta(C^{(n)} \times X^{(n)})\big)$. Let $X := \bigcup_{n \geq 0} X^{(n)}$ denote the space of dynamic choice problems.

(∗) University of Pennsylvania <ddill@sas.upenn.edu>
(†) Florida State University <rvk3570@gmail.com>
(‡) Duke University <p.sadowski@duke.edu>

It is easy to see that for each tuple $\big((u_s), \delta, \Pi, \mathscr{C}\big)$, there exists a *unique* function $V : X \times \Theta \times (S \cup \{s_0\}) \to \mathbb{R}$ that satisfies the functional equation

$$V(x, \theta, s) = \max_{P \in \mathscr{P}} \left[\sum_{J \in P} \Big(\max_{f \in x} \sum_{s' \in J} \pi(s' \mid J)\, \mathbf{E}^{f(s')} \big[u_{s'}(c) + \delta V(y, \tau(P, s, \theta), s') \big] \Big) - \rho(P, \theta) \right]$$

The proof is a straightforward variation of the proof of Proposition 2.2 in DKS because each $x \in X$ is an element of $X^{(n)}$ for some n (and hence is also a member of $X^{(m)}$ for every $m \geq n$), and the arguments in the proof of Proposition 2.2 of DKS apply to the subdomain $X^{(n)}$. Clearly, the particular choice of n does not matter.

1.2. Canonical Information Cost Structures

Recall that $\mathscr{P}$ is the space of all partitions of S. Consider the following definitions.

- $\mathfrak{X}_1 := \mathbb{R}_+^{\mathscr{P}}$
- $\mathfrak{X}_2 := \big(\mathbb{R}_+ \times \mathfrak{X}_1^S\big)^{\mathscr{P}}$
- $\mathfrak{X}_3 := \big(\mathbb{R}_+ \times \mathfrak{X}_2^S\big)^{\mathscr{P}}$
- $\vdots$
- $\mathfrak{X}_{n+1} := \big(\mathbb{R}_+ \times \mathfrak{X}_n^S\big)^{\mathscr{P}}$
- $\vdots$

It is easy to see that $\mathfrak{X}_1$ is a separable metric space (with the standard Euclidean metric). It follows then, that for each $n > 1$, $\mathfrak{X}_{n+1}$ is a separable metric space because $\mathfrak{X}_n$ is a separable metric space. A typical member of $\mathfrak{X}_n$ is ζ_n. Each ζ_n has the form $\zeta_n = \big((\rho^P, \boldsymbol{\zeta}_{n-1}^P)_{P \in \mathscr{P}}\big)$, where $\boldsymbol{\zeta}_{n-1}^P = (\zeta_{n-1,s}^P)_{s \in S}$, $\zeta_{n-1,s}^P \in \mathfrak{X}_{n-1}$ and $\rho^P \in \mathbb{R}_+$ for each $P \in \mathscr{P}$.

Let $\varphi_1 : \mathfrak{X}_2 \to \mathfrak{X}_1$ be given by $\varphi_1(\zeta_2) = \varphi_1(\rho, \boldsymbol{\zeta}_1) = \rho \in \mathfrak{X}_1$. Now define recursively, for $n > 1$, $\varphi_n : \mathfrak{X}_{n+1} \to \mathfrak{X}_n$ as follows: for each $(\rho, \boldsymbol{\zeta}_n) \in \mathfrak{X}_{n+1}$, $\varphi_n(\rho, \boldsymbol{\zeta}_n) = (\rho, \varphi_{n-1}(\boldsymbol{\zeta}_{n-1}))$.

Let

$$\mathfrak{X}^* := \bigtimes_{n=1}^{\infty} \mathfrak{X}_n$$

denote the collection of all *information cost hierarchies*. An information cost hierarchy $\zeta = (\zeta_1, \zeta_2, \dots) \in \mathfrak{X}^*$ is *consistent* if $\mathfrak{X}_{n-1} = \varphi_{n-1}(\mathfrak{X}_n)$ for all $n > 1$. The *canonical space* of information costs is

$$\mathfrak{X} := \big\{\zeta \in \mathfrak{X}^* : \zeta \text{ is consistent}\big\}$$

Theorem 1. *The set $\mathfrak{X}$ is homeomorphic to* $(\mathbb{R}_+ \times \mathfrak{X}^S)^{\mathscr{P}}$.

We write the homeomorphism as $\mathfrak{X} \simeq \big(\mathbb{R}_+ \times \mathfrak{X}^S\big)^{\mathscr{P}}$. The theorem can be proved by using the techniques in Epstein and Zin (1989).

Notice that each $\zeta \in \mathfrak{X}$ can then be written as $\zeta \simeq \big((\rho^P, \tilde{\boldsymbol{\zeta}}^P)_{P \in \mathscr{P}}\big)$ where $\tilde{\boldsymbol{\zeta}}^P = (\tilde{\zeta}_s^P)_{s \in S}$, $\tilde{\zeta}_s^P \in \mathfrak{X}$ and $\rho^P \in \mathbb{R}_+$ for all $P \in \mathscr{P}$.

1.3. Order for Information Strategies

Let $\mathfrak{p} : S \cup \{s_0\} \times \mathscr{X} \to \Delta^N$ be a *randomization device* (where N is such that $|\mathscr{P}| \leq N < \infty$) for each (s, ζ) with the property that $\mathfrak{p}(s_0, \zeta)$ is a degenerate probability measure. Let $\sigma : S \cup \{s_0\} \times \mathscr{X} \times \mathbb{N} \to \mathscr{P}$ denote an *information plan*. Intuitively, given (s, ζ), the randomization device reveals some further information, and can be viewed as a probability measure over $\mathscr{P}$ (because only finitely many choices of partitions are possible). Then, given the additional randomization, and as a function of (s, ζ), an information choice is made according to σ.

Definition 1.1. An information plan $\tilde{\sigma}$ is *more informative give the constraint* $\tilde{\zeta}$ *than the plan* σ *is given the constraint* ζ if

(a) $\tilde{\sigma}(s, \tilde{\zeta}, n)$ is finer than $\sigma(s, \zeta, n)$ for all $s \in S \cup \{s_0\}$ and $n \in \mathbb{N}$, and
(b) $\tilde{\sigma}$ is more informative given the constraint $\tilde{\zeta}'_{s'}\big(\tilde{\sigma}(s, \tilde{\zeta}, n)\big)$ than σ is given the constraint $\zeta'_{s'}\big(\sigma(s, \zeta, n)\big)$.

Each pair $(\sigma, \mathfrak{p})$ induces a *dynamic cost function* $W : S \cup \{s_0\} \times \mathscr{X} \to \mathbb{R}_+ \cup \{\infty\}$ defined recursively as

$$W(\zeta, s; \sigma, \mathfrak{p}) = \sum_n \mathfrak{p}(n) \left[\rho\big(\sigma(s, \zeta, n), \zeta\big) + \delta \sum_{s' \in S} \pi_s(s') W\big(\zeta'_s\big(\sigma(s, \zeta, n)\big), s; \sigma, \mathfrak{p}\big) \right]$$

Definition 1.2. The cost structure $\tilde{\zeta}$ *dominates* ζ if for all $(\sigma, \mathfrak{p})$, there exists $\tilde{\sigma}$ such that

(a) $\tilde{\sigma}$ is more informative give the constraint $\tilde{\zeta}$ than the plan σ is given the constraint ζ, and
(b) $W(\tilde{\zeta}, s_0; \tilde{\sigma}, \mathfrak{p}) \leq W(\zeta, s; \sigma, \mathfrak{p})$.

Definition 1.3. The plan σ is *undominated* given the constraint ζ if there exists $\mathfrak{p}$ such that for all σ where σ' is more informative than σ, $W(\zeta, s_0; \sigma, \mathfrak{p}) < W(\zeta, s_0; \sigma', \mathfrak{p})$.

1.4. Identifying Information Cost Structures

We shall first describe identification for one-period cost structures and then describe how to extend these ideas to two-period structures.

1.4.1. *Identifying One-Period Cost Structures*

We begin with some notation. Let $\zeta_1 \in \mathscr{X}_1$, and for each P, let

$$U(x \mid P) := \sum_{J \in P} \left[\max_{f \in x} \sum_s \pi_0(s \mid J) \, \mathbf{E}^{f(s)}[u_s(c)] \right] \pi_0(J)$$

$$V_1(x, \zeta_1, s_0; P) := U(x \mid P) - \rho(P, \zeta_1)$$

and

$$[\mathbf{1.1}] \qquad V_1(x, \zeta_1, s_0) = \max_{P \in \mathscr{P}} V_1(x, \zeta_1, s_0; P)$$

Recall the domain: C is sigma-compact, ie, there exists a collection of compact sets (C_m) such that (i) $C^{(m)} \subset C^{(m+1)}$ and (ii) $C = \bigcup_{m \geq 0} C^{(m)}$. Let $c^+_{s,m}, c^-_{s,m} \in C^{(m)}$ be the best and worst elements in $C^{(m)}$ in state s, respectively. (Existence is clear because $\succsim$ is continuous.) Then, ℓ^*_m is the consumption stream that pays $c^+_{s,m}$ in state s in each period. Similarly, $\ell_{*,m}$ is the consumption stream that pays $c^-_{s,m}$ in state s in each period.

Then a one-period problem is $X_1^{(m)} := \mathscr{K}\big(\mathscr{F}(\Delta(C^{(m)} \times \ell^\dagger))\big)$, while a t-period problem is $X_t^{(m)} := \mathscr{K}\big(\mathscr{F}(\Delta(C^{(m)} \times X_{t-1}^{(m)}))\big)$. As before, $X_m \simeq \mathscr{K}\big(\mathscr{F}(\Delta(C_m \times X_m))\big)$. As in DKS, let $x_t^{(m)}(P)$ be the t-period problem that pays either $c^+_{s,m}$ or $c^-_{s,m}$ and is aligned with P in the first period.

Lemma 1.4. Let σ be an information plan that chooses P given the cost structure ζ_1. Suppose this choice is undominated.[1] Then, there exists $m > 0$ such that

$$P \in \arg\max_{Q \in \mathscr{P}} V_1(x_1^{(m)}(P), \zeta_1; Q)$$

Proof. By the arguments in DKS, P is *strongly aligned* with $x_1^{(m)}$ for all m. Recall also that because the u_s's are unbounded, we have $u_s(c^{\pm}_{s,m}) \to \pm\infty$ as $m \to \infty$. We claim that for m large enough, we have

$$[\mathbf{1.2}] \qquad V_1(x_1^{(m)}(P), \zeta_1, s_0; P) > V_1(x_1^{(m)}(P), \zeta_1, s_0; Q)$$

for all $Q \neq P$. To see this, first consider the case where Q costs more than P. In that case, $\rho(P, \zeta_1) < \rho(Q, \zeta_1)$, but $U(x_1^{(m)}(P) \mid P) \geq U(x_1^{(m)}(P) \mid Q)$ which establishes [1.2].

Now consider the case where Q costs (weakly) less than P. Because P is undominated, this means Q is not finer than P. Then, although we have $\rho(P, \zeta_1) \geq \rho(Q, \zeta_1)$, $\max_{Q \in \mathscr{P}} \big[\rho(P, \zeta_1) - \rho(Q, \zeta_1)\big]$ is finite. Because Q is not finer than P, $U(x_1^{(m)}(P) \mid P) > U(x_1^{(m)}(P) \mid Q)$ and for large enough m, we have

$$U(x_1^{(m)}(P) \mid P) - U(x_1^{(m)}(P) \mid Q) > \max_{Q \in \mathscr{P}} \big[\rho(P, \zeta_1) - \rho(Q, \zeta_1)\big]$$

which establishes [1.2]. □

Proposition 1.5. The one-period cost structure ζ_1 in [1.1] is identified up to dominance.

Proof. Suppose P is undominated. Then, $P = \arg\max_Q V\big(x_1^{(m)}(P), \zeta, s_0; Q\big)$ by Lemma 1.4. Now suppose $\alpha \in [0, 1]$ is such that $(1-\alpha)\ell^*_m + \alpha \ell_{*,m} \sim x_1^{(m)}(P)$.[2] Then, $\rho_1(P, \zeta_1) = \alpha\big[V_1(\ell^*_m, \zeta_1, s_0) - V_1(\ell_{*,m}, \zeta_1, s_0)\big]$. □

(1) That is, for all Q finer than P, $\rho(Q, \zeta_1) > \rho(P, \zeta_1)$.

(2) Such an α exists because $V_1(\ell_{*,m}, \zeta_1) < V_1\big(x_m^{(1)}(P), \zeta_1\big) \leq V_1\big(\ell^*_m, \zeta_1\big)$.

1.5. Identifying Two-Period Cost Structures

Let $\zeta_2 \in \mathscr{Z}_2$. Then, for each information plan σ,

$$V_2(x, \zeta_2, s_0; \sigma) = \sum_{J \in P} \left[\max_{f \in x} \sum_s \pi_0(s \mid J)\, \mathbf{E}^{f(s)} \left[u_s(c) + \delta V_1(y, \zeta^P_{1,s}, s; \sigma)\right] \pi_0(J) \right] - \rho(P, \zeta_2)$$

where $\sigma(s_0, \zeta_2, \cdot) = P$. Let

$$V_2(x, \zeta_2, s_0) := \max_\sigma V_2(x, \zeta_2, s_0; \sigma) \tag{1.3}$$

Proposition 1.6. The two-period cost structure ζ_2 in [1.3] is identified up to dominance.

Proof. Suppose σ is undominated given the constraint ζ_2. Then, the $\mathfrak{p}$ in the definition can be explicitly computed because there are only finitely many partitions and finitely many (linear) inequalities that define dominance. For any information plan σ with Q as the first period choice of partition, define the two-period problem

$$x_2(\sigma) := \left\{ f_{2,J} : J \in \sigma(s_0, \zeta_2, \cdot) \right\}$$

where $f_{2,J}$ is defined as (following DKS)

$$f_{2,J}(s) := \begin{cases} \left(c^+_{s,m}, \left([x^{(1)}_m(P); \mathfrak{p}(P)]\right)_{P \in \mathscr{P}} \right) & s \in J \\ \ell^*_m(s) & s \notin J \end{cases}$$

where $[x^{(m)}_1(P); \mathfrak{p}(P)]$ is the lottery that delivers $x^{(m)}_1(P)$ with probability $\mathfrak{p}(P)$ for each $P \in \mathscr{P}$, and where $\mathfrak{p}(P) := \sum_n \mathfrak{p}(n; s, \zeta^P_{1,s}) \mathbb{1}_{\{\sigma(s, \zeta^P_{1,s}, n) = P\}}(n)$. Notice that $W_2(\zeta_2, \sigma, \mathfrak{p})$ is given by

$$W(\zeta_2, s_0, \sigma, \mathfrak{p}) = \rho(Q, \zeta_2) + \delta \sum_s \pi_0(s) W\left(\zeta_{1,s}(\sigma(s_0, \zeta, \cdot)), s, \sigma, \mathfrak{p}\right)$$

As in Lemma 1.4 and Proposition 1.5, for some $m > 0$ large enough, we can find α that solves $(1-\alpha)\ell^*_m + \alpha \ell_{*,m} \sim x_2(\sigma)$. Then, $W(\zeta_2, s_0, \sigma, \mathfrak{p}) = \alpha\left[V_2(\ell^*_m, \zeta_2, s_0) - V_2(\ell_{*,m}, \zeta_2, s_0)\right]$.

Now, for any P, s, and J such that $s \in J \in P$, let $x_2(\sigma; P, s, \alpha, \varepsilon) := \left(x_2(\sigma) \setminus f_{2,J}\right) \cup \hat{f}^{P,s,\alpha,\varepsilon}_{2,J}$, where

$$\hat{f}^{P,s,\alpha,\varepsilon}_{2,J} := \begin{cases} \left(c^+_{s,m}, \left([(1-\varepsilon)x^{(1)}_m(Q) + \varepsilon[(1-\alpha)\ell^*_m + \alpha \ell_{*,m}]; \mathfrak{p}(Q)]\right) \right) & s' = s,\ Q = P \\ \left(c^+_{s,m}, \left([x^{(m)}_1(Q); \mathfrak{p}(Q)]\right)_{Q \neq P} \right) & s' = s \\ f_{2,J}(s') & s' \neq s \end{cases}$$

Because σ is undominated, for $\varepsilon > 0$ small enough and m large enough (as in Lemma 1.4), we still have

$$\sigma = \arg\max_{\sigma'} V_2(x_2(\sigma; P, s, \alpha, \varepsilon), \zeta_2, s_0; \sigma')$$

so that the optimal first period choice of information is still Q.

Now, let α be such that $V_2(x_2(\sigma; P, s, \alpha, \varepsilon), \zeta_2, s_0) = V_2(x_2(\sigma), \zeta_2, s_0)$. Then, it follows immediately (from the representation in [1.3]) that $\rho(P, \zeta^P_{1,s}) = \alpha\big[V_1(\ell^*_m, \zeta^P_{1,s}, s) - V_1(\ell_{*,m}, \zeta^P_{1,s}, s)\big]$.

Finally, we find that $\rho(Q, \zeta_2) = W_2(\zeta_2, s_0, \sigma, \mathfrak{p}) - \delta \sum_s \pi_0(s) \sum_P \mathfrak{p}(P)\rho(P, \zeta^P_{1,s})$. □

1.6. Identifying t-Period and Infinite Horizon Cost Structures

Let $\zeta_t \in \mathfrak{X}_t$. Then, for each information plan σ,

$$V_t(x, \zeta_t, s_0; \sigma) = \sum_{J \in P} \left[\max_{f \in x} \sum_s \pi_0(s \mid J)\, \mathbf{E}^{f(s)} \big[u_s(c) + \delta V_{t-1}(y, \zeta^P_{t-1,s}, s; \sigma)\big]\pi_0(J)\right] - \rho(P, \zeta_t)$$

where $\sigma(s_0, \zeta_2, \cdot) = P$. Let

$$V_t(x, \zeta_t, s_0) := \max_\sigma V_t(x, \zeta_t, s_0; \sigma)$$

The cost structure ζ_t can be identified just as in Proposition 1.6. We omit the details.

Finally, to identify $\zeta \in \mathfrak{X}$, we observe that each ζ can be approximated by its finite horizon truncations. It is also clear that because all the costs are non-negative, the functions $W_t(\zeta_t, \cdot)$ increase to $W(\zeta, \cdot)$. Again, details are omitted because they are similar to the case of ICPs in DKS.

2. Convex Duality

We review some notions from convex analysis. Our review follows Ekeland and Turnbull (1983).

Let X be a Banach space, X^* its norm dual, $C \subset X$, and $f : C \to X$ a convex and Lipschitz function. The *subdifferential* of f at $x \in C$ is $\partial f(x) := \{x^* \in X^* : \langle y - x, x^* \rangle \leq f(y) - f(x) \text{ for all } y \in C\}$. A necessary and sufficient condition for the existence of a subdifferential at $x \in C$ is that there exists $K \geq 0$ such that for all $y \in X$, $f(x) - f(y) \leq K \|y - x\|$. To see this, recall that the set $\text{epi}(f) := \{(x, t) \in X \times \mathbb{R} : t \geq f(x)\}$, the *epigraph* of the function f, is a convex set (if, and only if, f is a convex function). For each $x \in C$, we define $A(x) := \{(y, t) \in X \times \mathbb{R} : f(x) - t > K \|y - x\|\}$. It is easy to see that the set $A(x)$ is (i) nonempty, (ii) convex, and (iii) open. It is also easy to show that $\text{epi}(f) \cap A(x) = \varnothing$, so there exists a non-vertical hyperplane that separates the two sets. Following the arguments in Gale (1967), we can conclude that $\partial f(x) \neq \varnothing$, and moreover, there exists $x^* \in \partial f(x)$ such that $\|x^*\| \leq K$. This is the content of the Duality Theorem of Gale (1967). (Indeed, Gale (1967) also shows that local Lipschitzness is a necessary condition for $\partial f(x)$ to be nonempty.) We will rely on the following result in the sequel.

Proposition 2.1 (Duality Theorem in Gale (1967))**.** Let $C \subset X$ be convex and suppose $f : C \to \mathbb{R}$ is convex and Lipschitz of rank K. Then, there exists $x^* \in \partial f(x)$ such that $\|x^*\| \leq K$.

In what follows, we will denote by $\partial_K f(x) := \{x^* \in \partial f(x) : \|x\| \leq K\}$. For each $x^* \in X^*$ and $a \in \mathbb{R}$, we can define the continuous affine functional $\varphi(\cdot, x^*) : X \to \mathbb{R}$ as $\varphi(y; x^*) := \langle y, x^* \rangle - a$. The function $\varphi \leq f$ for all $y \in C$ if, and only if, $\langle y, x^* \rangle - a \leq f(y)$, and is *exact* at $x \in C$ if $\varphi(x; x^*) = f(x)$. If φ is exact, the value of a which makes it so is given by $-a(x^*) := f(x) - \langle x, x^* \rangle$. Therefore, $x^* \in \partial f(x)$ if, and only if, the continuous affine functional $\varphi(y; x^*) = f(x) + \langle y - x, x^* \rangle \leq f(y)$ for all $y \in C$ with $\varphi(x; x^*) = f(x)$. In other words, $x^* \in \partial f(x)$ if, and only if, $\varphi(y; x^*) = f(x) + \langle y - x, x^* \rangle$ is a supporting hyperplane for the graph of f at x.

Notice that for any intercept $a \geq a(x^*)$, $\langle x, x^* \rangle - a < \langle x, x^* \rangle - a(x^*)$, so $a(x^*) = \inf[a \in \mathbb{R} : f(x) \geq \langle x, x^* \rangle - a] = \sup[x \in C : \langle x, x^* \rangle - f(x)]$. This smallest intercept is the *Fenchel conjugate* of f, and is denoted by $f^\star : X^* \to \mathbb{R} \cup \{-\infty, +\infty\}$, and is given by

$$f^\star(x^*) := \sup_{x \in C} \big[\langle x, x^* \rangle - f(x) \big]$$

Proposition 2 of Ekeland and Turnbull (1983) shows that $x^* \in \partial f(x)$ if, and only if, $f(x) + f^\star(x^*) = \langle x, x^* \rangle$.

By Proposition 2.1, it follows that for Lipschitz f, the conjugate function is given by $f^\star(x^*) := \max_{x \in C} \big[\langle x, x^* \rangle - f(x) \big]$. We now show that for positively homogeneous functions, the conjugate function $f^\star$ is identically 0.

Proposition 2.2. Let $C \subset X$ be a convex cone, and let $f : C \to \mathbb{R}$ be convex and Lipschitz. Then, the following are equivalent:

(a) f is positively homogeneous, ie, $f(\lambda x) = \lambda f(x)$ for all $\lambda > 0$;

(b) $f^\star(x^*) \in \mathbb{R}$ implies $f^\star(x^*) = 0$.

Proof. Suppose $f^\star = 0$. Fix $x \in C$, and recall that because f is convex and Lipschitz, there exists $x^* \in \partial f(x)$. This implies $f(x) = \langle x, x^* \rangle$. It is easy to see that $x^* \in \partial f(\lambda x)$ for all $\lambda > 0$, so that $f(\lambda x) = \lambda f(x)$. That is, f is positively homogeneous.

Now suppose f is positively homogeneous. Fix $x \in C$ and suppose $x^* \in \partial f(x)$. We will first show that for any $\lambda > 0$, $x^* \in \partial f(\lambda x)$. Then, by the definition of ∂f, for any $y \in C$, $\langle y - x, x^* \rangle \leq f(y) - f(x^*)$. Now let $\lambda > 0$ and let $y \in C$ be arbitrary. Because C is a cone, there exists $z \in C$ such that $\lambda z = y$. This implies $\langle y - \lambda x \rangle = \lambda \langle z - x^* \rangle \leq \lambda \big[f(z) - f(x) \big] = f(y) - f(\lambda x)$, which proves that $x^* \in \partial f(x)$ implies $x^* \in \partial f(\lambda x)$ for all $\lambda > 0$.

Now suppose x^* is such that $f^\star(x^*) \in \mathbb{R}$. Because f is positively homogeneous, we have $f(0) = 0$. (To see this, note that $f(0) = f(2 \times 0) = 2 f(0)$ which implies $f(0) = 0$.) Therefore, $f^\star(x^*) \geq \langle 0, x^* \rangle - f(0) = 0$. Now suppose $f^\star(x^*) > 0$. Then, for any $\varepsilon \in (0, f^\star(x^*))$, there exists $x \in C$ such that $f^\star(x^*) - \varepsilon = \langle x, x^* \rangle - f(x) > 0$. But then we can choose $\lambda > 0$ such that $\langle \lambda x, x^* \rangle - f(\lambda x) > f^\star(x^*)$, which is a contradiction. Therefore, it must be that $f^\star(x^*) = 0$. □

This allows us to establish the following corollary.

Corollary 2.3. Let $C \subset X$ be a convex cone, and $f \in \mathbb{R}^C$ be convex, Lipschitz, and positively homegeneous. Then, there exists a weak* compact set $\mathfrak{M} \subset X^*$ such that $f(x) = \max[\langle x, x^* \rangle : x^* \in \mathfrak{M}]$.

Proof. We have already established that for each $x \in C$, there exists $x^* \in \partial f(x)$ such that $\|x^*\| \leq K$, where K is the Lipschitz constant of f. We have also established that $x^* \in \partial f(\lambda x)$ for all $\lambda \geq 0$. Therefore, $f(y) \geq \langle y, x^* \rangle$ for all $y \in C$. Letting $\mathfrak{M} = \text{cl}(\{x^* \in \partial f(x) : x \in C, \ \|x^*\| \leq K\})$ (in the weak* topology) establishes the claim. □

If C is convex and $A \subset C$ is also convex, then $f : C \to \mathbb{R}$ is *A-affine* if for all $x \in C$, $a \in A$, and $t \in (0,1)$, we have $f(tx + (1-t)a) = tf(x) + (1-t)f(a)$.

For a fixed $x \in C$, notice that f is affine on the set $\text{ch}(\{x\} \cup A)$. Let $\mathscr{E}_x$ be the collection of all (convex) subsets of C such that if $E \in \mathscr{E}_x$ then (i) $x \in E$ and (ii) $f|_E$ is affine. A simple application of Zorn's lemma shows that for each $x \in C$, there is a largest set E_x that contains x and where $f|_{E_x}$ is affine.

Notice that there exist $x \in X$ such that this maximal set E_x is not unique. Indeed, for any $a \in A$, and $x, y \in C$ such that f is not affine on $[x, y]$ (the closed line segment joining x and y), then $a \in E_x \cap E_y$, but $E_x \cup E_y$ (or it's convex hull) is not a member of $\mathscr{E}_a$.

If f is Lipschitz continuous (as we shall assume below), then it is easy to see that the set E_x must be closed as well.

Proposition 2.4. Let $C \subset X$ be a convex set, and $f \in \mathbb{R}^C$ be convex and Lipschitz of rank K. Let $A \subset C$ be convex and suppose that $\mathbf{0} \in A$, $f(\mathbf{0}) = 0$, and that f is A-affine. Then, for each x, there exists $x^* \in X^*$ such that $x^* \in \partial f_K(y)$ for all $y \in E_x$ where E_x is defined above. Moreover, there exists a weak* compact set $\mathfrak{M}_f \subset X^*$ such that $f(x) = \max[\langle x, x^* \rangle : x^* \in \mathfrak{M}_f]$ and $\langle a, x^* \rangle$ is independent of $x^* \in \mathfrak{M}_f$ for all $a \in A$.

Proof. Fix $x \in C$, let $y_1, \ldots, y_n \in E_x$, and define $y := \frac{1}{n} \sum_i n$. Then, by Proposition 2.1, there exists $y^* \in \partial_K f(y)$. Recall the affine function $\varphi(\cdot, y^*) X :\to \mathbb{R}$ given by

$$\varphi(x; y^*) := \langle x - y, y^* \rangle + f(y)$$

The affine function φ satisfies the following two properties:

- $f(x) \geq \varphi(x; y^*)$ for all $x \in C$, and
- $f(y) = \varphi(y; y^*)$.

The first requirement implies that $f(y_i) \geq \varphi(y_i; y^*)$ for all $i = 1, \ldots, n$. Summing up and dividing by n, we see that $\frac{1}{n} \sum_i f(y_i) \geq \frac{1}{n} \sum_i \varphi(y_i; y^*)$. However, f restricted to E_x is affine which implies $\frac{1}{n} \sum_i f(y_i) = f(y)$; similarly, φ is affine, which implies $\frac{1}{n} \sum_i \varphi(y_i; y^*) = \varphi(y; y^*)$.

But we have noted above that $f(y) = \varphi(y; y^*)$, which is possible if, and only if, $f(y_i) = \varphi(y_i; y^*)$ for all $i = 1, \ldots, n$. But this is equivalent to saying that $y^* \in \partial_K f(y_i)$.

For any $y \in E_x$, $\partial_K f(y)$ is a (nonempty) closed (and hence compact) subset of $\{x^* \in X^* : \|x^*\| \leq K\}$.[3] Thus, $(\partial_K f(y))_{y \in E_x}$ is a collection of closed subsets of the compact set $\{x^* \in X^* : \|x^*\| \leq K\}$. But we have just established that for any $y_1, \ldots, y_n \in E_x$, $\bigcap_{i=1}^n \partial_K f(y_i) \neq \varnothing$. In other words, the collection of closed sets $(\partial_K f(y))_{y \in E_x}$ has the finite intersection property. The compactness of $\{x^* \in X^* : \|x^*\| \leq K\}$ then implies that $\bigcap_{y \in E_x} \partial_K f(y) \neq \varnothing$. Thus, there exists $\zeta_x \in \bigcap_{y \in E_x} \partial_K f(y)$ which proves the first part.

Fix this ζ_x and notice that $\varphi(y; \zeta_x) = f(y)$ for all $y \in E_x$. Because $\mathbf{0} \in A$, this implies $\varphi(\mathbf{0}; \zeta_x) = 0$. In other words, $f^\star(\zeta_x) = 0$. (In geometric terms, the supporting hyperplane determined by ζ_x passes through the origin.) Now, let $\mathfrak{M}_f := \mathrm{cl}\,(\{\zeta_x \in X^* : x \in C\})$. It is immediate that $\mathfrak{M}_f$ is closed. Because $f(a) = \langle a, \zeta_x \rangle$ for all $x \in C$, it follows that the same holds for all $x^* \in \mathfrak{M}_f$, which completes the proof. □

We end with an easy observation.

Lemma 2.5. Let $C \subset X$ be a convex set, and $f \in \mathbb{R}^C$, and $\mathfrak{M}_f$ a weak* compact subset of X^* such that for all $x \in C$, $f(x) = \max[\langle x, x^* \rangle : x^* \in \mathfrak{M}_f]$. (This implies f is convex and Lipschitz of rank K for some K.) Let $C_0 \subset C$ be convex. Then, the following are equivalent.

(a) The function $f|_{C_0}$ is linear.

(b) There exists $x_0^* \in \mathfrak{M}_f$ such that $x_0^* \in \bigcap_{x \in C_0} \partial_K f(x)$ (which is equivalent to saying that $f(x) = \langle x, x_0^* \rangle$ for all $x \in C_0$).

Proof. It is easy to see that (b) implies (a). To prove that (a) implies (b), we shall prove the contrapositive. So, suppose $\bigcap_{x \in C_0} \partial_K f(x) = \varnothing$. Then, there exist $x_1, \ldots, x_n \in C_0$ such that $\bigcap_{i=1}^n \partial_K f(x_i) = \varnothing$. Let $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.

Then, for all $x^* \in \mathfrak{M}_f$ we have

- $\langle x_i, x^* \rangle \leq \langle x_i, x_i^* \rangle = f(x_i)$ for all $i = 1, \ldots, n$, and
- $\langle x_i, x^* \rangle < \langle x_i, x_i^* \rangle = f(x_i)$ for some $i \in \{1, \ldots, n\}$

This implies $\frac{1}{n} \sum_i \langle x_i, x^* \rangle = \langle \bar{x}, x^* \rangle < \frac{1}{n} \sum_i f(x_i)$. Since this is true for all $x^* \in \mathfrak{M}_f$, and because $\mathfrak{M}_f$ is compact, it follows that $f(\bar{x}) = \max[\langle \bar{x}, x^* \rangle : x^* \in \mathfrak{M}_f] < \frac{1}{n} \sum_i f(x_i)$, which proves that f is not linear on C_0, as claimed. □

3. Static Representation

We establish the existence of a static representation as in [D.1] of DKS. Formally, we establish the following.

(3) By the Banach-Alaoglu Theorem — see, for instance, Theorem 6.25 of Aliprantis and Border (1999) — the set $\{x^* \in X^* : \|x^*\| \leq K\}$ is a weak* compact subset of the dual X^*.

Theorem 2. *A binary relation $\succsim$ satisfies Axioms 1–5 if, and only if, there exists a function $V : X \to \mathbb{R}$ that represents $\succsim$ and has a representation of the form*

$$V(x) = \max_{P \in \mathfrak{M}_p^\sharp} \sum_{J \in P} \left[\max_{f \in x} \sum_s \pi_0(s \mid J) \left[u_s(f_1(s)) + v_s(f_2(s), P) \right] \pi_0(J) \right]$$

where $\mathfrak{M}_p^\sharp$ is a finite collection of partitions P of S, $u_s \in \mathbf{C}(C)$, and $v_s(\cdot, P) \in \mathbf{C}(X)$ for each $s \in S$ and $P \in \mathfrak{M}_p^\sharp$, with the property that for all $P, P' \in \mathfrak{M}_p^\sharp$, $s \in S$, $v_s(\cdot, P)|_L = v_s(\cdot, P')|_L$.

3.1. Abstract Static Representation

With a view towards proving Theorem 2, in this section we provide an abstract static representation.

Let Y be a compact metric space. Then, $\Delta(Y)$ is the space of probability measures on Y. For compact metric spaces $Y_1, \dots, Y_n$, we will consider the product space $Z := \Delta(Y_1) \times \dots \times \Delta(Y_n)$. We are interested in the space of closed subsets of Z, $\mathscr{K}(Z)$ (endowed with the Hausdorff metric), and also in the space of closed and convex subsets $\mathscr{K}_c(Z)$. It is well known that $\mathscr{K}_c(Z)$ is a closed subset of $\mathscr{K}(Z)$.

The convex hull of a set A (in the relevant ambient vector space) is denoted by ch A. If the ambient vector space has a topology, then cch A denotes the closed convex hull of A.

Recall that $\mathbf{C}(Y_i)$ is the space of all uniformly continuous functions on Y_i and for $\alpha_i \in \Delta(Y_i)$ and $\mathfrak{u}_i \in \mathbf{C}(Y_i)$, $\mathfrak{u}_i(\alpha_i) := \int_{Y_i} \mathfrak{u}_i(y_i) \, \mathrm{d}\alpha_i(y_i) =: \langle \alpha_i, \mathfrak{u}_i \rangle$; endowed with the supremum norm, $\mathbf{C}(Y_i)$ is a Banach space. For each $s \in S$, let $L_s \subset \Delta(Y_s)$ be a closed subset, and define $L := \bigtimes_{s \in S} L_s$. Fix $\ell_s^\dagger \in L_s$, and define $\mathfrak{U}_{Y_s, \ell_s^\dagger} := \{\mathfrak{u}_s \in \mathbf{C}(Y_s) : \mathfrak{u}_s(\ell_s^\dagger) = 0, \ \|\mathfrak{u}\|_\infty = 1\}$. Finally, define $\mathfrak{U} := \left\{(p_1 \mathfrak{u}_1, \dots, p_n \mathfrak{u}_n) : \mathfrak{u}_s \in \mathfrak{U}_{Y_s, \ell_s^\dagger}, \ p_s \geq 0, \sum_s p_s = 1\right\}$. The space $\mathfrak{U}$ will serve as our *subjective state space* below. It is useful to reconsider $\mathfrak{U}$ as $\mathfrak{U} := \left\{(p, \mathfrak{u}) : p := (p_1, \dots, p_n) \in \Delta(S), \ \mathfrak{u} := (\mathfrak{u}_1, \dots, \mathfrak{u}_n) \in \bigtimes_{s \in S} \mathfrak{U}_{Y_s, \ell_s^\dagger}\right\}$.

Specifically, if we consider the domain X, then each $Y_s := C \times X$, which then results in a corresponding definition of $\mathfrak{U}$, in which case we write $\mathfrak{U}_{s, \ell^\dagger(s)}$ for $\mathfrak{U}_{Y_s, \ell_s^\dagger}$.

Theorem 3. *Let $\succsim$ be a binary relation on X. Then, the following are equivalent:*

(a) *$\succsim$ satisfies Basic Properties (Axiom 1) and L-Independence (Axiom 2(a)).*

(b) *There exists a metric space of continuous functions $\mathfrak{U}$ (as defined above) and a minimal set $\mathfrak{M}$ of finite, normal, and positive charges[4] on $\mathfrak{U}$ that is weak* compact such that*

  [i] *For all $\ell \in L$ and $s \in S$, $\int_{\mathfrak{U}} p_s \mathfrak{u}_s(\ell_s) \, \mathrm{d}\mu(p, \mathfrak{u})$ is independent of $\mu \in \mathfrak{M}$, and*

  [ii] *The function $V : X \to \mathbb{R}$ given by*

[♦] $$V(x) := \max_{\mu \in \mathfrak{M}} \left[\int_{\mathfrak{U}} \max_{f \in x} \sum_s p_s \mathfrak{u}_s(f(s)) \, \mathrm{d}\mu(p, \mathfrak{u}) \right]$$

  represents $\succsim$.

(4) A charge is a finitely additive measure.

The proof of Theorem 3 follows immediately from Propositions 3.10, 3.11, and 3.12 below.

3.1.1. *Algebraic Representation*

Recall that our domain is $X \simeq \mathcal{K}(\mathcal{F}(\Delta(C \times X)))$. We shall first show that under our assumptions, every closed subset is indifferent to its closed convex hull.

Lemma 3.1. If $\succsim$ satisfies Axiom 1, then for each $x \in \mathcal{K}(Z)$, $x \sim \text{cch}(x)$.

Proof. First consider $x \in X$ that is finite and follow Ergin and Sarver (2010a, Lemma 2). Notice that $\text{cch}(x) \succsim x$ by Monotonicity (Axiom 1(d)). Let $x^0 := x$, and for each $k \geq 1$, define $x^k := \frac{1}{2}x^{k-1} + \frac{1}{2}x^{k-1}$. Then, by Aversion to Randomization (Axiom 1 (e)), $x^{k-1} \succsim x^k$. In other words, by Order (Axiom 1(a)), $x \succsim x^k$ for all $k \geq 1$. But notice that $d(x^k, \text{cch}(x)) \to 0$ as $k \to \infty$. Therefore, by Continuity (Axiom 1(b)), it follows that $x \succsim \text{cch}(x)$, which proves that $x \sim \text{cch}(x)$ for all finite subsets of X.

Now consider the general case, where $x \in X$ is arbitrary. Then, there exists a sequence of finite sets (x_m) such that (i) $x_m \subset x$ for all m, and (ii) $d(x_m, x) \to 0$ (in the Hausdorff metric). But each $x_m \sim \text{cch}(x_m)$. It is also easy to see that $d(\text{cch}(x), \text{cch}(x_m)) \to 0$ as $m \to \infty$. Continuity (Axiom 1(b)) now implies that $x \sim \text{cch}(x)$, which proves the claim. □

In light of lemma 3.1, in what follows, we may restrict attention to the space $\mathcal{K}_c(X)$.

Lemma 3.2. If $\succsim$ satisfies Continuity (Axiom 1(b)) and L-Independence (Axiom 2(a)), then there exists a continuous and affine function $\zeta : L \to \mathbb{R}$ such that ζ represents $\succsim|_L$, ie, for all $\ell, \ell' \in L$, $\ell \succsim \ell'$ if, and only if, $\zeta(\ell) \geq \zeta(\ell')$.

Proof. Independence and Continuity hold on L, so by the Expected Utility Theorem, the claim follows. □

Corollary 3.3. If $\succsim$ satisfies Axiom 1, there exist $\ell^\sharp, \ell_\sharp \in L$ such that $\ell^\sharp \succ \ell_\sharp$.

Proof. Consider $\ell^\sharp, \ell_\sharp \in L$ that exist by Lipschitz continuity (Axiom 1(c)). Set $x = y = \{\ell^\sharp\}$ and $\alpha = \frac{1}{2}$. Lipschitz continuity then implies $\ell^\sharp \succ \frac{1}{2}\ell^\sharp + \frac{1}{2}\ell_\sharp$. Similarly, let $x = y = \{\ell_\sharp\}$ and $\alpha = \frac{1}{2}$, so Lipschitz continuity implies $\frac{1}{2}\ell^\sharp + \frac{1}{2}\ell_\sharp \succ \ell_\sharp$. It follows immediately that $\ell^\sharp \succ \ell_\sharp$. □

Lemma 3.4. Given the function $\zeta : L \to \mathbb{R}$ from lemma 3.2 above, there exists $V : X \to \mathbb{R}$ such that

(a) $x \succsim y$ if, and only if, $V(x) \geq V(y)$ for all $x, y \in X$,
(b) for all $\ell \in L$, $V(\ell) = \zeta(\ell)$, and
(c) V is continuous.

Proof. By Corollary 3.3, $\ell^* \succ \ell_*$. First, consider the case where $x \in X$ is such that $\ell^* \succsim x \succsim \ell_*$. By Continuity (Axiom 1(b)), there exists $a \in [0, 1]$ such that $x \sim a\ell^* + (1-a)\ell_*$.

Define $V(x) := \zeta\big(a\ell^* + (1-a)\ell_*\big) = a\zeta(\ell^*) + (1-a)\zeta(\ell_*)$. It is easy to see that for all $\ell \in L$, $V(\ell) = \zeta(\ell)$.

Next, consider the case where $x \succ \ell^*$. By Continuity, for any $\ell \in L$, there exists $a \in [0,1]$ such that $ax + (1-a)\ell_* \sim \ell$. Now, set $V(x) = [V(\ell) - (1-a)V(\ell_*)]/a$.

To see that $V(x)$ is independent of the choice of ℓ, suppose $\ell' \in L$ and $a' \in [0,1]$ are such that $\ell \succsim \ell'$ and $a'x + (1-a')\ell_* \sim \ell'$, so that $V(x) = [V(\ell') - (1-a')V(\ell_*)]/a'$. Because $ax + (1-a)\ell_* \sim \ell$, for all $b \in [0,1]$, $b\big(ax + (1-a)\ell_*\big) + (1-b)\ell_* \sim b\ell + (1-b)\ell_*$. Now, choose b such that $b\ell + (1-b)\ell_* \sim \ell'$. Then, $b\big(ax + (1-a)\ell_*\big) + (1-b)\ell_* \sim \ell'$, which implies $ba = a'$. Using the fact that $V(\ell') = bV(\ell) + (1-b)V(\ell_*)$, we see that

$$
\begin{aligned}
V(x) &= \frac{V(\ell') - (1-a')V(\ell_*)}{a'} \\
&= \frac{\big[bV(\ell) + (1-b)V(\ell_*)\big] - (1-ba)V(\ell_*)}{ba} \\
&= \frac{V(\ell) - (1-a)V(\ell_*)}{a}
\end{aligned}
$$

which is independent of the choice of b, or equivalently, the choice of ℓ'.

We can deal with case where $\ell_* \succ x$ in a similar fashion. The continuity of V follows immediately from the continuity of $\succsim$ and from the continuity of ζ, which completes the proof. □

Lemma 3.5. If $tx + (1-t)\ell \succ ty + (1-t)\ell$ then $x \succ y$.

Proof. Suppose not. Then, by L-Independence, there are x, y, ℓ, and t such that $x \sim y$ and $tx + (1-t)\ell \succ ty + (1-t)\ell$. By Lipschitz Continuity (Axiom 1(c)), and because $d(x,x) = 0$, we have $t'x + (1-t')\ell^\sharp \succ t'x + (1-t')\ell_\sharp$ for all $t' > 0$. Observe that by Negative Transitivity of the strict relation $\succ$, it must be that for all t', either $t'x + (1-t')\ell^\sharp \succ x$ or $x \succ t'x + (1-t')\ell_\sharp$ holds, and the same for y. There are three cases to consider.

Case 1: For all $\varepsilon > 0$ there is $(1-t') < \varepsilon$ with $x \succ t'x + (1-t')\ell_\sharp$. Then, since $x \sim y$, L-Independence implies that $ty + (1-t)\ell \succ t\left(t'x + (1-t')\ell_\sharp\right) + (1-t)\ell$ for all such $(1-t') > 0$. At the same time, by continuity, we can pick $\left(1-\bar{t}\right) > 0$ small enough, such that by replacing x with $\bar{t}x + (1-\bar{t})\ell_\sharp$, $t\left(\bar{t}x + (1-\bar{t})\ell_\sharp\right) + (1-t)\ell \succ ty + (1-t)\ell$ still holds. Taking $\varepsilon \leq (1-\bar{t})$ establishes a contradiction.

Case 2: For all $\varepsilon > 0$ there is $(1-t') < \varepsilon$ with $t'y + (1-t')\ell^\sharp \succ y$. This case is analogous to case 1.

Case 3: There is $\varepsilon > 0$ such that for all $(1-t') < \varepsilon$, both $t'x + (1-t')\ell_\sharp \succsim x$ and $y \succsim t'y + (1-t')\ell^\sharp$. We claim that this case can never occur. To see this, first observe that by continuity, if $t'x + (1-t')\ell_\sharp \succsim x$ for all $(1-t') < \varepsilon$ then $\ell_\sharp \succsim x$; and if $y \succsim t'y + (1-t')\ell^\sharp \succsim x$ for all $(1-t') < \varepsilon$ then $y \succsim \ell^\sharp$. But then we have $y \succsim \ell^\sharp \succ \ell_\sharp \succsim x$, which contradicts the premise that $x \sim y$. □

Corollary 3.6. It follows immediately from L-Independence and Lemma 3.5 that $tx + (1-t)\ell \succ ty + (1-t)\ell$ if, and only if, $x \succ y$.

Lemma 3.7. $\ell \succ \ell'$ if, and only if, $tx + (1-t)\ell \succ tx + (1-t)\ell'$.

Proof. If $x \succ \ell_*$, by continuity there are $\alpha \in (0,1)$ and $\overline{\ell} \in L$ with $\alpha x + (1-\alpha)\ell_* \sim \overline{\ell}$. Applying Corollary 3.6 repeatedly yields that $\ell \succ \ell'$ if, and only if, $t'[\alpha x + (1-\alpha)\ell_*] + (1-t')\ell \sim t'\overline{\ell} + (1-t')\ell \succ t'\overline{\ell} + (1-t')\ell' \sim t'[\alpha x + (1-\alpha)\ell_*] + (1-t')\ell'$ for all $t' \in (0,1)$. Again by Corollary 3.6, and for $t' = \frac{t}{\alpha + t(1-\alpha)}$, this is equivalent to $tx + (1-t)\ell \succ tx + (1-t)\ell'$. The case where $\ell^* \succ x$ is similar and hence omitted. $\square$

Lemma 3.8. The function V defined in the proof of Lemma 3.4 has the following properties:

(a) V is monotone, ie, $V(x \cup y) \geq V(x)$ for all $x, y \in X$;

(b) V is L-affine, ie, for all $x \in X$, $\ell \in L$ and $a \in [0,1]$, $V\big(ax + (1-a)\ell\big) = aV(x) + (1-a)V(\ell)$;

(c) V is midpoint convex, ie, $V\left(\frac{1}{2}x_1 + \frac{1}{2}x_2\right) \leq \frac{1}{2}V(x_1) + \frac{1}{2}V(x_2)$;

(d) V is convex.

Proof. To ease notational burden, we shall assume only in this part of the proof, and without loss of generality, that $V(\ell^*) = 1$ while $V(\ell_*) = 0$. We prove the claims in turn.

(a) V represents $\succsim$, so it is clear that it is monotone.

(b) Let $x \in X$ and $\ell \in L$. Consider first the case where $\ell^* \succsim x \succsim \ell_*$. Then, there exists $\ell_x \in L$ such that $x \sim \ell_x$. Then, by L-Independence, for all $a \in (0,1]$, $ax + (1-a)\ell \sim a\ell_x + (1-a)\ell$. Therefore, $V\big(ax + (1-a)\ell\big) = V\big(a\ell_x + (1-a)\ell\big) = aV(\ell_x) + (1-a)V(\ell) = aV(x) + (1-a)V(\ell)$, as required.

Now consider the case where $x \succ \ell^*$, the case where $\ell_* \succ x$ being analogous. Because $\ell \succsim \ell_*$, Lemma 3.7 yields $t\ell_* + (1-t)\ell \succsim \ell_*$, and then, by Corollary 3.6, $tx + (1-t)\ell \succ \ell_*$. By continuity, there are $\alpha \in (0,1)$ and $\overline{\ell}$, such that $\ell^* \succ \alpha\,(tx + (1-t)\ell) + (1-\alpha)\ell_* \sim \overline{\ell} \succ \ell_*$. Further, let $\beta \in [0,1]$ be such that $\ell \sim \beta\ell^* + (1-\beta)\ell_*$ (so that $V(\ell) = \beta$), and let $\gamma \in (0,1)$ be such that $\overline{\ell} \sim \gamma\ell^* + (1-\gamma)\ell_*$. First, from Corollary 3.6 and the definition of V it is easy to verify that $V(tx + (1-t)\ell) = \frac{\gamma}{\alpha}$ (independent of whether $tx + (1-t)\ell \succsim \ell^*$ or not). Next, by Lemma 3.7, $tx + (1-t)\ell \sim tx + (1-t)\,(\beta\ell^* + (1-\beta)\ell_*)$. Then, by Corollary 3.6,

$$\alpha\,(tx + (1-t)\,(\beta\ell^* + (1-\beta)\ell_*)) + (1-\alpha)\ell_* \sim \gamma\ell^* + (1-\gamma)\ell_*$$

or

$$\alpha tx + \alpha(1-t)\beta\ell^* + [1 - \alpha t - \alpha(1-t)\beta]\,\ell_* \sim \gamma\ell^* + (1-\gamma)\ell_*$$

Because $x \succ \ell^*$, Corollary 3.6 and Lemma 3.7 further imply that $\alpha(1-t)(1-\beta) + (1-\alpha) > (1-\gamma)$ or $\gamma - \alpha\,(1-t)\,\beta > \alpha t > 0$. This implies that $\gamma > \alpha\,(1-t)\,\beta$. Corollary 3.6 then yields that

$$\frac{\alpha t}{D_1}x + \frac{1 - \alpha t - \alpha(1-t)\beta}{D_1}\ell_* \sim \frac{\gamma - \alpha(1-t)\beta}{D_1}\ell^* + \frac{1-\gamma}{D_1}\ell_*$$

where $D_1 = \gamma - \alpha(1-t)\beta + (1-\gamma) = 1 - \alpha(1-t)\beta$.

It follows that $1-\gamma < 1-\alpha t-\alpha(1-t)\beta$, and hence, again by Corollary 3.6,

$$\frac{\alpha t}{D_2}x+\frac{1-\alpha t-\alpha(1-t)\beta-(1-\gamma)}{D_2}\ell_* \sim \ell^*$$

where $D_2=\alpha t+1-\alpha t-\alpha(1-t)\beta-(1-\gamma)=\gamma-\alpha(1-t)\beta$.
Hence, $\frac{\alpha t}{\gamma-\alpha(1-t)\beta}x+\left[1-\frac{\alpha t}{\gamma-\alpha(1-t)\beta}\right]\ell_* \sim \ell^*$, so that $V(x)=\frac{\gamma-\alpha(1-t)\beta}{\alpha t}$. Putting everything together establishes the lemma, ie,

$$tV(x)+(1-t)V(\ell)=\frac{\gamma}{\alpha}=V(tx+(1-t)\ell)$$

(c) Suppose first that $x_1 \sim x_2$. Then, by Aversion to Randomization (Axiom 1 (e)), $x_1 \succsim \frac{1}{2}x_1+\frac{1}{2}x_2$, from which it follows immediately that $V(\frac{1}{2}x_1+\frac{1}{2}x_2)\le\frac{1}{2}V(x_1)+\frac{1}{2}V(x_2)$. Let us now suppose that $x_1 \succ x_2$ and consider the case where $\ell^* \succ x_1$. By continuity, there exists $\lambda\in(0,1)$ such that $y:=\lambda x_2+(1-\lambda)\ell^* \sim x_1$. Notice that because V is L-affine, $V(y)=\lambda V(x_2)+(1-\lambda)V(\ell^*)=V(x_1)$. Let $\bar{x}:=\frac{\lambda}{1+\lambda}x_1+\frac{1}{1+\lambda}y=\frac{2\lambda}{1+\lambda}(\frac{1}{2}x_1+\frac{1}{2}x_2)+\frac{1-\lambda}{1+\lambda}\ell^*$, so that $V(\bar{x})=\frac{2\lambda}{1+\lambda}V(\frac{1}{2}x_1+\frac{1}{2}x_2)+\frac{1-\lambda}{1+\lambda}V(\ell^*)$, where we have used the L-affinity of V. But notice also that $V(\bar{x})\le\frac{\lambda}{1+\lambda}V(x_1)+\frac{1}{1+\lambda}V(y)$ by Aversion to Randomization (Axiom 1 (e)) because $x_1\sim y$. We also have $\frac{\lambda}{1+\lambda}V(x_1)+\frac{1}{1+\lambda}V(y)=\frac{\lambda}{1+\lambda}\big(V(x_1)+V(x_2)\big)+\frac{1-\lambda}{1+\lambda}V(\ell^*)$. Substituting in the value of $V(\bar{x})$ obtained above, we see that $V(\frac{1}{2}x_1+\frac{1}{2}x_2)\le\frac{1}{2}V(x_1)+\frac{1}{2}V(x_2)$, as claimed.
Now consider the case where $x_1\succ x_2$ but $x_1\succ\ell^*$. Then, by continuity, there exists $a\in[0,1]$ such that $y=ax_1+(1-a)\ell_*\sim x_2$. Therefore, $V(y)=aV(x_1)+(1-a)V(\ell_*)=V(x_1)$. Set $\bar{x}=\frac{a}{1+a}x_2+\frac{1}{1+a}y=\frac{2a}{1+a}(\frac{1}{2}x_1+\frac{1}{2}x_2)+\frac{1-a}{1+a}\ell_*$. Then, using the L-affinity of V, we obtain $V(\bar{x})=\frac{2a}{1+a}V(\frac{1}{2}x_1+\frac{1}{2}x_2)+\frac{1-a}{1+a}V(\ell_*)$.
But notice that $x_2\sim y$, so that by Aversion to Randomization (Axiom 1 (e)), $V(\bar{x})\le\frac{a}{1+a}V(x_2)+\frac{1}{1+a}V(y)$. We also have $\frac{a}{1+a}V(x_1)+\frac{1}{1+a}V(y)=\frac{a}{1+a}\big(V(x_1)+V(x_2)\big)+\frac{1-a}{1+a}V(\ell_*)$. Substituting in the value of $V(\bar{x})$ obtained above, we see that $V(\frac{1}{2}x_1+\frac{1}{2}x_2)\le\frac{1}{2}V(x_1)+\frac{1}{2}V(x_2)$, as claimed.

(d) As noted above, V is continuous, and because it is midpoint convex, it is convex. □

Recall that V is Lipschitz if there exists a constant $K>0$ such that for all $x,y\in X$, $|V(x)-V(y)|\le Kd(x,y)$, where $d(\cdot,\cdot)$ is the metric on X.

Lemma 3.9. If $\succsim$ satisfies Lipschitz continuity (Axiom 1(c)) and is represented by a continuous and L-affine V, then V is Lipschitz. Conversely, if V is Lipschitz, non-trivial, L-affine, and represents $\succsim$, then it satisfies Lipschitz continuity.

Proof. Let $N>0$ be as given in Lipschitz continuity. Fix $\beta\in(0,1)$ such that $N\beta<1$. First consider the case where $x,y\in X$ are such that $0<d(x,y)\le\beta$ and let $\alpha=Nd(x,y)$. Then, by Lipschitz Continuity, $(1-\alpha)x+\alpha\ell^\sharp\succ(1-\alpha)y+\alpha\ell_\sharp$. By the L-affinity of V,

it follows that $V(y) - V(x) < \frac{\alpha}{1-\alpha}\big[V(\ell^\sharp) - V(\ell_\sharp)\big]$. But notice that $\alpha/N \leq \beta$, so setting $K = N/(1-N\beta)\big[V(\ell^\sharp) - V(\ell_\sharp)\big]$, we find that

$$\begin{aligned} V(y) - V(x) &< \frac{\alpha}{1-\alpha}\big[V(\ell^\sharp) - V(\ell_\sharp)\big] \\ &< \frac{N}{1-\alpha}\big[V(\ell^\sharp) - V(\ell_\sharp)\big]d(x,y) \\ &< Kd(x,y) \end{aligned}$$

We now follow Dekel et al. (2007) and remove the restriction on the x and y. For arbitrary $x, y \in X$, let $0 =: \lambda_0 < \lambda_1 < \cdots < \lambda_{J+1} = 1$ such that $(\lambda_{j+1} - \lambda_j)d(x,y) \leq \beta$ for all $j = 0, \ldots, J+1$. Define $x_j := \lambda_j x + (1-\lambda_j)y$, so $d(x_{j+1}, x_j) = (\lambda_{j+1} - \lambda_j)d(x,y) < \beta$. From the result established above, we see that $V(x_{j+1}) - V(x_j) \leq Kd(x_{j+1}, x_j) = K(\lambda_{j+1} - \lambda_j)d(x,y)$. Summing over j, we find $V(y) - V(x) \leq Kd(x,y)$. Interchanging the roles of x and y, it follows that $|V(x) - V(y)| \leq Kd(x,y)$, as claimed. The converse is as in Dekel et al. (2007) and is omitted. □

In sum, we have proven that (a) implies (b) in the following representation result.

Proposition 3.10. Let $\succsim$ be a binary relation. Then, the following are equivalent.

(a) $\succsim$ satisfies Basic Properties (Axiom 1) and L-Independence (Axiom 2(a)).

(b) There exists a function $V : X \to \mathbb{R}$ that represents $\succsim$ and is L-affine, Lipschitz Continuous, and convex. Moreover, any such representation of $\succsim$ is unique up to a positive affine transformation.

The proof that (b) implies (a) is standard and is omitted.

3.1.2. *Abstract Convex and Monotone Representation*

Every $f \in \mathcal{F}(\Delta(C \times X))$ is a product lottery of the form $f(1) \times \cdots \times f(n)$. A function $\mathfrak{u} \in \mathfrak{U}$ acts on $\mathcal{F}(\Delta(C \times X))$ as follows: $\mathfrak{u}(f) := \sum_i p_i \mathfrak{u}_i(f(i))$. For any $x \in \mathcal{K}_c(\mathcal{F}(\Delta(C \times X)))$, define its *support function* $H_x : \mathfrak{U} \to \mathbb{R}$ as $H_x(\mathfrak{u}) := \max_{f \in x} \mathfrak{u}(f)$. The *extended* support function of $x \in \mathcal{K}_c(\mathcal{F}(\Delta(C \times X)))$ is the unique extension of the support function H_x to $\mathrm{span}(\mathfrak{U})$ by positive homogeneity. Theorem 5.102 and Corollary 6.27 of Aliprantis and Border (1999) imply that a function defined on $\mathrm{span}(\mathfrak{U})$ is sublinear, norm continuous, and positively homogeneous if, and only if, it is the extended support function of some weak* closed, convex subset of $\mathcal{F}(\Delta(C \times X))$. Therefore, a function $H : \mathfrak{U} \to \mathbb{R}$ is a support function if its unique extension to $\mathrm{span}(\mathfrak{U})$ by positive homogeneity is sublinear and norm continuous.

Given a function $H : \mathfrak{U} \to \mathbb{R}$ whose extension to $\mathrm{span}(\mathfrak{U})$ by positive homogeneity is sublinear and norm continuous, we may define $x_H := \{f \in \mathrm{aff}\big(\mathcal{F}(\Delta(C \times X))\big) : \mathfrak{u}(f) \leq H(\mathfrak{u}) \text{ for all } \mathfrak{u} \in \mathfrak{U}\}$. Support functions enjoy the following duality: For any weak* compact, convex subset x of $\mathrm{aff}\big(\mathcal{F}(\Delta(C \times X))\big)$, $x_{H_x} = x$, and for any function H as defined above, $H_{x_H} = H$.

For weak* compact, convex subsets x and x' of X, support functions exhibit the following properties: (i) $x \subset x'$ if, and only if, $H_x \leq H_{x'}$, (ii) $H_{tx+(1-t)x'} = tH_x + (1-t)H_{x'}$ for all $t \in (0,1)$, (iii) $H_{x \cap x'} = H_x \wedge H_{x'}$, and (iv) $H_{\mathrm{ch}(x \cup x')} = H_x \vee H_{x'}$. (By Lemma 5.14 of Aliprantis and Border (1999), $\mathrm{ch}(x \cup x')$ is compact because x and x' are compact, which ensures that $H_{\mathrm{ch}(x \cup x')}$ is well defined.) Finally, observe that for $\ell^\dagger := \ell_i^\dagger \times \cdots \times \ell_n^\dagger$, $H_{\ell^\dagger} = \mathbf{0}$.

Proposition 3.11. Let $V : X \to \mathbb{R}$ be Lipschitz, convex, and L-affine. Then, there exists a minimal set $\mathfrak{M}$ of finite normal charges on $\mathfrak{U}$ so that V can be written as

$$[\bullet] \qquad V(x) = \max_{\mu \in \mathfrak{M}} \left[\int_{\mathfrak{U}} \max_{f \in x} \sum_i p_i \mathrm{u}_i(f(i)) \, \mathrm{d}\mu(p, \mathrm{u}) \right]$$

where the set $\mathfrak{M} \subset ba_{\mathrm{n}}(\mathfrak{U})$ is weak* compact and $\int_{\mathfrak{U}} \max_{f \in x} \sum_i p_i \mathrm{u}_i(f(i)) \, \mathrm{d}\mu(p, \mathrm{u})$ is independent of μ for all $x \in L$.[5] Moreover, for a dense set of points in X, there is a unique $\mu \in \mathfrak{M}$ that achieves the maximum in [•].

In Proposition 3.11 above, $ba_{\mathrm{n}}(\mathfrak{U})$ is the space of bounded additive (or finitely additive) measures (ie, charges) on $\mathfrak{U}$ that are also normal (ie, inner and outer regular). The last part of the proposition reflects the fact that V is linear on L. The set $\mathfrak{M}$ is minimal in the sense that if $\mathcal{N} \subset \mathfrak{M}$ is compact, then there exists $x \in X$ such that $V(x) > \max_{\mu \in \mathcal{N}} \left[\int_{\mathfrak{U}} \max_{f \in x} \sum_i p_i \mathrm{u}_i(f(i)) \, \mathrm{d}\mu(p, \mathrm{u}) \right]$.

Proof. By Lemma 3.1, for every $x \in \mathcal{K}(\mathcal{F}(\Delta(C \times X)))$, $V(x) = V(\mathrm{cch}(x))$. Therefore, we may restrict attention to convex menus.

Let $\Psi : \mathcal{K}_c(\mathcal{F}(\Delta(C \times X))) \to \mathbf{C}_b(\mathfrak{U})$ be the map that associates each compact, convex subset x of $\mathcal{F}(\Delta(C \times X))$ with its support function, $\Psi : x \mapsto H_x$. Note that Ψ is invertible. Moreover, Ψ is an isometry because $d(x, x') = \|H_x - H_{x'}\|_\infty$ for all $x, x' \in \mathcal{K}_c(\mathcal{F}(\Delta(C \times X)))$. Thus Ψ is an affine isometric embedding of $\mathcal{K}_c(\mathcal{F}(\Delta(C \times X)))$ in $\mathbf{C}_b(\mathfrak{U})$. Moreover, $\Psi(\{\ell^*\}) = \mathbf{0}$. In sum, $\Psi\big(\mathcal{K}_c(\mathcal{F}(\Delta(C \times X)))\big)$ is a compact and convex subset of $\mathbf{C}_b(\mathfrak{U})$ that contains the origin.

Let $\bar{V} : \Psi\big(\mathcal{K}_c(\mathcal{F}(\Delta(C \times X)))\big) \to \mathbb{R}$ be defined as follows: $\bar{V}(H) := V(x)$ where $H = H_x$ for some x. Because Ψ is injective, it follows that $\bar{V}$ is well defined. Thus, $\bar{V}$ is Lipschitz, convex, and $\Psi(L)$-affine. Recall that by definition, $V(\{\ell^*\}) = 0 = \bar{V}(H_{\{\ell^*\}})$, and $\Psi(\{\ell^*\}) = \mathbf{0}$. Therefore, $\bar{V}$ is positively homogeneous. Extending $\bar{V}$ to $\mathrm{cone}(\Psi(\mathcal{K}_c(\mathcal{F}(\Delta(C \times X)))))$ by positive homogeneity, it follows by Proposition 2.4 that $\bar{V}$ (and hence V) has the desired representation. □

Proposition 3.12. Let $V : \mathcal{K}(\mathcal{F}(\Delta(C \times X))) \to \mathbb{R}$ be as in [•]. Then, the following are equivalent.

(a) V is monotone, in the sense that $x \subset x'$ implies $V(x) \leq V(x')$.
(b) Every charge $\mu \in \mathfrak{M}$ is *positive*, ie, $\mu(E) \geq 0$ for all (Borel) measurable $E \subset \mathfrak{U}$.

(5) Recall that $ba_{\mathrm{n}}(\mathfrak{U})$ is the space of finite normal charges on $\mathfrak{U}$.

Proof. That (b) implies (a) is easy to see. That (a) implies (b) follows from Theorem S.2 of Ergin and Sarver (2010b) after observing that $\bar{V}$ (defined in the proof of 3.12) is monotone. We note that a similar statement is contained in the proof of Lemma 3.5 of Gilboa and Schmeidler (1989). □

The following corollary follows immediately from Lemma 2.5.

Corollary 3.13. Let $V : \mathcal{K}(\mathcal{F}(\Delta(C \times X))) \to \mathbb{R}$ have a representation as in [•]. Suppose $E \subset \mathcal{K}(\mathcal{F}(\Delta(C \times X)))$ is convex and $V|_E$ is linear. Then, there exists $\mu \in \mathfrak{M}$ such that $V(x) = \int_{\mathfrak{U}} \max_{f \in x} \sum_i p_i \mathfrak{u}_i(f(i))\, d\mu(p, \mathfrak{u})$ for all $x \in E$.

This establishes Theorem 3 of Section 3.1. Notice that by definition, V is (i) convex, (ii) Lipschitz continuous, and (iii) L-affine in the sense that for all $x \in X$, $\ell \in L$ and $t \in [0, 1]$, $V\big((1-t)x + t\ell\big) = (1-t)V(x) + tV(\ell)$. We shall use these properties in the sequel.

Each of the following subsections will introduce a new axiom which will, in turn, impose further restrictions on the set $\mathfrak{M}$, eventually leading us to the desired representation in Theorem 2.

3.2. Partitional Representation

In this section, we consider the representation in [♦] of $\succsim$ and impose Indifference to Incentivized Contingent Commitment (henceforth IICC, Axiom 4).

The main consequence of assuming IICC (Axiom 4) is that instead of considering arbitrary finitely additive measures $\mu \in \mathfrak{M}$ over $\mathfrak{U}$ in the representation [♦], we can replace each μ by a pair $(P, \mathfrak{u})$ along with a prior belief π_0 over S, where P is a partition of S and $\mathfrak{u} \in \mathbf{C}(C \times X)$.

Proposition 3.14. Consider a preference relation $\succsim$ on X, and suppose $V : X \to \mathbb{R}$ represents $\succsim$ and has the form in [♦]. Then, the following are equivalent:

(a) $\succsim$ satisfies IICC (Axiom 4).

(b) The function V has the form

[3.1]
$$V(x) = \max_{(P,\mathfrak{u}) \in \mathfrak{M}_p} \left[\sum_{J \in P} \Big(\max_{f \in x} \sum_{s \in J} \pi_0(s \mid J)\, \mathfrak{u}_s(f(s)) \Big) \pi_0(J) \right]$$

where $\mathfrak{M}_p$ is a collection of pairs $(P, \mathfrak{u})$ where P is a partition and $\mathfrak{u} = (\mathfrak{u}_s)_{s \in S}$ is a collection of state dependent (vN-M) utility functions on $C \times X$ with the property that for all $s \in S$, $\mathfrak{u}_s(\alpha) = \mathfrak{u}'_s(\alpha)$ for all $(P, \mathfrak{u}), (P', \mathfrak{u}') \in \mathfrak{M}_p$ and $\alpha \in \Delta(C \times L)$.

Notice that each partition P along with a prior π_0 is equivalent to a posterior belief over S, while $\mathfrak{u}$ corresponds to a Dirac measure over $\mathfrak{U}$, both of which are countably additive.

Thus, an essential part of the proof of Proposition 3.14 is to show that IICC (Axiom 4) allows us to replace each $\mu \in \mathfrak{M}$ by a countably additive measure without affecting the representation. The proof is lengthy precisely due to the complications that arise from dealing with $\mu \in \mathfrak{M}$ in [♦] that are finitely additive. If we knew beforehand that each μ was countably additive, the proof would simply formalize the intuition behind IICC (Axiom 4) and be considerably shorter. The rest of this section proves Proposition 3.14. We begin with some lemmas.

Let $\tilde{\Xi}_x := \{\xi \in \Xi_x : \mathcal{I}(\xi) \sim x\}$. By IICC (Axiom 4), $\tilde{\Xi}_x$ is non-empty. It follows from the definition of $\tilde{\Xi}_x$ that for each $\xi \in \tilde{\Xi}_x$, there exist $f_1, \ldots, f_m \in x$ such that for each $i = 1, \ldots, m$, $f_i = \xi(s)$ for some $s \in S$. The collection $\{f_1, \ldots, f_m\}$ denotes a set of *generators* of the set x according to ξ. We shall also say that $\{f_1, \ldots, f_m\}$ *generates* x according to ξ.

Lemma 3.15. For $x \in X^*$, let $\{f_1, \ldots, f_m\}$ generate x according to $\xi \in \tilde{\Xi}_x$. Then, $x \sim \{f_1, \ldots, f_m\}$.

Proof. Notice that

$$
\begin{aligned}
x &\succsim \{f_1, \ldots, f_m\} && \text{by Monotonicity (Axiom 1(d))} \\
&\succsim \mathcal{I}(\xi) && \text{by IICC(a) and Continuity} \\
&\sim x && \text{by IICC(b)}
\end{aligned}
$$

which establishes the claim. □

Definition 3.16. A menu x is *nice* if $x \in X^*$ and there is a unique $\xi \in \tilde{\Xi}_x$. X_0 denotes the space of nice menus. A menu x is *minimal* if $x \succ x \setminus \{f\}$ for all $f \in x$.

Let x be a nice menu, $\xi \in \tilde{\Xi}_x$, and $f_1, \ldots, f_m$ the corresponding generators of x. Each such ξ induces a partition $J_1, \ldots, J_m$ of S wherein $\xi(s) = f_k$ if, and only if, $s \in J_k$. In this case, we shall say that f_k is *active* in state $s \in J_k$, so that J_k denotes all the states where f_k is active.

Proposition 3.17. The space X_0 is dense in X.

Proof. It is easy to see that the space X^* is dense in X. Therefore, it will suffice to show that X_0 is dense in X^*. For any $x \in X^*$, it is easy to see that IICC (Axiom 4) implies the existence of a minimal set of generators, $\{f_1, \ldots, f_m\}$. Let $x_\varepsilon := (1-\varepsilon)\,x + \varepsilon \ell_*$ and $y := \{f_1, \ldots, f_m\} \cup x_\varepsilon$. By Monotonicity (Axiom 1(d)), $y \succsim x$. Obviously $d(y, x) \to 0$ as $\varepsilon \to 0$. Because $x \in X^*$ and $\varepsilon > 0$ are arbitrary, it suffices to establish that (some perturbation of) $y \in X_0$.

Because $x \in X^*$, we also have $x_\varepsilon \in X^*$ and, because $\{f_1, \ldots, f_m\} \subset x$, also $\{f_1, \ldots, f_m\} \in X^*$, which then implies $y \in X^*$. We now show that there must be a unique $\xi \in \tilde{\Xi}_y$ (perhaps after further perturbing y) to establish the proposition.

Suppose there is $\xi \in \tilde{\Xi}_y$ with generator set

$$\left\{f_1', \ldots, f_j', (1-\varepsilon)\, f_{j+1}' + \varepsilon \ell_*, \ldots, (1-\varepsilon)\, f_k' + \varepsilon \ell_*\right\} \sim y$$

(indifference follows from Lemma 3.15) where $f_a' \in x$ for all $a \in \{1, \ldots, k\}$. Consider, now, $\mathcal{I}(\xi)$ and note that it can be generated inductively from $y_0 := \left\{f_1', \ldots, f_k'\right\}$ as follows, where the induction is over the set of states $S = \{s_1, \ldots, s_n\}$. For $i \in \{1, \ldots, n\}$, let $e_i : y \to [0, 1]$ be defined by

$$e_i\,(f) := \begin{cases} 0 & \text{if } f = \xi(s_i) \text{ and } f \in \{f_1, \ldots, f_m\} \\ \varepsilon & \text{if } (1-\varepsilon) f + \varepsilon \ell_* = \xi(s_i) \notin \{f_1, \ldots, f_m\} \\ 1 & \text{otherwise} \end{cases}$$

Given y_i, let

$$y_{i+1} := y_i \oplus_{(e_{i+1}, s_{i+1})} \ell_*$$

Observe that, indeed, $y_n = \mathcal{I}(\xi)$. Note, further, that by IICC (part a) and Continuity (Axiom 1(b)), $y_i \succsim y_{i+1}$, with $y_i \succ y_{i+1}$ if $\xi\,(s_i) \in x_\varepsilon$. Suppose, now, that $k > j$. In that case, $y_0 \succ y_n = \mathcal{I}(\xi) \sim y$. By Monotonicity, $x \succsim y_0$, and hence $x \succ y$, which contradicts the observation above that $y \succsim x$. Therefore, $m = j$. But then $y \succsim x$ and the minimality of $\{f_1, \ldots, f_m\}$ implies that the generator set that corresponds to ξ must be $\{f_1, \ldots, f_m\}$. Because ξ was chosen arbitrarily among the $\xi \in \tilde{\Xi}_y$, any such ξ must have generator set $\{f_1, \ldots, f_m\}$.

Suppose, then, that there are $\xi, \xi' \in \tilde{\Xi}_y$ with the same generator set $\{f_1, \ldots, f_m\}$, and $f_b = \xi\,(s) \neq \xi'\,(s)$ for some $s \in S$ and $b \in \{1, \ldots, m\}$. Let

$$\hat{f}_b(s') := \begin{cases} f_b(s') & s' \neq s \\ (1-t)\, f_b + t \ell_* & s' = s \end{cases}$$

Note that, by Continuity, for $t > 0$ small enough, $\{f_1, \ldots, \hat{f}_b, \ldots, f_m\}$ remains the unique generator set for $\hat{y} := [y \backslash \{f_b\}] \cup \{\hat{f}_b\}$. Let $\hat{\xi} \in \Xi_{\hat{y}}$ be the contingent plan with

$$\hat{\xi}(s') := \begin{cases} \hat{f}_b(s') & \xi(s') = f_b \\ \xi(s') & \text{otherwise} \end{cases}$$

and analogously for $\hat{\xi}'$ and ξ'. Then IICC (part a) implies that $y \succ \mathcal{I}(\hat{\xi})$. At the same time $\mathcal{I}(\hat{\xi}') = \mathcal{I}\,(\xi') \sim y$. It is also clear that, for $t > 0$ small enough and by Continuity, for any $\xi'' \in \Xi_y$ with $\mathcal{I}\,(\xi'') \nsim y$, also $\mathcal{I}\left(\hat{\xi}''\right) \nsim \hat{y}$,where $\hat{\xi}''$ is again defined analogously. Hence, $\tilde{\Xi}_{\hat{y}}$ has at least one element less than $\tilde{\Xi}_y$. In finitely many steps we arrive at an (arbitrarily small) perturbation of y that is in X_0. This establishes the proposition. □

A (static) *strategy* for DM at a menu x given $\mu \in \mathfrak{M}$ is a mapping $\zeta_x^\mu : \mathfrak{U} \to x$. The strategy ζ_x^μ is *partitional* if there is a finite partition (E_i) of $\mathfrak{U}$, such that for each E_i there exists $f_i \in x$ with $\zeta_x^\mu(E_i) = f_i$. The value of this strategy is

$$V(x, \mu, \zeta_x^\mu) = \sum_i \int_{E_i} \sum_s p_s \mathrm{u}_s(f_i(s)) \, \mathrm{d}\mu(p, \mathrm{u})$$

A strategy ζ_x^μ is *optimal* at x if there is no other strategy that gives a higher payoff. A partitional optimal strategy ζ_x^μ is an optimal strategy that is partitional, ie, one where

$$\begin{aligned} V(x, \mu, \zeta_x^\mu) &= \sum_i \int_{E_i} \langle (p, \mathrm{u}), f_i \rangle \, \mathrm{d}\mu(p, \mathrm{u}) \\ &= \max_{\mu \in \mathfrak{M}} \left[\int_{\mathfrak{U}} \max_{f \in x} \langle (p, \mathrm{u}), f \rangle \, \mathrm{d}\mu(p, \mathrm{u}) \right] \end{aligned}$$

where $\langle (p, \mathrm{u}), f \rangle = \sum_s p_s \mathrm{u}_s(f_i(s))$. Notice that if a partitional strategy ζ_x^μ is optimal at x and if f_i is the act chosen in the cell E_i, we must necessarily have, for all $(p, \mathrm{u}) \in E_i$, $\langle (p, \mathrm{u}), f_i \rangle \geq \langle (p, \mathrm{u}), f \rangle$ for all $f \in x$.

In the sequel, ζ_x^μ denotes an optimal partitional strategy when one exists. It is easy to see that for a finite x, an optimal strategy is always partitional, though there may be many such strategies that are optimal. If ζ_x^μ induces the partition (E_i), we refer to (E_i) as an optimal partition for μ at x.

Definition 3.18. Let $\{f_1, \ldots, f_m\}$ be a set of generators of x, and let $(E_i)_{i=1}^m$ be a partition of $\mathfrak{U}$. Then, (E_i) is a *partition of* $\mathfrak{U}$ *consistent with* $\{f_1, \ldots, f_m\}$ if $(p, \mathrm{u}) \in E_i$ implies $\langle (p, \mathrm{u}), f_i \rangle \geq \left\langle (p, \mathrm{u}), f_j \right\rangle$ for all $j = 1, \ldots, m$.

Intuitively, a partition (E_i) of $\mathfrak{U}$ is consistent with $\{f_1, \ldots, f_m\}$ if there is some optimal μ such that it is optimal to choose f_i when $(p, \mathrm{u}) \in E_i$.

For each $\mu \in \mathfrak{M}$, let $V(x, \mu) := \int_{\mathfrak{U}} \max_{f \in x} \sum_s p_s \mathrm{u}_s(f(s)) \, \mathrm{d}\mu(p, \mathrm{u})$ be the utility from choosing the measure μ. Let $\Upsilon : X \rightrightarrows \mathfrak{M}$ be the mapping selecting the maximizing μ for each x; that is, $\Upsilon(x) := \arg\max_{\mu \in \mathfrak{M}} V(x, \mu)$. It is easy to see that $V(x, \mu)$ is continuous in μ, so it follows that Υ is a correspondence that is closed valued. The following lemma implies that finite menus always have consistent partitions.

Lemma 3.19. Let $x \in X$ be finite and suppose $\{f_1, \ldots, f_m\}$ is a set of generators of x. Then, $\mu \in \Upsilon(\{f_1, \ldots, f_m\})$ implies $\mu \in \Upsilon(x)$.

Proof. Consider the following string of inequalities:

$$\begin{aligned} V(x) &= V(\{f_1, \ldots, f_m\}) && \text{because } \{f_1, \ldots, f_m\} \text{ generates } x \\ &= V(\{f_1, \ldots, f_m\}, \mu) && \text{definition of } \mu \\ &\leq V(x, \mu) && V(\cdot, \mu) \text{ is monotone} \\ &\leq V(x) && \text{definition of } V \end{aligned}$$

which proves that $\mu \in \Upsilon(x)$, as claimed. □

Lemma 3.20. Let x be finite. For any $\ell \in L$ and $\varepsilon > 0$, (i) $\Upsilon(x) = \Upsilon\big((1-\varepsilon)x + \varepsilon\ell\big)$, (ii) if x is nice, then $(1-\varepsilon)x + \varepsilon\ell$ is also nice, and (iii) if $\mu \in \Upsilon(x)$ and (E_i) is an optimal partition for μ at x, then it is also an optimal partition for μ at $(1-\varepsilon)x + \varepsilon\ell$.

Proof. Let x be finite and $\mu \in \Upsilon(x)$. Then, $V(x) = V(x,\mu) \geq V(x,\mu')$ for all $\mu' \in \mathfrak{M}$. We also have

$$\begin{aligned} V\big((1-\varepsilon)x + \varepsilon\ell, \mu\big) &= (1-\varepsilon)V(x,\mu) + \varepsilon V(\ell,\mu) \\ &\geq (1-\varepsilon)V(x,\mu') + \varepsilon V(\ell,\mu) \\ &= (1-\varepsilon)V(x,\mu') + \varepsilon V(\ell,\mu') \\ &= V\big((1-\varepsilon)x + \varepsilon\ell, \mu'\big) \end{aligned}$$

where the inequality uses the fact that $V(x,\mu) \geq V(x,\mu')$ and the second equality follows because $V(\ell,\mu) = V(\ell,\mu')$ for all $\mu,\mu' \in \mathfrak{M}$ and $\ell \in L$. This proves part (i). Part (ii) follows immediately from the definition.

To see part (iii), let ζ_x^μ be a partitional optimal strategy with optimal partition (E_i). Then,

$$V(x) = V(x,\mu,\zeta_x^\mu) = \sum_i \int_{E_i} \langle (p,\mathrm{u}), f_i\rangle \, \mathrm{d}\mu(p,\mathrm{u})$$

For the menu $(1-\varepsilon)x + \varepsilon\ell$, consider the strategy $\zeta^\mu_{(1-\varepsilon)x+\varepsilon\ell}(E_i) = (1-\varepsilon)f_i + \varepsilon\ell$. Then,

$$\begin{aligned} &V\big((1-\varepsilon)x + \varepsilon\ell, \mu, \zeta^\mu_{(1-\varepsilon)x+\varepsilon\ell}\big) \\ &\quad= (1-\varepsilon)\sum_i \int_{E_i} \langle (p,\mathrm{u}), f_i\rangle \, \mathrm{d}\mu(p,\mathrm{u}) + \varepsilon \sum_i \int_{E_i} \langle (p,\mathrm{u}), \ell\rangle \, \mathrm{d}\mu(p,\mathrm{u}) \\ &\quad= (1-\varepsilon)V(x) + \varepsilon V(\ell) \\ &\quad\geq V\big((1-\varepsilon)x + \varepsilon\ell, \mu'\big) \end{aligned}$$

for all $\mu' \in \mathfrak{M}$ where the second equality follows from part (i). This proves that $\zeta^\mu_{(1-\varepsilon)x+\varepsilon\ell}$ is a partitional optimal strategy at the menu x given the optimal $\mu \in \mathfrak{M}$ and completes the proof. □

For a fixed partition (E_i) of $\mathfrak{U}$, $\mu \in \mathfrak{M}$, and $s \in S$, consider the map

$$(\mu, E_i, s) \mapsto \int_{E_i} p_s \mathrm{u}_s(\cdot)\, \mathrm{d}\mu(p,\mathrm{u})$$

Each tuple (μ, E_i, s) induces a continuous and linear preference functional $\int_{E_i} p_s \mathrm{u}_s(\cdot)\, \mathrm{d}\mu(p,\mathrm{u})$ on $\Delta(C \times X)$. By the Expected Utility Theorem, this linear functional has a vN-M utility representation which we denote by $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\cdot)$, where $\|\bar{\mathrm{u}}_{i,s}\|_\infty = 1$. Thus, for all $\alpha \in \Delta(C \times X)$, we have

$$\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\alpha) = \int_{E_i} p(s)\mathrm{u}_s(\alpha)\, \mathrm{d}\mu(p,\mathrm{u})$$

Then, $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}$ is a *local EU* representation of μ on E_i for state s. We do not index $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}$ by the relevant (E_i) and μ because these should be clear from the context.

Definition 3.21. Let $\mu \in \mathfrak{M}$ and (E_i) a partition of $\mathfrak{U}$. Then,

- A measure μ is *Type Ia* on E_i in state s if $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s} = \mathbf{0}$, ie, if $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}$ is trivial.
- A measure μ is *Type Ib* on E_i in state s if $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}$ is non-trivial, $\bar{p}_i(s)\bar{\mathrm{u}}_s$ is constant on $\Delta(C \times L)$, and ℓ_* maximizes $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}$ on $\Delta(C \times X)$.
- A measure μ is *Type IIa* on E_i in state s if $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}$ is non-trivial and not constant on $\Delta(C \times L)$.
- A measure μ is *Type IIb* on E_i in state s if $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}$ is non-trivial, constant on $\Delta(C \times L)$, and there exists $\alpha \in \Delta(C \times X)$ such that $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\alpha) > \bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\beta)$ for some (and hence all) $\beta \in \Delta(C \times L)$.

It is easy to see that the above taxonomy of measures is both mutually exclusive and exhaustive. Analogous to the definition in Section C.1 of DKS (and abusing notation), for any $\alpha \in \Delta(C \times X)$ we define

$$\big(f \oplus_{\varepsilon,s} \alpha\big)(s') := \begin{cases} (1-\varepsilon) f(s) + \varepsilon\alpha & \text{if } s' = s \\ f(s) & \text{otherwise} \end{cases}$$

Lemma 3.22. Let x be a finite menu, $\mu \in \Upsilon(x)$, and suppose there is a partitional optimal strategy ζ_x^μ with optimal partition (E_i), where $\zeta_x^\mu(E_i) = f_i \in x$. Suppose μ is Type II (a or b) on some E_i in state $s \in S$ and there exists $\alpha \in \Delta(C \times X)$ such that

$$\int_{E_i} p(s)\mathrm{u}_s\big(\alpha - f_i(s)\big)\, \mathrm{d}\mu(p, \mathrm{u}) > 0$$

Then, the menu $z := x \setminus \{f_i\} \cup \{f_i \oplus_{\varepsilon,s} \alpha\}$ is such that $V(z) > V(x)$ for all $\varepsilon > 0$.

Proof. Let $\mu \in \Upsilon(x)$ so that $V(x) = V(x, \mu)$. If

$$\int_{E_i} p(s)\mathrm{u}_s\big(\alpha - f_i(s)\big)\, \mathrm{d}\mu(p, \mathrm{u}) > 0$$

then it must necessarily be that $\mu(E_i) > 0$. The measure μ and the set E_i induce the functional

$$V_i(x, \mu, E_i) := \int_{E_i} \max_{f \in x} \sum_s p(s)\mathrm{u}_s(f(s))\, \mathrm{d}\mu(p, \mathrm{u})$$

on X. Let V_i^0 denote the restriction of V_i to $\mathscr{F}(\Delta(C \times X))$. By construction,

$$V_i^0(f) = \int_{E_i} \sum_s p(s)\mathrm{u}_s(f(s))\, \mathrm{d}\mu(p, \mathrm{u})$$

and because $\mu(E_i) > 0$, V_i^0 is non-trivial. By hypothesis, we have $V_i^0(f \oplus_{\varepsilon,s} \alpha) > V_i^0(f_i)$.

Consider the menu z and the strategy which entails the choice of f_j for $(p, \mathrm{u}) \in E_j$ when $j \neq i$, and the choice of $f_i \oplus_{\varepsilon,s} \alpha$ when $(p, \mathrm{u}) \in E_i$. This strategy delivers utility bounded above by $V(z,\mu)$, ie,

$$\begin{aligned}V(z,\mu) &\geq \sum_{j\neq i}\left[\int_{E_j}\sum_s p(s)\mathrm{u}_s f_j(s))\,\mathrm{d}\mu(p,\mathrm{u})\right] + \int_{E_i}\sum_s p(s)\mathrm{u}_s(f_i(s))\,\mathrm{d}\mu(p,\mathrm{u}) \\ &\qquad + \varepsilon\int_{E_i} p(s)\mathrm{u}_s(\alpha - f_i(s))\,\mathrm{d}\mu(p,\mathrm{u}) \\ &= V(x) + \varepsilon\int_{E_i} p(s)\mathrm{u}_s(\alpha - f_i(s))\,\mathrm{d}\mu(p,\mathrm{u}) \\ &> V(x)\end{aligned}$$

because $\int_{E_i} p(s)\mathrm{u}_s(\alpha - f_i(s))\,\mathrm{d}\mu(p,\mathrm{u}) > 0$ by hypothesis. Noting that $V(z) \geq V(z,\mu)$ by the definition of V completes the proof. □

Let $\mathfrak{M}_0 := \{\Upsilon(\{f_1,\dots,f_m\}) : \{f_1,\dots,f_m\}$ generates x for some $x \in X\}$. It follows from Lemma 3.19 that for all finite x,

$$\max_{\mu\in\mathfrak{M}_0} V(x,\mu) = \max_{\mu\in\mathfrak{M}} V(x,\mu)$$

In what follows, we shall restrict attention to finite menus and, therefore, it suffices to consider the set $\mathfrak{M}_0$. Let $\Upsilon_0 : X_0 \rightrightarrows \mathfrak{M}_0$ be defined as $\Upsilon_0(x) = \Upsilon(x) \cap \mathfrak{M}_0$.

Lemma 3.23. Let $x_0 := \{f_1,\dots,f_m\}$ be the generator set for some nice menu x, and suppose $\mu \in \Upsilon(x_0)$. Let J_i denote the states where f_i is active, and also let (E_i) represent an optimal partitional strategy (for μ) at x so that act f_i is chosen in the cell E_i. Then, μ is not Type II (a or b) at E_i in state s for all $i = 1,\dots,m$ and $s \in J_i^c$.

Proof. Let $\mu \in \Upsilon(x_0)$ so that $V(x) = V(x_0) = V(x_0,\mu)$ and suppose μ is Type II (a or b) at E_i in state $s \in J_i^c$. Note also that because x is nice, there is a unique $\xi \in \varXi_x$ such that $x \sim \mathcal{J}(\xi)$, and the generator of x is unique.

Case 1: First consider the case where $f_i(s)$ is not a maximizer for $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\cdot)$ on $\Delta(C\times X)$. Let f_i^* be the act such that (i) $f_i^*(s') = f_i(s')$ for all $s' \neq s$, and (ii) $f_i^*(s)$ maximizes $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\cdot)$ on $\Delta(C\times X)$, so that $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}\big(f_i^*(s)\big) > \bar{p}_i(s)\bar{\mathrm{u}}_{i,s}\big(f_i(s)\big)$. An act satisfying (ii) exists because μ is Type II at E_i in state s.

Now, consider the menu $x_{i,\varepsilon} := \{f_1,\dots,(1-\varepsilon)f_i + \varepsilon f_i^*,\dots,f_m\}$. By Lemma 3.22, $V(x_{i,\varepsilon}) > V(x)$ for all $\varepsilon > 0$. Notice also that $x_{i,\varepsilon} \to x$ as $\varepsilon \to 0$.

For any $\varepsilon > 0$, consider $\varXi_{x_{i,\varepsilon}}$, and notice that the set-valued map $\varepsilon \mapsto \varXi_{x_{i,\varepsilon}}$ is a continuous, closed, and compact valued correspondence. By IICC (Axiom 4), there exists $\xi \in \tilde{\varXi}_{x_{i,\varepsilon}}$. Consider the maximization problem (parametrized by ε)

$$\textbf{[P1]} \qquad W(\varepsilon) := \max V\big(\mathcal{J}(\xi)\big) \qquad \text{s.t. } \xi \in \varXi_{x_{i,\varepsilon}}$$

Notice that $W(0) = V(x)$ and that because $\varXi_{i,\varepsilon}$ is finite, a solution to [**P1**] always exists. We claim that for any $\varepsilon > 0$, the value of problem [**P1**] is precisely the value of $x_{i,\varepsilon}$, ie, $W(\varepsilon) = V(x_{i,\varepsilon})$.

To see this, notice that from the proof of Lemma 3.15, it follows that $V(x_{i,\varepsilon}) \geq V(\mathcal{J}(\xi))$ for all $\xi \in \Xi_{i,\varepsilon}$. By IICC (Axiom 4), there exists $\xi \in \tilde{\Xi}_{x_{i,\varepsilon}}$ such that $V(\mathcal{J}(\xi)) = V(x_{i,\varepsilon})$. Therefore, $W(\varepsilon) \geq V(x_{i,\varepsilon})$. Combining the two inequalities establishes that $W(\varepsilon) = V(x_{i,\varepsilon})$ for all $\varepsilon > 0$.

By the Theorem of the Maximum — see for instance, Ok (2007, p306) — W is continuous in ε. The Theorem of the Maximum also implies that the maximizer correspondence is upper hemicontinuous, and therefore for any ξ^*_ε that is optimal for the problem [**P1**], the limit $\xi^*_0 := \lim_{\varepsilon \to 0} \xi^*_\varepsilon$ is also a maximizer. (The limit always exists because $\Xi_{x_{i,\varepsilon}}$ is a continuous, closed, and compact valued correspondence.) The continuity of W then implies that $W(0) = V(\mathcal{J}(\xi^*_0))$.

There are two possibilities now. The first is that for all $\varepsilon^\circ > 0$, there exists $\varepsilon \in (0, \varepsilon^\circ)$ such that $\xi^*_\varepsilon(s) = (1-\varepsilon) f_i + \varepsilon f^*_i$ is active in state s. Because $\xi^*_0 = \lim_{\varepsilon \to 0} \xi^*_\varepsilon$, it follows that $\xi^*_0(s) = f_i$, ie, f_i is active in state s. In other words, $\xi^*_0 \neq \xi$. But we have already established that $W(0) = V(x) = V(\mathcal{J}(\xi^*_0))$, which contradicts the assumption that x is nice, which rules out this first possibility.

The other possibility is that there exists an $\varepsilon_\circ > 0$ such that for all $\varepsilon < \varepsilon_\circ$, the act $(1-\varepsilon) f_i + \varepsilon f^*_i$ is *in*active in every such state $s \in J^c_i$, ie, $\xi^*_\varepsilon(s) \neq (1-\varepsilon) f_i + \varepsilon f^*_i$. In this case, for all $\varepsilon < \varepsilon_\circ$, we have $\xi^*_0 = \xi^*_\varepsilon$. Because x is nice, it must necessarily be that $\xi^*_0 = \xi$. This implies that for all such ε, $V(x_{i,\varepsilon}) = W(\varepsilon) = W(0) = V(x)$. But this contradicts our earlier observation (which follows from Lemma 3.22) that $V(x_{i,\varepsilon}) > V(x)$ if μ is Type II at E_i in state s whenever f_i is active in state $s \in J_i$. This contradiction rules out the second possibility, and completes the proof of the first case.

Case 2: Suppose that $f_i(s)$ is a maximizer for $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\cdot)$ on $\Delta(C \times X)$. If μ is of Type IIa on E_i in state $s \in J^c_i$, then $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\cdot)$ is not constant on $\Delta(C \times L)$. If μ is of Type IIb on E_i in state $s \in J^c_i$, then $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\cdot)$ is constant on $\Delta(C \times L)$. However, in either case, there exists $\ell \in L$ such that $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(f_i(s)) > \bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\ell(s))$. (Such an ℓ exists because $f_i(s)$ is a maximizer of $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\cdot)$ and by hypothesis that μ is of Type II, there exists some $\beta \in \Delta(C \times L)$ that is *not* a maximizer.)

Consider the menu $\frac{1}{2}x + \frac{1}{2}\ell$. By Lemma 3.20, we see that $\mu \in \Upsilon(x)$ implies $\mu \in \Upsilon(\frac{1}{2}x + \frac{1}{2}\ell)$. Because x is nice, x_0, which satisfies $V(x_0) = V(x)$, is the unique generator set of x. L-Independence now implies that $V(\frac{1}{2}x_0 + \frac{1}{2}\ell) = V(\frac{1}{2}x + \frac{1}{2}\ell)$. Moreover, Lemma 3.20 says that $\frac{1}{2}x + \frac{1}{2}\ell$ is nice. It follows immediately that $\frac{1}{2}x_0 + \frac{1}{2}\ell$ is a generator set for $\frac{1}{2}x + \frac{1}{2}\ell$.

Now consider the nice menu $\frac{1}{2}x + \frac{1}{2}\ell$ with generator $\frac{1}{2}x_0 + \frac{1}{2}\ell$, and let $\mu \in \Upsilon(\frac{1}{2}x_0 + \frac{1}{2}\ell)$. By construction, $\frac{1}{2}f_i(s) + \frac{1}{2}\ell(s)$ is not a maximizer of $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}$ on $\Delta(C \times X)$ (although $f_i(s)$ is), which means that we now satisfy the hypotheses of Case 1. Lemma 3.20 ensures that $\Upsilon(\frac{1}{2}x + \frac{1}{2}\ell) \cap \Upsilon(x) \neq \varnothing$ and that a partitional optimal strategy at x is also optimal at $\frac{1}{2}x + \frac{1}{2}\ell$. These facts allow us to establish that even in this case, μ cannot be of Type II, which completes the proof. □

Let x be nice and let $\mu \in \Upsilon_0(x)$. Let $(E^{\mu,x}_i)$ be the partition induced by an optimal strategy (for instance, one coming from the generators of x) given μ and consider the mapping

$$(\mu, E_i^{\mu,x}, s) \mapsto \bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(\cdot) = \int_{E_i^{\mu,x}} p(s)\mathfrak{u}_s(\cdot)\, \mathrm{d}\mu(p, \mathfrak{u})$$

Let $\{f_1, \ldots, f_k\}$ be the unique generator set of x, and let J_i denote the set of states where f_i is active so (J_i) is a partition of S. Now define

$$\begin{aligned} \gamma^i_{\mu,x} &:= \sum_{s \in J_i} \bar{p}_i(s) \\ p_i(s) &:= \begin{cases} \bar{p}_i(s)/\gamma^i_{\mu,x} & \text{if } s \in J_i \\ 0 & \text{otherwise} \end{cases} \\ \hat{\mathfrak{u}}_s &:= \gamma^i_{\mu,x}\bar{\mathfrak{u}}_{i,s} \qquad \text{where } i \text{ is such that } s \in J_i \end{aligned} \qquad [\clubsuit]$$

and let

$$\hat{\mathfrak{M}} := \{\hat{\mu} \in \Delta(\mathfrak{U}) : \operatorname{supp}(\hat{\mu}) = \{(p_i, \hat{\mathfrak{u}}) : i = 1, \ldots, k \text{ where } k \leq n = |S|\}\}$$

Note that $\gamma^i_{\mu,x} \neq 0$ so that p_i is well defined. To see this, suppose that $\gamma^i_{\mu,x} = 0$. Then, $\bar{p}_i(s) = 0$ for all $s \in J_i$. This implies that $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(f) = 0$ for all acts f, which implies that $\{f_1, \ldots, f_k\} \sim \{f_1, \ldots, f_k\} \setminus \{f_i\}$. That is, we can drop the act f_i from the set $\{f_1, \ldots, f_k\}$ without any loss in utility, contradicting the assumption that $\{f_1, \ldots, f_k\}$ is the unique generator set of x.

Consider the mapping

$$\mathfrak{D}(\mu, x, (E_i^{\mu,x})) \mapsto \hat{\mu} \in \hat{\mathfrak{M}}$$

where $\operatorname{supp} \hat{\mu} = \{(p_i, \hat{\mathfrak{u}}) : i = 1, \ldots, k\}$, p_i for $i = 1, \ldots, k$ and $\hat{\mathfrak{u}}$ are defined in [♣], and $\hat{\mu}$ itself is defined as

$$\hat{\mu}\big((p_i, \hat{\mathfrak{u}})\big) = \mu\big(E_i^{\mu,x}\big)$$

Let $\hat{\mathfrak{M}}_p \subset \hat{\mathfrak{M}}$ be the image of $\mathfrak{D}$. (The domain of $\mathfrak{D}$ is easily defined, but notationally cumbersome, and because omitting it will not cause any confusion in the sequel, we refrain from a formal definition.)

A collection of probability measures $\{p_1, \ldots, p_k\}$ on S (so each $p_i \in \Delta(S)$) forms a *partitional system* if (i) for all $s \in S$, $p_i(s) > 0$ implies $p_j(s) = 0$ for all $j \neq i$, and (ii) for all s, $\sum_{i=1}^k p_i(s) > 0$. In other words, every state s is supported by exactly one p_i in the collection.

A positive measure μ on $\mathfrak{U}$ is *elementary* if its support is Dirac (degenerate) on $\mathfrak{U}_{s,\ell^\dagger(s)}$ (see Section 3.1 above for a definition) and the support on $\Delta(S)$ is a partitional system of probability measures on S. In other words, there exist $p_1, \ldots, p_k \in \Delta(S)$ and $\mathfrak{u}_s \in \mathfrak{U}_{s,\ell^\dagger(s)}$ for all s such that μ is supported on the finite collection $(p_1, \mathfrak{u}), \ldots, (p_k, \mathfrak{u})$ where $\mathfrak{u} = (\mathfrak{u}_s)_{s \in S}$.

Rather than saying that the marginal of μ on $\Delta(S)$ has support $\{p_1, \dots, p_k\}$, we will often say in the sequel that μ supports the partitional system (p_i).

With these definitions, it is clear that each $\hat{\mu} \in \hat{\mathfrak{M}}_p$ is elementary. The following proposition says that it is without loss of generality to restrict attention to elementary measures. Towards this end, let us define $\hat{V} : X_0 \to \mathbb{R}$ as

$$\hat{V}(x) := \sup_{\mu \in \hat{\mathfrak{M}}_p} \left[\sum_i \Big[\max_{f \in x} \sum_s p_i(s) \mathrm{u}_s(f(s)) \Big] \mu(p_i, \mathrm{u}) \right]$$

Proposition 3.24. For all nice x, $\hat{V}(x) = V(x)$. Moreover, the supremum in the definition of $\hat{V}$ is attained.

Proof. Let x be nice, $\mu \in \Upsilon_0(x)$, and $\{f_1, \dots, f_k\}$ the unique generator set of x. Let us first prove that $V(x) \leq \hat{V}(x)$. Let $(E_i^{\mu,x})$ be an optimal partition for μ at x, and let $\hat{\mu} = \mathfrak{D}(\mu, x, (E_i^{\mu,x}))$. Then,

$$\begin{aligned} V(x,\mu) &= \sum_i \max_{f \in x} \left[\sum_s \int_{E_i^{\mu,x}} p(s) \mathrm{u}_s(f(s))\, \mathrm{d}\mu(p, \mathrm{u}) \right] \\ &= \sum_i \max_{f \in x} \sum_s \bar{p}_i(s) \bar{\mathrm{u}}_{i,s}(f(s)) \end{aligned}$$

Lemma 3.23 says that μ cannot be of Type II (a or b) if $s \in J_i^c$, and hence must be either Type Ia or Type Ib. In either case, $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(f(s)) \leq 0 = \bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(\ell^\dagger(s))$ for all $s \in J_i^c$. Therefore, it must be that

$$V(x) = V(x,\mu) \leq \sum_i \max_{f \in x} \sum_s p_i(s) \hat{\mathrm{u}}_s(f(s)) = \hat{V}(x, \hat{\mu}) \leq \hat{V}(x)$$

We now prove that $\hat{V}(x) \leq V(x)$ for all nice x. Suppose, by way of contradiction, that $\hat{V}(x, \hat{\mu}) > V(x)$ for some nice x and $\hat{\mu} \in \hat{\mathfrak{M}}_p$. Suppose the optimal strategy here is to choose $f_i \in x$ whenever the 'interim information' is (p_i, u).

Now recall that $\hat{\mu} = \mathfrak{D}(\mu, y, (E_i^{\mu,y}))$ for some $\mu \in \Upsilon_0$ and $y \in X_0$. Consider the strategy ζ^μ that is constant on $E_i^{\mu,y}$, ie, satisfies $\zeta^\mu(E_i^{\mu,y}) = f_i \in x$ for each i (where f_i is the optimal choice when presented with the interim information (p_i, u)). The value of this strategy, $V(x, \mu, \zeta^\mu)$, is given by

$$\begin{aligned} V(x,\mu,\zeta^\mu) &= \sum_i \left[\sum_s \int_{E_i^{\mu,y}} p(s) \mathrm{u}_s(f_i(s))\, \mathrm{d}\mu(p, \mathrm{u}) \right] \\ &= \sum_i \left[\sum_s \bar{p}_i(s) \bar{\mathrm{u}}_{i,s}(f_i(s)) \right] \end{aligned}$$

It follows from Lemma 3.23 that μ is not Type II (a or b) at $E_i^{\mu,y}$ in state s for all $s \in J_i^c$. (Note that the partition (J_i) is generated by the unique $\xi \in \tilde{\Xi}_y$. Thus, (J_i) does not depend on x.) Therefore, for all such $s \in J_i^c$, it must be that $\bar{p}_i(s)\bar{\mathrm{u}}_{i,s}(f_i(s)) \leq 0$. For such an $s \in J_i^c$, if we replace $f_i(s)$ by ℓ_*, we obtain the new menu x', which has the property that $V(x', \mu, \zeta^{\mu}_{x'}) = \hat{V}(x, \hat{\mu})$. But this implies $V(x') \geq \hat{V}(x, \hat{\mu}) > V(x)$, where the strict inequality follows from our hypothesis. This violates IICC (Axiom 4) and Continuity because x' is obtained from x by replacing payoffs in acts in x by ℓ_*, so that $x \succsim x'$. This proves that $\hat{V}(x) = V(x)$ for all nice x.

Now, to show that the maximum is achieved in the definition of $\hat{V}(x)$, observe that for each nice x, there exists $\mu \in \Upsilon_0(x)$, so that

$$
\begin{aligned}
V(x) &= V(x, \mu) && \text{definition of } \mu \\
&\leq \hat{V}(x, \hat{\mu}) && \text{from proof of } V(x) \leq \hat{V}(x) \text{ above} \\
&\leq \hat{V}(x) && \text{definition of } \hat{V} \\
&\leq V(x) && \text{because } \hat{V}(x) \leq V(x) \text{ as proved above}
\end{aligned}
$$

where $\hat{\mu} = \mathfrak{D}(\mu, x, (E_i^{\mu,x}))$, $\mu \in \Upsilon_0(x)$, and $(E_i^{\mu,x})$ is an optimal partition strategy for μ at x. Therefore, $\hat{\mu}$ is $\hat{V}$-optimal for x, as claimed. □

Because V is Lipschitz, it follows immediately that $\hat{V}$ is also Lipschitz on X_0. By Proposition 3.17, X_0 is dense in X, so that $\hat{V}$ uniquely extends to X. It is easy to see that in the representation of $\hat{V}$, this amounts to replacing $\hat{\mathfrak{M}}_p$ with its closure. In what follows, we shall therefore assume that $\hat{\mathfrak{M}}_p$ is closed and that $\hat{V}$ is defined on X.

Thus far, we have shown that $\succsim$ is represented by a function $V : X \to \mathbb{R}$ that has the form

$$V(x) = \max_{\mu \in \mathfrak{M}} V(x, \mu) \tag{3.2}$$

where
- each $\mu \in \mathfrak{M}$ is a positive elementary measure,
- $V(x, \mu) = \left[\sum_{p \in \Delta(S)} \left(\max_{f \in x} \sum_{s \in S} p(s)\mathrm{u}_s\big(f(s)\big)\right)\mu(p; \mathrm{u})\right]$, and
- $V(\ell; \mu) = V(\ell; \mu')$ for all $\mu, \mu' \in \mathfrak{M}$ and $\ell \in L$.

Our first result establishes that we can replace an elementary measure by an elementary probability measure.

Lemma 3.25. Let μ be an elementary measure. Then, there exists an elementary *probability* measure $\hat{\mu}$ such that for all $x \in X$, $V(x, \mu) = V(x, \hat{\mu})$.

Proof. Let μ be supported on $(p_1, \mathrm{u}), \ldots, (p_k, \mathrm{u})$, and let $\|\mu\|_1$ be the total weight of μ. (That is, $\|\mu\|_1 := \sum_i \mu((p_i, \mathrm{u}))$.) For any $s \in S$, define $\hat{\mathrm{u}}_s := \|\mu\|_1 \mathrm{u}_s$, and for any $p \in \Delta(S)$, let $\hat{\mu}(p, \hat{\mathrm{u}}) := \mu(p, \mathrm{u}) / \|\mu\|_1$ where $\hat{\mathrm{u}} = (\hat{\mathrm{u}}_s)_{s \in S}$. It is easy to see that $\hat{\mu}$ so defined is elementary and is also a probability measure.

Moreover, we have

$$\begin{aligned} V(x,\hat{\mu}) &= \sum_p \max_{f\in x} \sum_s \hat{\mu}(p,\hat{\mathrm{u}})p(s)\hat{\mathrm{u}}_s(f(s)) \\ &= \sum_p \max_{f\in x} \sum_s \frac{\mu(p,\mathrm{u})}{\|\mu\|_1} p(s)\,\|\mu\|_1\,\mathrm{u}_s(f(s)) \\ &= V(x,\mu) \end{aligned}$$

which establishes the claim. □

Two partitional systems of probability measures $\{p_1,\dots,p_k\}$ and $\{q_1,\dots,q_k\}$ are *similar* if for all $i=1,\dots,k$, $\mathrm{supp}(p_i)=\mathrm{supp}(q_i)$.

Every elementary probability measure μ on $\Delta(S)$ supports a partitional system. We now show that we can replace, ie, without affecting utility considerations, μ by another elementary probability measure $\hat{\mu}$ that supports another partitional system that is similar to the partitional system supported by μ.

Lemma 3.26. Let μ be an elementary probability measure whose support is $(p_1,\mathrm{u}),\dots,(p_k,\mathrm{u})$. Let $\{\tilde{p}_1,\dots,\tilde{p}_k\}$ be a partitional system on $\Delta(S)$ that is similar to $\{p_1,\dots,p_k\}$. Then, there exists an elementary probability measure $\tilde{\mu}$ with support $(\tilde{p}_1,\tilde{\mathrm{u}}),\dots,(\tilde{p}_k,\tilde{\mathrm{u}})$ such that for all $x\in X$ we have $V(x,\mu)=V(x,\tilde{\mu})$.

Proof. Define $\tilde{\mathrm{u}}_s := (p_i(s)/\tilde{p}_i(s))\mathrm{u}_s$, and set $\mu(p_i,\mathrm{u})=\tilde{\mu}(\hat{p}_i,\tilde{\mathrm{u}})$, where $\tilde{\mathrm{u}}=(\tilde{\mathrm{u}}_s)_{s\in S}$. Then, we have

$$\begin{aligned} V(x,\tilde{\mu}) &= \sum_i \max_{f\in x} \sum_s \tilde{\mu}(\tilde{p}_i,\tilde{\mathrm{u}})\,\tilde{p}_i(s)\,\tilde{\mathrm{u}}_s(f(s)) \\ &= \sum_i \max_{f\in x} \sum_s \mu(p_i,\mathrm{u})\,p_i(s)\,\mathrm{u}_s(f(s)) \\ &= V(x,\mu) \end{aligned}$$

which completes the proof. □

Let μ be an elementary probability measure and define $\pi_\mu\in\Delta(S)$ as

$$\pi_\mu(s) := \sum_p \mu(p)\,p(s)$$

Let $\pi_0\in\Delta(S)$ and $P:=(J_i)$ be a partition of S. Then, the conditional probability induced by J_i is $q_i(\cdot,\pi_0\mid J_i)$ where

$$q_i(s;\pi_0\mid J_i) := \pi_0(s\mid J_i)$$

for all $J_i\in P$. It is easy to see that $\big(q_i(\cdot,\pi_0\mid J_i)\big)$ is a partitional system of probabilities on S. Conversely, let μ be an elementary measure that supports the partitional system (p_i). This induces the partition $P_\mu:=(J_i)$ of S where $J_i:=\mathrm{supp}(p_i)$.

Lemma 3.27. Let $\pi_0 \in \Delta(S)$, μ an elementary probability measure that supports the partitional system (p_i), and let (J_i) be the partition of S induced by (p_i). Then, there exists an elementary probability measure $\hat{\mu}$ such that
(a) μ^* supports the partitional system $\big(q_i(\cdot, \pi_0 \mid J_i)\big)$,
(b) $\pi_{\mu^*} = \pi_0$, and
(c) $V(x, \mu) = V(x, \mu^*)$ for all $x \in X$.

Proof. Let μ and π_0 be as hypothesized and consider the induced partitional system $\big(q_i(\cdot; \pi_0 \mid J_i)\big)$. By Lemma 3.26, there exists an elementary probability measure $\tilde{\mu}$ that supports $\big(q_i(\cdot; \pi_0 \mid J_i)\big)$ while keeping utilities unaltered.

For each s, define the utility function

$$\mathrm{u}^*_s := \left[\frac{\sum_i \tilde{\mu}\big(q_i(s; \pi_0 \mid J_i), \tilde{\mathrm{u}}\big)\mathbb{1}_{\{s\in J_i\}}}{\sum_i \pi_0(J_i)\mathbb{1}_{\{s\in J_i\}}}\right]\tilde{\mathrm{u}}_s$$

and observe that in the sums in both the numerator and denominator, only one term is non-zero. Now, define the elementary probability measure μ^* as follows: If s is supported by $q_i(\cdot; \pi_0 \mid J_i)$, set

$$\mu^*\big(q_i(\cdot; \pi_0 \mid J_i), \mathrm{u}^*\big) := \pi_0(J_i)$$

and 0 otherwise, which proves (a). With this definition, $\pi_{\mu^*}(s) = \sum_i \mu^*\big((q_i(\cdot; \pi_0 \mid J_i), \mathrm{u}^*)\big) \cdot q_i(s; \pi_0 \mid J_i) = \pi_0(s)$, as desired for the proof of (b). To see (c), notice that we have

$$\begin{aligned}
V(x, \mu^*) &= \sum_i \max_{f\in x} \sum_s \mu^*\big(q_i(\cdot; \pi_0 \mid J_i), \mathrm{u}^*\big)\, q_i(s; \pi_0 \mid J_i)\, \mathrm{u}^*_s(f(s)) \\
&= \sum_i \max_{f\in x} \sum_s \pi_0(J_i)\, q_i(s; \pi_0 \mid J_i)\, \frac{\tilde{\mu}\big(q_i(\cdot; \pi_0 \mid J_i), \tilde{\mathrm{u}}\big)}{\pi_0(J_i)}\, \tilde{\mathrm{u}}_s(f(s)) \\
&= \sum_i \max_{f\in x} \sum_s q_i(s; \pi_0 \mid J_i)\, \tilde{\mu}\big(q_i(\cdot; \pi_0 \mid J_i), \tilde{\mathrm{u}}\big)\, \tilde{\mathrm{u}}_s(f(s)) \\
&= V(x, \tilde{\mu}) = V(x, \mu)
\end{aligned}$$

which completes the proof. □

We are now in a position prove Proposition 3.14.

Proof of Proposition 3.14. We shall first prove (a) implies (b). We have shown that given the representation [♦] in Theorem 3 and IICC (Axiom 4), V has the form in [**3.2**], where every $\mu \in \mathfrak{M}$ is an elementary (positive, but finite) measure. Lemma 3.25 shows that it is without loss of generality to consider μ that are elementary *probability* measures. Consider such a μ and suppose it supports the partitional system (p_i). Let $J_i = \mathrm{supp}(p_i)$, and notice that (J_i) is a partition of S. Lemma 3.27 says that it is without loss of generality to assume that every μ supports the partitional system $\big(q_i(\cdot; \pi_0 \mid J_i)\big)$ (recall that $q_i(s; \pi_0 \mid J_i) = \pi_0(s \mid J_i)$) and also has the feature that $\pi_\mu(s) := \sum_i \mu\big(q_i(s;\pi_0 \mid J_i)\big) q_i(s;\pi_0 \mid J_i) = \pi_0(s)$ for all s. (To ease notational burden, in what follows we shall write $q_i(s;\pi_0 \mid J_i)$ as $q_i(s)$.)

In particular, this last property implies that $\mu(q_i, \mathfrak{u}) = \pi_0(J_i)$ and $\mu(q_i, \mathfrak{u}) q_i(s) = \pi_0(J_i)\pi_0(s \mid J_i)$. This implies

$$\begin{aligned} V(x,\mu) &:= \sum_i \left[\max_{f \in x} \sum_s q_i(s)\, \mathfrak{u}_s(f(s)) \right] \mu(q_i, \mathfrak{u}) \\ &= \sum_{J_i \in P} \left[\max_{f \in x} \sum_s \pi_0(s \mid J_i)\, \mathfrak{u}_s(f(s)) \right] \pi_0(J_i) \\ &= \sum_{J_i \in P} \left[\max_{f \in x} \sum_{s \in J_i} \pi_0(s \mid J_i)\, \mathfrak{u}_s(f(s)) \right] \pi_0(J_i) \\ &=: V'(x, \pi_0, (P, \mathfrak{u})) \end{aligned}$$

In other words, the informational content of the elementary probability measure μ is now encoded into the prior π_0, the partition $P = (J_i)$, and the utility functions $\mathfrak{u} = (\mathfrak{u}_s)$. Let $\mathfrak{M}'$ be the collection of all such pairs $(P, \mathfrak{u})$ induced by elementary probability measures in $\mathfrak{M}$. Then, we can write

$$\begin{aligned} V(x) &= \max_{\mu} V(x,\mu) \\ &= \max_{(P,\mathfrak{u}) \in \mathfrak{M}'} V'(x, \pi_0, (P, \mathfrak{u})) \\ &=: V'(x) \end{aligned}$$

where $V'(x) = V(x)$ for all $x \in X$; this proves the representation part.

Observe now — see [3.2] — that for all $\ell \in L$ and $\mu, \mu' \in \mathfrak{M}$, we have $V(\ell, \mu) = V(\ell, \mu')$. This implies that, for all $\ell \in L$ and $(P, \mathfrak{u}), (P', \mathfrak{u}') \in \mathfrak{M}'$, we have $V'(\ell, \pi_0, (P, \mathfrak{u})) = V(\ell, \pi_0, (P', \mathfrak{u}'))$.

Recall that $\ell^\dagger \in L$ is such that $\mathfrak{u}_s(\ell^\dagger(s)) = 0$ for all $s \in S$. For any $\alpha \in \Delta(C \times L)$, define $\hat{\ell}^s_\alpha \in L$ as

$$\hat{\ell}^s_\alpha(s') = \begin{cases} \alpha & \text{if } s' = s \\ \ell^\dagger(s') & \text{otherwise} \end{cases}$$

For all $(P, \mathfrak{u}), (P', \mathfrak{u}') \in \mathfrak{M}'$, we then have $V(\hat{\ell}^s_\alpha, \pi_0, (P, \mathfrak{u})) = V(\hat{\ell}^s_\alpha, \pi_0, (P', \mathfrak{u}'))$. Notice that $V(\hat{\ell}^s_\alpha, \pi_0, (P, \mathfrak{u})) = \pi_0(s)\mathfrak{u}_s(\alpha) = \pi_0(s)\mathfrak{u}'_s(\alpha) = V(\hat{\ell}^s_\alpha, \pi_0, (P', \mathfrak{u}'))$. Since this is true for all $\alpha \in \Delta(C \times L)$, it follows that $\mathfrak{u}_s$ and $\mathfrak{u}'_s$ are identical on $C \times L$ for all $(P, \mathfrak{u}), (P', \mathfrak{u}') \in \mathfrak{M}'$. This proves that (a) implies (b).

That (b) implies (a) follows immediately from Lemma 3.27 which shows how to construct an elementary measure μ given the prior π_0, the partition $P_\mu = (J_i)$, and the utility function $\mathfrak{u} = (\mathfrak{u}_s)$. □

3.3. Separable Representation

We now investigate the implication of imposing State-Contingent Indifference to Correlation (henceforth SCIC, Axiom 3). Suppose $V : X \to \mathbb{R}$ represents $\succsim$ and takes the form [3.1]. For each $(P, \mathfrak{u})$, define

$$V(x, (P, \mathfrak{u})) := \sum_{J \in P} \Big(\max_{f \in x} \sum_{s \in J} \pi_0(s \mid J)\, \mathfrak{u}_s(f(s)) \Big)\, \pi_0(J)$$

to be the expected utility when the pair $(P, \mathfrak{u})$ is chosen from $\mathfrak{M}_p$.

For each $\alpha \in \Delta(C \times X)$, define the equivalence class $[\alpha] := \{\alpha' \in \Delta(C \times X) : \alpha_1 = \alpha'_1,\ \alpha_2 = \alpha'_2\}$ of lotteries with identical marginals over C and X. Consider now the collection

$$\mathfrak{M}'_p := \Big\{ (P, \mathfrak{u}') : (P, \mathfrak{u}) \in \mathfrak{M}_p,\ \mathfrak{u}'_s(\alpha) = \min_{\alpha' \in [\alpha]} \mathfrak{u}_s(\alpha'),\ \text{and } \alpha \in \Delta(C \times X) \Big\}$$

and observe that $\mathfrak{u}'_s : \Delta(C \times X) \to \mathbb{R}$ is continuous and linear[6] so that $\mathfrak{u}'_s \in \mathbf{C}(C \times X)$. Moreover, for all $(P, \mathfrak{u}'), (\hat{P}, \hat{\mathfrak{u}}') \in \mathfrak{M}_p$, $\mathfrak{u}'_s|_{C \times L} = \hat{\mathfrak{u}}'_s|_{C \times L}$. This implies that $V(\ell, (P, \mathfrak{u}'))$ is independent of $(P, \mathfrak{u}') \in \mathfrak{M}'_p$.

Now define $V' : X \to \mathbb{R}$ as

[3.3] $$V'(x) := \max_{(P, \mathfrak{u}') \in \mathfrak{M}'_p} V(x, (P, \mathfrak{u}'))$$

Observe that V' is monotone, ie, $x \subset x'$ implies $V'(x) \leq V'(x')$. This follows immediately from the form of V' in [3.3]. We claim that V' also represents $\succsim$.

Lemma 3.28. Let V and V' be defined as in [3.1] and [3.3] respectively. Then, for all $x \in X$, $V(x) = V'(x)$.

Proof. Because V is Lipschitz, it suffices to show that $V(x) = V'(x)$ for all finite x. Notice first that for all $x \in X$, $V'(x) \leq V(x)$. To see this, fix x and let $(P, \mathfrak{u}')$ be a maximizing pair for V'. That is, $V'(x) = V(x, (P, \mathfrak{u}'))$. But $V(x, (P, \mathfrak{u}')) \leq V(x, (P, \mathfrak{u})) \leq V(x)$, where the first inequality follows from the definition of $\mathfrak{u}'_s$, which entails that for each $\alpha \in \Delta(C \times X)$, $\mathfrak{u}'_s(\alpha) \leq \mathfrak{u}_s(\alpha)$.

We shall now show that for all finite $x \in X$, $V(x) \leq V'(x)$. Note first that for each x and for any $(P, \mathfrak{u})$ that is optimal for x with $P = \{J_1, \ldots, J_m\}$, for $i = 1, \ldots, m$ we can define the acts

$$f_i := \arg\max_{f \in x} \sum_{s} \pi_0(s \mid J_i)\, \mathfrak{u}_s(f(s))$$

Then, we see that $V(x) = V\big(\{f_1, \ldots, f_m\}\big)$, ie, $\{f_1, \ldots, f_m\}$ is the *generator set* of x.

(6) It is easy to see that for all $\alpha' \in [\alpha]$ and $\beta' \in [\beta]$, $(\frac{1}{2}\alpha' + \frac{1}{2}\beta')_i = \frac{1}{2}\alpha_i + \frac{1}{2}\beta_i$ for $i = 1, 2$. This, the continuity of $\mathfrak{u}'_s(\cdot; P)$, and the fact that $\mathfrak{u}_s(\alpha'; P)$ is linear in α', immediately imply that $\mathfrak{u}'_s(\cdot; P)$ is linear.

Now define the act $\hat{f}_i$ so that for each $s \in S$,

$$\hat{f}_i(s) = \arg\min_{\alpha \in [f_i(s)]} \mathfrak{u}_s(\alpha)$$

With this definition, we make the following observations.

(a) $V(\{f_1, \ldots, f_m\}) = V(\{\hat{f}_1, \ldots, \hat{f}_m\})$ by repeated application of SCIC (Axiom 3).

(b) $V(\{\hat{f}_1, \ldots, \hat{f}_m\}, (P, \mathfrak{u})) = V(\{\hat{f}_1, \ldots, \hat{f}_m\}, (P, \mathfrak{u}'))$ for all pairs $(P, \mathfrak{u})$ and $(P, \mathfrak{u}')$. This follows from the definitions of $\mathfrak{u}'_s$ and $\hat{f}_i$, which imply that in any state s, $\mathfrak{u}_s(\hat{f}_i(s)) = \mathfrak{u}'_s(\hat{f}_i(s))$.

(c) $V(\{\hat{f}_1, \ldots, \hat{f}_m\}) = V(\{\hat{f}_1, \ldots, \hat{f}_m\}, (\hat{P}, \hat{\mathfrak{u}}))$ where $(\hat{P}, \hat{\mathfrak{u}})$ is a maximizing pair in $\mathfrak{M}'_p$ for $\{\hat{f}_1, \ldots, \hat{f}_m\}$ under V.

(d) $V(\{f_1, \ldots, f_m\}, (\hat{P}, \hat{\mathfrak{u}}')) = V(\{\hat{f}_1, \ldots, \hat{f}_m\}, (\hat{P}, \hat{\mathfrak{u}}'))$. This follows from the definitions of $\hat{\mathfrak{u}}'$ and $\hat{f}_i$, which imply that in any state s, $\hat{\mathfrak{u}}'_s(f_i(s)) = \hat{\mathfrak{u}}'_s(\hat{f}_i(s))$.

We can now use these equalities to form the following chain.

$$\begin{aligned}
V(x) &= V(\{f_1, \ldots, f_m\}) && \text{definition of } \{f_1, \ldots, f_m\} \\
&= V(\{\hat{f}_1, \ldots, \hat{f}_m\}) && \text{established in (a) above} \\
&= V(\{\hat{f}_1, \ldots, \hat{f}_m\}, (\hat{P}, \hat{\mathfrak{u}}')) && \text{established in (c) above} \\
&= V(\{\hat{f}_1, \ldots, \hat{f}_m\}, (\hat{P}, \hat{\mathfrak{u}}')) && \text{established in (b) above} \\
&= V(\{f_1, \ldots, f_m\}, (\hat{P}, \hat{\mathfrak{u}}')) && \text{established (d) above} \\
&\le V'(\{f_1, \ldots, f_m\}) && \text{definition of } V' \\
&\le V'(x) && \text{monotonicity of } V'
\end{aligned}$$

which completes the proof. □

We can now state the main result of this section.

Proposition 3.29. Let V be as in [3.1] and suppose V represents $\succsim$. Then, the following are equivalent.

(a) $\succsim$ satisfies SCIC (Axiom 3).

(b) There exist functions $u_s \in \mathbf{C}(C)$ and a set $\mathfrak{M}''_p$ consisting of pairs of $(P, (v_s))$ where P is a partition and $v_s \in \mathbf{C}(X)$ for each s such that $(P, (v_s)), (P', (v'_s)) \in \mathfrak{M}''_p$ implies $v_s|_L = v'_s|_L$ for all $s \in S$, and V can be written as

$$V(x) = \max_{(P,(v_s)) \in \mathfrak{M}''_p} \sum_{J \in P} \pi_0(J) \max_{f \in x} \sum_s \pi_0(s \mid J) \left[u_s(f_1(s)) + v_s(f_2(s))\right] \quad [\mathbf{3.4}]$$

Proof. It is easy to see that (b) implies (a). We now show that (a) implies (b).

Lemma 3.28 implies we can replace V in [3.1] by V' in [3.3]. Moreover, from the definition of V in [3.1], $\mathfrak{u}_s(\alpha) = \mathfrak{u}'_s(\alpha)$ for all $(P, \mathfrak{u}), (P', \mathfrak{u}') \in \mathfrak{M}'_p$ and for all $\alpha \in \Delta(C \times L)$.

For any $\alpha \in \Delta(C \times X)$ with marginals α_1 and α_2, let $\alpha_1 \otimes \alpha_2 \in \Delta(C) \times \Delta(X)$ denote the product lottery with the same marginals. Recall that $\ell^\dagger \in L$ is such that $\mathfrak{u}_s(\ell^\dagger(s)) = 0$ for all s. Given $(P, \mathfrak{u})$, now define

- $u_s(\alpha_1) := \mathfrak{u}_s\big(\alpha_1 \otimes \ell_2^\dagger(s)\big)$ (and notice $\mathfrak{u}_s(\alpha) = \mathfrak{u}'_s(\alpha)$ for all $(P, \mathfrak{u}), (P', \mathfrak{u}') \in \mathfrak{M}_p$ and for all $\alpha \in \Delta(C \times L)$ because $\alpha_1 \otimes \ell_2^\dagger(s) \in \Delta(C \times L)$); and
- $v_s(\alpha_2) := \mathfrak{u}_s\big(\ell_1^\dagger(s) \otimes \alpha_2\big)$.

With these definitions, $u_s \in \mathbf{C}(C)$ while $v_s(\cdot) \in \mathbf{C}(X)$. Notice that the lotteries $\frac{1}{2}\big(\alpha_1 \otimes \alpha_2\big) + \frac{1}{2}\ell^\dagger(s)$ and $\frac{1}{2}\big(\alpha_1 \otimes \ell_2^\dagger(s)\big) + \frac{1}{2}\big(\ell_1^\dagger(s) \otimes \alpha_2\big)$ have identical marginals, which implies that for every $(P, \mathfrak{u})$,

$$\mathfrak{u}_s\left(\tfrac{1}{2}\big(\alpha_1 \otimes \alpha_2\big) + \tfrac{1}{2}\ell^\dagger(s)\right) = \mathfrak{u}_s\left(\tfrac{1}{2}\big(\alpha_1 \otimes \ell_2^\dagger(s)\big) + \tfrac{1}{2}\big(\ell_1^\dagger(s) \otimes \alpha_2\big)\right)$$

This means we can write

$$\begin{aligned}\tfrac{1}{2}\mathfrak{u}_s\big(\alpha_1 \otimes \alpha_2\big) + \tfrac{1}{2}\mathfrak{u}_s\big(\ell^\dagger(s)\big) &= \mathfrak{u}_s\left(\tfrac{1}{2}\big(\alpha_1 \otimes \alpha_2\big) + \tfrac{1}{2}\ell^\dagger(s)\right) = \mathfrak{u}_s\left(\tfrac{1}{2}\big(\alpha_1 \otimes \ell_2^\dagger(s)\big) + \tfrac{1}{2}\big(\ell_1^\dagger(s) \otimes \alpha_2\big)\right) \\ &= \tfrac{1}{2}\mathfrak{u}_s\big(\alpha_1 \otimes \ell_2^\dagger(s)\big) + \tfrac{1}{2}\mathfrak{u}_s\big(\ell_1^\dagger(s) \otimes \alpha_2\big) = \tfrac{1}{2}u_s(\alpha_1) + \tfrac{1}{2}v_s(\alpha_2)\end{aligned}$$

where the second equality holds because $\mathfrak{u}_s(\cdot)$ is constant on the equivalence class of lotteries with identical marginals. The first and third equalities from the linearity of $\mathfrak{u}_s(\cdot)$, while the last equality follows from the definitions of u_s and $v_s(\cdot)$.

But we have already stipulated that $\mathfrak{u}_s\big(\ell^\dagger(s)\big) = 0$, which implies that for all s, we have

$$\mathfrak{u}_s\big(\alpha_1 \otimes \alpha_2\big) = u_s(\alpha_1) + v_s(\alpha_2)$$

Substituting in [3.3] and invoking Lemma 3.28 gives us [3.4], as desired. □

As always, for each $(P, (v_s)) \in \mathfrak{M}''_p$, define $V(x, (P, (v_s))$ as

$$V(x, (P, (v_s))) = \sum_{J \in P} \pi_0(J) \max_{f \in x} \sum_s \pi_0(s \mid J)\big[u_s(f_1(s)) + v_s(f_2(s))\big]$$

3.4. Representation with Deterministic Continuation Utilities

Thus far, we have seen that $\succsim$ has a representation as in [3.4]. We now impose Concordant Independence (Axiom 5) and show that $\succsim$ then has a representation of the form

$$V(x) = \max_{P \in \mathfrak{M}_p^\sharp} \sum_{J \in P} \left[\max_{f \in x} \sum_s \pi_0(s \mid J)\big[u_s(f_1(s)) + v_s(f_2(s), P)\big]\pi_0(J)\right] \qquad [3.5]$$

where $\mathfrak{M}_p^\sharp$ is a finite collection of partitions P of S, $u_s \in \mathbf{C}(C)$, and $v_s(\cdot, P) \in \mathbf{C}(X)$ for each $s \in S$ and $P \in \mathfrak{M}_p^\sharp$, with the property that for all $P, P' \in \mathfrak{M}_p^\sharp$, $s \in S$, $v_s(\cdot, P)|_L = v_s(\cdot, P')|_L$.

For a fixed P in the representation in [3.4], let X'_P and $\hat{X}_P$ be defined as follows:

$$\begin{aligned}X'_P &:= \big\{x : V(x) = V\big(x, (P, (v_s))\big) \text{ for some } (P, (v_s)) \in \mathfrak{M}''_p \text{ and} \\ &\qquad V(x) > V\big(x, (Q, (v'_s))\big) \text{ for all } (Q, (v'_s)) \in \mathfrak{M}''_p \text{ such that } P \neq Q\big\} \\ \hat{X}_P &:= \big\{x : V(x) = V\big(x, (P, (v_s))\big) \text{ for some } (P, (v_s)) \in \mathfrak{M}''_p \text{ and} \\ &\qquad V(x) \geq V\big(x, (Q, (v'_s))\big) \text{ for all } (Q, (v'_s)) \in \mathfrak{M}''_p \text{ such that } P \neq Q\big\}\end{aligned}$$

Recall that $x_1(P) := x(P, \hat{\omega})$ as in [♠] in Section 3.1. That is, for any partition P, $x_1(P) \in X$ is a *one-period* problem where the choice of P is optimal.

Lemma 3.30. Let $x \in X'_P$. Then, for all $\lambda \in (0,1)$, $(1-\lambda)x + \lambda x_1(P) \in X'_P$. Moreover, $V\big((1-\lambda)x + \lambda x_1(P)\big) = V\big((1-\lambda)x + \lambda \ell^*\big) > V\big((1-\lambda)x + \lambda x_1(Q)\big)$ if P is not finer than Q.

Proof. We begin by establishing three claims.

(i) In the representation [3.4], $v_s(\ell^*) > v_s(\ell_*)$ for all $s \in S$.

(ii) $V(x_1(Q)) \leq V(\ell^*)$ for all $Q \in \mathcal{P}$.

(iii) $V\big(x_1(Q), (P, (v_s))\big) = V(\ell^*)$ if, and only if, P is finer than Q.

To see (i), observe that by the RAA representation of Appendix D.1, $\big[\ell^* \oplus_{(1,S\backslash s)} \ell_*\big] \succ \big[\ell_* \oplus_{(1,S\backslash s)} \ell_*\big] = \ell_*$ for all $s \in S$. Because we have $v_s(\ell) = v'_s(\ell)$ for all $\ell \in L$ and $(P, (v_s)), (P', (v'_s)) \in \mathfrak{M}''_p$ in [3.4], $v_s(\ell^*) > v_s(\ell_*)$ follows for all $s \in S$.

Given claim (i), and because $u_s(c_s^+) \geq u_s(c_s^-)$ for all s, claim (ii) follows by evaluating V in [3.4] at $x_1(Q)$.

To establish claim (iii), consider first P finer than Q, then

$$\begin{aligned} V\left(x_1(Q), (P, (v_s))\right) &= \sum_{J \in P} \pi_0(J) \max_{f \in x_1(Q)} \sum_s \pi_0(s \mid J)\left[u_s(f_1(s)) + v_s(f_2(s))\right] \\ &= \sum_{J \in P} \pi_0(J) \sum_s \pi_0(s \mid J)\left[u_s(c_s^+) + v_s(\ell^*)\right] \\ &= V\left(\ell^*, (P, (v_s))\right) = V(\ell^*) \end{aligned}$$

Now suppose instead that P is not finer than Q. Then there must be $J \in P$ with $s \in J$ such that

$$\left[\arg\max_{f \in x_1(Q)} \left(\sum_{s'} \pi_0(s' \mid J)\left[u_{s'}\left(f_1(s')\right) + v_{s'}\left(f_2(s')\right)\right]\right)\right](s) = \ell_*(s)$$

Then, by claim (i) and because $u_s(c_s^+) \geq u_s(c_s^-)$ for all s by construction, we find that $V(\ell^*) > V\big(x_1(Q), (P, (v_s))\big)$.

With the claims in hand, observe that

$$V\left((1-\lambda)x + \lambda x_1(P)\right) \geq V\left((1-\lambda)x + \lambda x_1(P, (P', \cdot))\right)$$

for all $(P', \cdot) \in \mathfrak{M}''_p$. Let (v_s) be such that $(P, (v_s)) \in \mathfrak{M}''_p$ and $V(x) = V(x, (P, (v_s)))$. Then

$$\begin{aligned} V\left((1-\lambda)x + \lambda x_1(P), (P, (v_s))\right) &= (1-\lambda)V(x) + \lambda V\left(x_1(P), (P, (v_s))\right) \\ &= (1-\lambda)V(x) + \lambda V(\ell^*) = V\left((1-\lambda)x + \lambda \ell^*\right) \end{aligned}$$

by claims (ii) and (iii). Moreover, for any other $(Q, (v'_s)) \in \mathfrak{M}''_p$,

$$\begin{aligned} V((1-\lambda)x + \lambda x_1(P), (Q, (v'_s))) &= (1-\lambda)V(x, (Q, (v'_s))) + \lambda V(x_1(P), (Q, (v'_s))) \\ &< (1-\lambda)V(x) + \lambda V(\ell^*) \end{aligned}$$

where the strict inequality is because $V(x,(Q,(v'_s))) < V(x) = V(x,(P,\cdot))$ (recall that $x \in X'_P$) and $V(x_1(P),(Q,(v'_s))) \le V(\ell^*)$ (claim (ii) above). This implies $(1-\lambda)x + \lambda x_1(P) \in X'_P$. Moreover, it now follows immediately that $V((1-\lambda)x + \lambda x_1(P)) = V((1-\lambda)x + \lambda\ell^*)$.

Finally, suppose P is not finer than Q. Consider the menu $(1-\lambda)x + \lambda x_1(Q)$ and suppose $(P',\cdot) \in \mathfrak{M}''_p$ is optimal for this menu. Notice that if $P' \neq P$, then $V(x,(P',\cdot)) < V(x,(P,\cdot)) = V(x)$ by virtue of $x \in X'_P$, and that if $P = P'$, then $V(x_1(Q),(P,\cdot)) < V(\ell^*)$ by case (iii) because P is not finer than Q. Thus,

$$\begin{aligned} V\big((1-\lambda)x + \lambda x_1(Q)\big) &= V\big((1-\lambda)x + \lambda x_1(Q),(P',\cdot)\big) \\ &= (1-\lambda)V(x,(P',\cdot)) + \lambda V(x_1(Q),(P',\cdot)) \\ &< (1-\lambda)V(x,(P,\cdot)) + \lambda V(\ell^*) = V\big((1-\lambda)x + \lambda\ell^*\big) \end{aligned}$$

which completes the proof. □

Lemma 3.31. X'_P is convex and consists of concordant choice problems.

Proof. Because V is L-affine, any $(P,(v_s))$ that is optimal for x is also optimal for $(1-t)x + t\ell$ for all $t \in [0,1]$ and $\ell \in L$, and vice versa. Thus, $x \in X'_P$ if, and only if, $(1-t)x + t\ell \in X'_P$.

Let $x, y \in X'_P$ be such that $x \succsim y$. It follows from IICC (Axiom 4) that $x \succsim y \succsim \ell_*$. By Continuity of $\succsim$ (Axiom 1 (b)), there exists a $t \in [0,1]$ such that $(1-t)x + t\ell_* \sim y$ and $(1-t)x + t\ell_* \in X'_P$ (as we observed above). Thus, it is without loss of generality to consider $x, y \in X'_P$ such that $x \sim y$.

By Lemma 3.30, $(P,(\tilde v_s)) \in \mathfrak{M}''_p$ remains optimal for both $(1-\lambda)x + \lambda x_1(P)$ and $(1-\lambda)y + \lambda x_1(P)$, for all $\lambda \in (0,1)$. It follows that x and y are λ-concordant (Definition C.1), and by Concordant Independence (Axiom 5), so are x and $\frac{1}{2}x + \frac{1}{2}y$. It follows that x, y, and $\frac{1}{2}x + \frac{1}{2}y$ are concordant.

Now suppose $(Q,(v'_s)) \in \mathfrak{M}''_p$ is optimal for $\frac{1}{2}x + \frac{1}{2}y$. Then, $V(x,(Q,(v'_s))) \le V(x)$ and $V(y,(Q,(v'_s))) \le V(y)$. Because x, y, and $\frac{1}{2}x + \frac{1}{2}y$ are λ-concordant, $V(x) = V(y) = V(\frac{1}{2}x + \frac{1}{2}y)$, ie, $(Q,(v'_s))$ is optimal at x and y. But $x, y \in X'_P$, which implies that $Q = P$. That is, $\frac{1}{2}x + \frac{1}{2}y \in X'_P$.

Standard arguments now imply that every $z \in [x,y]$ is concordant with x and y and the argument above establishes that $Q = P$ for any maximizer $(Q,(v'_s))$ at z, ie, X'_P is convex. □

Lemma 3.32. For each $x \in X$, there exists $(P,(v_s)) \in \mathfrak{M}''_p$ such that $x \in \mathrm{cl}(X'_P)$.

Proof. Let $x \in \hat X_{P_1} \cap \cdots \cap \hat X_{P_n}$ and suppose $n \ge 2$ (because if $n = 1$, then $x \in X'_P \subset \mathrm{cl}(X'_P)$). Without loss of generality, suppose that none of $P_2, \ldots, P_n$ are finer than P_1. In analogy to the arguments in the proof of Lemma 3.30, we find that $V((1-\lambda)x + \lambda x_1(P_1),(P_1,(v^1_s))) = V((1-\lambda)x + \lambda\ell^*) > V((1-\lambda)x + \lambda x_1(P_1),(P_i,(v^i_s)))$ for some v^1_s with $(P,(v^1_s)) \in \mathfrak{M}''_p$ and all $(P_i,(v^i_s)) \in \mathfrak{M}''_p$ for $i = 2, \ldots, n$. That is, $(1-\lambda)x + \lambda x_1(P_1) \in X'_{P_1}$ for all $\lambda \in (0,1)$, which implies $x \in \mathrm{cl}(X'_{P_1})$ as claimed. □

Lemma 3.33. Let $x \in X'_P$ and let Y_x denote the set of choice problems that (i) are concordant with x, and (ii) have a unique optimal partition. Then, $Y_x = X'_P$.

Proof. By hypothesis, P is uniquely optimal for x. Let $Q \neq P$ be optimal for $y \in Y_x$. Because V is L-affine, we may assume without loss of generality, that $x \sim y$. (This is made clear in the proof of Lemma 3.31.) If P is not finer than Q, by Lemma 3.30, $(1-\lambda)y + \lambda x_1(Q) \succ (1-\lambda)x + \lambda x_1(Q)$, which contradicts our assumption that x and y are concordant. Conversely, if Q is not finer than P, then an analogous argument establishes that $(1-\lambda)x + \lambda x_1(P) \succ (1-\lambda)y + \lambda x_1(P)$, which also contradicts our assumption that x and y are concordant. Therefore, P must be the unique optimal partition for any $y \in Y_x$. Thus, $Y_x \subset X'_P$. That $X'_P \subset Y_x$ is an immediate consequence of Lemma 3.31. □

Notice that replacing $\mathfrak{M}''_p$ with its weak* closure (in the event that it is not weak* compact) in [3.4] does not affect the representation. Therefore, we shall now assume that $\mathfrak{M}''_p$ is weak*-compact.

Lemma 3.34. Let $x \in \text{cl}(X'_P)$. Then, there exists (v_s) such that $(P, (v_s)) \in \mathfrak{M}''_p$ is optimal for all $y \in \text{cl}(X'_P)$.

Proof. By Lemma 3.33, $Y_x \subset X'_P$ which, by Lemma 3.31, is convex. By Concordant Independence, $\succsim|_{X'_P}$ satisfies Independence. That is, $V|_{X'_P}$ is linear. It follows from Lemma 2.5 above that there exists (v_s) such that $(P, (v_s))$ is optimal for all $x \in X'_P$. Continuity now implies that $(P, (v_s))$ is optimal for all $x \in \text{cl}(X'_P)$. □

It follows that we can replace the set $\mathfrak{M}''_p$ by a finite collection $\{(P_1, (v^1_s)), \ldots, (P_n, (v^n_s))\} = \mathfrak{M}^\sharp_p$ as in [3.5]. Thus, we have shown that (a) implies (b) in the following proposition. That (b) implies (a) is clear.

Corollary 3.35. Let V be as in [3.4] and suppose V represents $\succsim$. Then, the following are equivalent.

(a) $\succsim$ satisfies Concordant Independence (Axiom 5).
(b) V can be written as in [3.5].

As always, for any partition $P \in \mathfrak{M}^\sharp_p$, we define

$$V(x, P) = \sum_{J \in P} \max_{f \in x} \left[\sum_s \pi_0(s \mid J) \left[u_s(f_1(s)) + v_s(f_2(s), P) \right] \pi_0(J) \right]$$

Note that the proof of the existence of an ICP representation in DKS begins with such a representation in hand.

To see that $L \simeq \mathcal{F}(\Delta(C \times L))$, note that we can define $L^{(1)} := \mathcal{F}(\Delta(C))$ and then recursively define $L^{(n)} := \mathcal{F}(\Delta(C \times L^{(n-1)}))$ as the space of consumption streams of length n. Just as with the definition of the space of choice problems X in Appendix A.2 of DKS, we say that L is the space of all *consistent* sequences in $\bigtimes_{n=1}^{\infty} L^{(n)}$.

The *support* of a consumption stream $\ell \in L$ is a set $B \subset C$ such that at any date and in any state, the realized consumption lies in B. A consumption stream has *finite support* if its support in C is finite. For any finite set $B \subset C$, we can define L_B as the space of all consumption streams with prizes in B. Formally, $L_B \simeq \mathcal{F}(\Delta(C \times L_B))$. Let L_0 be the space of all consumption streams with finite support. That is, $L_0 := \bigcup \{L_B : B \subset C,\ B \text{ finite}\}$.

Recall the consumption stream $\ell^\dagger \in L$ which delivers $c^\dagger(s)$ in state s at every date. Clearly, the support of $\ell^\dagger$ is finite. Analogous to L_0, we can define $L_0^{(n)}$ as the space of consumption streams of length n with finite support. For any $\ell^{(n)} \in L_0^{(n)}$, $\ell^{(n)} \diamond \ell^\dagger \in L_0$, where $\ell^{(n)} \diamond \ell^\dagger$ is the concatenation of $\ell^\dagger$ to $\ell^{(n)}$. In other words, each $L^{(n)}$ is naturally embedded in L_0.

Proposition 4.1. The space L_0 is dense in L.

Proof. Because probability measures on C with finite support are dense in $\Delta(C)$, it follows that for all $n \geq 1$, $L_0^{(n)}$ is dense in $L^{(n)}$. (The metrics defined on $L^{(n)}$ make this clear — see Appendix A.2 of DKS for a formal definition.) By the definition of the product metric (see Appendix A.2 of DKS), this means that for any $\ell \in L$ and $\varepsilon > 0$, there exists an n and an $\ell^{(n)} \in L^{(n)}$ such that $d(\ell, \ell^{(n)} \diamond \ell^\dagger) < \varepsilon$, where $\ell^{(n)} \diamond \ell^\dagger$ is the concatenation of $\ell^\dagger$ to $\ell^{(n)}$. This completes the proof. □

It follows immediately from Lipschitz Continuity (Axiom 1(c)) that $\succsim|_L$ is non-trivial, see Corollary 3.3. We now show that $\succsim_s$ (as defined in Appendix C.1 of DKS) is also non-trivial for each $s \in S$.

Lemma 4.2. Let $\ell^0, \ell^1 \in L$. Then, $\ell^0(s) \sim_s \ell^1(s)$ for all $s \in S$ implies $\ell^0 \sim|_L \ell^1$.

Proof. By definition of $\succsim_s$, $\ell^0(s) \sim_s \ell^1(s)$ if, and only if, $\ell^0 \oplus_{(1,S\setminus s)} \ell_* \sim|_L \ell^1 \oplus_{(1,S\setminus s)} \ell_*$. Repeatedly applying L-Independence, we find

$$\frac{1}{n}\ell^0 + \frac{n-1}{n}\ell_* = \frac{1}{n}\sum_{s\in S} \ell^0 \oplus_{(1,S\setminus s)} \ell_* \sim|_L \frac{1}{n}\sum_{s\in S} \ell^1 \oplus_{(1,S\setminus s)} \ell_* = \frac{1}{n}\ell^1 + \frac{n-1}{n}\ell_*$$

By L-Independence, we find $\ell^0 \sim|_L \ell^1$. (More precisely, this follows immediately once we note that, by the Mixture Space Theorem, $\succsim|_L$ has an affine representation.) □

Lemma 4.3. There exists $s \in S$ such that $\ell^*(s) \nsim_s \ell_*(s)$. For all $s \in S$, there exists $s' \in S$ such that $(c, \ell^* \oplus_{(1,S\setminus s)} \ell_*) \nsim_{s'} (c, \ell_*)$.

Proof. Corollary 3.3 says that $\ell^* \succ |_L \ell_*$. Therefore, by (the contrapositive to) Lemma 4.2, there must exists an s such that $\ell^*(s) \not\sim_s \ell_*(s)$. In particular, then, $\ell^* \oplus_{(1,S\setminus s)} \ell_* \not\sim |_L \ell_*$.

To see the second part, let us suppose by way of contradiction that for all $s' \in S$, $(c, \ell^* \oplus_{(1,S\setminus s)} \ell_*) \sim_{s'} (c, \ell_*)$. Now, set ℓ^0, ℓ^1 such that $\ell^0(s') = (c, \ell^* \oplus_{(1,S\setminus s)} \ell_*)$, while $\ell^1(s') = (c, \ell_*)$. It follows from Lemma 4.2 that $(c, \ell^* \oplus_{(1,S\setminus s)} \ell_*) \sim |_L (c, \ell_*)$.

Now, L-Stationarity (Axiom 2(c)) and the fact that $\ell^* \oplus_{(1,S\setminus s)} \ell_* \not\sim |_L \ell_*$ imply that we have $(c, \ell^* \oplus_{(1,S\setminus s)} \ell_*) \not\sim |_L (c, \ell_*)$, which yields the desired contradiction. □

Proposition 4.4. For all $s \in S$, $\succsim_s$ is non-trivial.

Proof. Lemma 4.3 and (the contrapositive to) L-History Independence (Axiom 2(b)) imply $(c, \ell^* \oplus_{(1,S\setminus s)} \ell_*) \not\sim_{s''} (c, \ell_*)$ for all $s'' \in S$, as claimed. □

Proposition 4.5. The preference $\succsim |_L$ on L has a standard RAA representation. Moreover, Π and δ are unique and the collection $(u_s)_{s\in S}$ is unique up to a common positive scaling.

As described in Appendix C.1 of DKS, for each $s \in S$, $\succsim_s$ is an induced preference over $\Delta(C \times L)$. Let $\succsim_s^C$ denote the induced preference over $\Delta(C)$ in state s. It is clear that $\succsim_s^C$ is well defined, continuous on $\Delta(C)$, and satisfies Independence. These properties imply there exist $\succsim_s^C$-maximal and -minimal lotteries that are degenerate; denote them by $c^*(s)$ and $c_*(s)$. Let F_0 be the finite set of consumption defined as

$$F_0 := \{c_*(s), c^\dagger(s), c^*(s) : s \in S\}$$

Lemma 4.6. For any finite set $B \subset C$, the induced preference $\succsim |_{L_B}$ satisfies the Axioms stated in Corollary 5 of Krishna and Sadowski (2014, henceforth KS).

Proof. It follows from Proposition 4.4 that each $\succsim_s$ is non-trivial. That is, $\succsim |_L$ is state-wise nontrivial. In addition, $\succsim |_L$ is continuous, satisfies Independence, and is separable in ℓ_1 and ℓ_2, thereby satisfying Axioms 2, 3, and 5 in KS. Axioms 6, 7, and 9 in KS correspond to properties (c), (d), and (b) of Consumption Stream Properties (Axiom 2). □

We now proceed to the proof of Proposition 4.5.

Proof of Proposition 4.5. Let $B \subset C$ be finite. By Lemma 4.6, $\succsim |_{L_B}$ satisfies the Axioms in Corollary 5 of KS. This implies there exists a tuple $\big((u_s^B)_{s\in S}, \delta^B, \Pi^B\big)$ that is an RAA representation of $\succsim |_{L_B}$. If $F_0 \subset B$, then we may assume, without loss of generality, that $u_s^B(c^\dagger(s)) = 0$ for all $s \in S$. Then, Corollary 5 in KS says that the collection of utilities (u_s^B) is uniquely identified up to a joint scaling, and that Π^B and δ^B are also uniquely determined.

Now, consider any other finite set D such that $F_0 \subset B \subset D$. By Lemma 4.6, $\succsim |_{L_D}$ also has an RAA representation $\big((u_s^D)_{s\in S}, \delta^D, \Pi^D\big)$. As before, if we set $u_s^D(c^\dagger(s)) = 0$ for all $s \in S$, then the collection of utilities (u_s^D) is identified up to a common scaling. Now, because $B \subset D$, we have $L_B \subset L_D$. Therefore, the RAA representation $\big((u_s^D)_{s\in S}, \delta^D, \Pi^D\big)$ of $\succsim |_{L_D}$ when restricted to L_B, is also a representation of $\succsim |_{L_B}$. And this representation has the feature that $u_s^D(c^\dagger(s)) = 0$ for all $s \in S$. Once again, the uniqueness of the RAA representation implies that a single joint scaling of the collection (u_s^D) results in $u_s^D|_B = u_s^B$ for all $s \in S$, $\Pi^B = \Pi^D$, and $\delta^B = \delta^D$.

Recall that $c^*(s) \succsim_s^C \alpha \succsim_s^C c_*(s)$ for all $\alpha \in \Delta(C)$. Because u_s^B and u_s^D represent, respectively, $\succsim_s^C |_{\Delta(B)}$ and $\succsim_s^C |_{\Delta(D)}$, it must be that $\lambda^*(s) := u_s^j(c^*(s))$ and $\lambda_*(s) := u_s^j(c_*(s))$ for $j = B, D$. Since B and D are arbitrary, it follows that it holds for all finite B that contains F_0. In other words, the Markov transition operator Π has been identified uniquely, as has the discount factor $\delta \in (0,1)$.

Let $u_s \in \mathbf{C}(C)$ be a vN-M utility representation of $\succsim_s^C$ such that $u_s(c^\dagger(s)) = 0$. Both $u_s|_{\Delta(B)}$ and u_s^B are vN-M representations of $\succsim^C |_{\Delta(B)}$ and by the Mixture Space Theorem, differ at most by a positive affine transformation. Because they agree on $c^\dagger(s)$, they differ at most by a positive scaling. Therefore, if we scale u_s so that $u_s(c^*(s)) = \lambda^*(s)$, we must necessarily have $u_s(c_*(s)) = \lambda_*(s)$ for all $s \in S$.

Consider, now, the tuple $((u_s), \Pi, \delta)$, and the functional $W_0 : L \to \mathbb{R}$ defined as $W_0(\ell) := \sum_s \pi_0(s) W(\ell, s)$, where

$$W(\ell, s) := \sum_{s'} \Pi(s, s')\big[u_{s'}(\ell_1(s')) + \delta W(\ell_2(s'), s')\big]$$

It is easy to see that the function W_0[7] is uniquely determined by the tuple $\big((u_s)_{s\in S}, \delta, \Pi\big)$. As established above, W_0 represents $\succsim |_{L_B}$ for every finite B. In other words, W_0 represents $\succsim |_{L_0}$. Proposition 4.1 says that L_0 is dense in L, and because W_0 is (uniformly) continuous, it also represents $\succsim$ on L. The uniqueness of the RAA representation of $\succsim |_L$ (given our normalizations) follows immediately, which concludes the proof. □

5. A Metric on the Space of Partitions

In this section, we define a natural metric on the space of partitions that is related to the informational content of the partitions. The metric we introduce is fairly standard. However, we have been unable to find a formulation suitable for our purposes, so we prove that our proposed metric is indeed a metric. It is also worth noting that all the results in this section remain valid if the state space S is an arbitrary countable set, μ is a countably additive measure on S, and $\mathscr{P}$ represents the space of all partitions of S with countably many (measurable) cells.

Let S be a finite set, and $\mathscr{P}$ be the space of all partitions of S. Let μ be a probability measure on S. Define the *entropy* of the partition $P \in \mathscr{P}$ as

$$H(P) := -\sum_{J\in P} \mu(J) \log \mu(J)$$

(7) As always, W_0 also denotes the linear extension of W_0 to $\Delta(L)$.

Let $\geq$ be a partial order on $\mathscr{P}$, wherein $P \geq Q$ if P is coarser than Q (or equivalently, Q is finer than P). We shall say that $P > Q$ if $P \geq Q$ and $P \neq Q$.

We may also define the coarsest refinement of P and Q, denoted by $P \wedge Q$. If $P = (I_m)$ and $Q = (J_n)$, then $P \wedge Q = (I_m \cap J_n)_{m,n}$, so

$$H(P \wedge Q) = -\sum_m \sum_n \mu(I_m \cap J_n) \log \big(\mu(I_m \cap J_n)\big)$$

Similarly, $P \vee Q$ is the finest partition coarser than P and Q. Then, $(\mathscr{P}, \geq, \vee, \wedge)$ is a *lattice*, with greatest (coarsest) element $\{S\}$, and least (finest) element $\big\{\{s\} : s \in S\big\}$. Notice that $H(\{S\}) = 0$, while $H(P) > 0$ for all other partitions P. Define the *conditional entropy* $H(P \mid Q)$ as

$$H(P \mid Q) := H(P \wedge Q) - H(Q)$$

It is easy to see that

$$H(P \mid Q) = -\sum_n \mu(J_n) \sum_m \frac{\mu(I_m \cap J_n)}{\mu(J_n)} \log \left(\frac{\mu(I_m \cap J_n)}{\mu(J_n)}\right)$$

We now come to the main result of this section.

Proposition 5.1. The function

$$\mathsf{d}(P, Q) := 2H(P \wedge Q) - H(P) - H(Q) = H(P \mid Q) + H(Q \mid P)$$

is a metric on $\mathscr{P}$.

We begin with some lemmata.

Lemma 5.2. H is anti-monotone, ie, $P \geq Q$ implies $H(Q) \geq H(P)$. Moreover, H is strictly anti-monotone, ie, $P > Q$ implies $H(Q) > H(P)$.

The proof is trivial and is omitted.

Lemma 5.3. The function $H(P \mid Q)$ is anti-monotone in P, and is monotone in Q.

Proof. Notice that if $P' \geq P$, then $P' \wedge Q \geq P \wedge Q$, so the anti-monotonicity of H implies that $H(P \mid Q)$ is anti-monotone in P. We say that Q is an *elementary refinement* of Q' if $Q' = \{J_1, \ldots, J_N\}$ and $Q = \{\tilde{J}_1, \ldots, \tilde{J}_{n-1}, \tilde{J}_N, \tilde{J}_{N+1}\}$, where $\tilde{J}_n := J_n$ for all $n = 1, \ldots, N-1$, while $J_N = \tilde{J}_N \cup \tilde{J}_{N+1}$. In other words, Q and Q' are identical except that there exists a cell $J_N \in Q'$ that is the union of exactly two cells in Q.

Let $Q' \geq Q$. Then, there exist $Q_1, \ldots, Q_k \in \mathscr{P}$ such that $Q' = Q_k \geq Q_{k-1} \geq \cdots \geq Q_1 = Q$, and where Q_i is an elementary refinement of Q_{i+1}. Thus, in order to show that $H(P \mid Q)$ is monotone in Q, it suffices to consider Q and Q' where Q is an elementary refinement of Q'.

Let $P = \{I_1, \ldots, I_M\}$ and Q and Q' be as above. In what follows, we shall let $\eta(x) = x \log x$ for all $x > 0$ and $\eta(0) = 0$. Then $\eta \in \mathbb{R}^{\mathbb{R}_+}$ is strictly convex and continuous on its domain. Let

$$\Lambda = -\sum_{n=1}^{N-1} \mu(J_n) \sum_m \eta\left(\frac{\mu(I_m \cap J_n)}{\mu(J_n)}\right)$$

This allows us to write

$$\begin{aligned}
H(P \mid Q) &= -\sum_n \mu(\tilde{J}_n) \sum_m \frac{\mu(I_m \cap \tilde{J}_n)}{\mu(\tilde{J}_n)} \log\left(\frac{\mu(I_m \cap \tilde{J}_n)}{\mu(\tilde{J}_n)}\right) \\
&= \Lambda - \sum_{n=N,N+1} \mu(\tilde{J}_n) \sum_m \eta\left(\frac{\mu(I_m \cap \tilde{J}_n)}{\mu(\tilde{J}_n)}\right) \\
&= \Lambda - \mu(J_N) \sum_m \sum_{n=N,N+1} \frac{\mu(\tilde{J}_n)}{\mu(J_N)} \eta\left(\frac{\mu(I_m \cap \tilde{J}_n)}{\mu(\tilde{J}_n)}\right) \\
&\leq \Lambda - \mu(J_N) \sum_m \eta\left(\sum_{n=N,N+1} \frac{\mu(I_m \cap \tilde{J}_n)}{\mu(J_N)}\right) \\
&= \Lambda - \mu(J_N) \sum_m \eta\left(\frac{\mu(I_m \cap J_n)}{\mu(J_n)}\right) \\
&= H(P \mid Q')
\end{aligned}$$

where we have used the fact that $-\eta$ is concave to establish the inequality. □

Lemma 5.4. The function H is submodular, ie, $H(P \wedge Q) + H(P \vee Q) \leq H(P) + H(Q)$.

Proof. Fix P and Q, and let $Q \leq Q'$. We shall use the fact that the function $H(P \mid Q)$ is anti-monotone in P and monotone in Q. Then, $H(P \wedge Q) - H(Q) = H(P \mid Q) \leq H(P \mid Q') = H(P \wedge Q') - H(Q')$. Now set, $Q' := P \vee Q$, so that $P \wedge (P \vee Q) = P$, which implies $H(P \wedge Q) - H(Q) \leq H(P) - H(P \vee Q)$. Therefore, H is submodular. □

We now list some properties of the lattice $(\mathscr{P}, \geq, \vee, \wedge)$.

Lemma 5.5. For $P, Q, R \in \mathscr{P}$, the following hold:
(a) $R \geq (P \wedge R) \vee (Q \wedge R)$.
(b) $(P \vee Q) \wedge R \geq (P \wedge R) \vee (Q \wedge R)$.

Proof. Note that $R \geq P \wedge R$ and $R \geq Q \wedge R$, so $R \geq (P \wedge R) \vee (Q \wedge R)$, which establishes (a). To see (b), note that $P \geq P \wedge R$, while $Q \geq Q \wedge R$. Therefore, $P \vee Q \geq (P \wedge R) \vee (Q \wedge R)$. But we also have that $R \geq (P \wedge R) \vee (Q \wedge R)$, from (a). The definition of $\wedge$ then implies that $(P \vee Q) \wedge R \geq (P \wedge R) \vee (Q \wedge R)$, as required. □

Proof of Proposition 5.1. The proof relies on the fact that conditional entropy $H(P \mid Q)$ is anti-monotone (Lemmas 5.2 and 5.3) and submodular (Lemma 5.4). Because H is anti-monotone (Lemma 5.2), $\mathsf{d}(P, Q) \geq 0$ for all P, Q. We have already established that $P < Q$ implies $H(P) > H(Q)$. If P and Q are distinct, then $P \wedge Q$ is distinct from either P or Q, so that $\mathsf{d}(P, Q) > 0$.

It is easy to see that $\mathsf{d}(P, Q) \leq \mathsf{d}(P, R) + \mathsf{d}(R, Q)$ if, and only if,

$$H(P \wedge Q) + H(R) \leq H(P \wedge R) + H(Q \wedge R) \qquad [\heartsuit]$$

By Lemma 5.5, we see that $R \geq (P \wedge R) \vee (Q \wedge R)$ and $(P \vee Q) \wedge R \geq (P \wedge R) \vee (Q \wedge R)$. Set $P' = P \wedge R$ and $Q' = Q \wedge R$. The submodularity of H implies $H(P' \vee Q) + H(P' \wedge Q') \leq H(P') + H(Q')$. That is, $H\big((P \wedge R) \vee (Q \wedge R)\big) + H(P \wedge Q \wedge R) \leq H(P \wedge R) + H(Q \wedge R)$. But $R \geq (P \wedge R) \vee (Q \wedge R)$, so $H(R) \leq H\big((P \wedge R) \vee (Q \wedge R)\big)$. Similarly, $P \wedge Q \geq P \wedge Q \wedge R$, which implies $H(P \wedge Q) \leq H(P \wedge Q \wedge R)$. These observations imply [♡], so that d is a metric. □

References

Aliprantis, Charalambos, and Kim C. Border. 1999. *Infinite Dimensional Analysis: A Hitchhiker's Guide, 2/e*. Springer, New York. (Cit. on pp. 9, 15, 16).

Dekel, Eddie, Barton L. Lipman, Aldo Rustichini, and Todd Sarver. 2007. "Representing Preferences with a Unique Subjective State Space: A Corrigendum". *Econometrica* 75 (2): 591–600. (Cit. on p. 15).

Dillenberger, David, R. Vijay Krishna, and Philipp Sadowski. 2017. *Subjective Information Choice Processes*. Tech. rep. Duke University. (Cit. on p. 1).

Ekeland, Ivar, and Thomas Turnbull. 1983. *Infinite-Dimensional Optimization and Convexity*. Chicago: University of Chicago Press. (Cit. on pp. 6, 7).

Epstein, Larry G, and Stanley E Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework". *Econometrica* 57 (4): 937–969. (Cit. on p. 2).

Ergin, Haluk, and Todd Sarver. 2010a. "A Unique Costly Contemplation Representation". *Econometrica* 78 (4): 1285–1339. (Cit. on p. 11).

– . 2010b. "Supplement to 'A Unique Costly Contemplation Representation'". *Econometrica Supplemental Material* 78. (Cit. on p. 17).

Gale, David. 1967. "A Geometric Duality Theorem with Economic Applications". *Review of Economic Studies* 34 (1): 19–24. (Cit. on pp. 6, 7).

Gilboa, Itzhak, and David Schmeidler. 1989. "Maxmin Expected Utility with a Non-Unique Prior". *Journal of Mathematical Economics* 18:141–153. (Cit. on p. 17).

Krishna, R. Vijay, and Philipp Sadowski. 2014. “Dynamic Preference for Flexibility”. *Econometrica* 82 (2): 655–703. (Cit. on pp. 1, 38).

Ok, Efe A. 2007. *Real Analysis with Economic Applications*. Princeton University Press. (Cit. on p. 24).