Information-dependent Utilities and Beliefs*

David Dillenberger† R. Vijay Krishna‡ Philipp Sadowski§

April 21, 2020 at 23:53 hrs

Abstract

We axiomatize a model of preferences over menus of acts in which not only beliefs but also state-dependent utilities depend on the individual's choice of information. Our most general model features both contemplation about the appropriate way to evaluate alternatives as well as acquisition of information about the payoff relevant state of the world, before a choice is made. We then focus on the special case where the value of alternatives depends directly and exclusively on the state of the world and on the choice of information about that state.

Key Words: Contemplation, Information Acquisition

JEL Classification: D80, D81, D90

1. Introduction

We study an individual who has preference over menus of acts, defined on some state space S. We begin by axiomatizing a representation, according to which the value of a menu x is given by

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

(*) We thank the National Science Foundation for support under awards SES 1461469 and SES 1461502. This work began while Dillenberger and Krishna were visiting the Economics departments at NYU and Duke University, respectively. They are most grateful to these institutions for their hospitality.

(†) University of Pennsylvania <ddill@sas.upenn.edu>

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

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

The representation [$\blacklozenge$] features *simultaneous* information acquisition about the state of the world $s \in S$ and contemplation about how to evaluate alternatives via a utility $\mathfrak{u}$, that is randomly drawn from a set of possible state-dependent utilities $\mathfrak{U}$. The interpretation is that the individual chooses a joint distribution over beliefs and utilities from a specific feasible set, and for each realized pair chooses the act that maximizes the corresponding expected utility. This is a constrained optimization problem, with the constraint being specified by properties of the set $\mathfrak{M}$. This model unifies the approach of de Oliveira et al. (2017) that study information uncertainty and that of Ergin and Sarver (2010a) that study uncertainty about future tastes. In Section 3 we characterize this representation, highlighting the way in which it behaviorally relaxes each of the aforementioned models.

In Section 4 we proceed to study a special case of [$\blacklozenge$], where the individual can choose only a partition P of S, prior to choosing an act. Any act f has two components, f_1 and f_2. Upon learning the cell $J \in P$, the individual updates the prior π_0 according to Bayes' law. The value of any set of acts is given by

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

The representation suggests that the choice of information about the state of the world directly affects the individual's value for the second component of consumption, the outcome of f_2. Below we briefly discuss two examples of motives that could be accommodated by this representation.

Example 1.1. Future consumption choice: The second consumption component is itself a consumption choice problem for a not explicitly modeled future, and today's information choice may affect the value of such continuation problems by affecting the future information constraint. This is our leading interpretation and we will refer to it when discussing the formal model. It is further developed in Dillenberger, Krishna, and Sadowski (2020), where the present static model of menu choice is the starting point for an infinite horizon model of repeated choice with intertemporal information constraints.

Example 1.2. Vindication or Repudiation: The individual may feel repudiated if the chosen information was "misleading" in terms of the second consumption component and may feel vindicated if it was not. To give one simple example, suppose the value of vindication or repudiation depends only on the potential value of information about the second component, independently of actual consumption utility. Specifically, for

$$\Gamma(s, P, x) := \arg\max_{f \in x} \sum_{J \in P} \sum_{s' \in J} \mathbf{1}_{\{s \in J\}} \pi_0(s') v(f_2(s'))$$

where $\mathbf{1}_A$ is 1 if the set A obtains and 0 otherwise. Define

$$v_s^+ = \max_{Q \in \mathfrak{M}} \max_{f \in \Gamma(s,Q,x)} v(f_2(s))$$

$$v_s^- = \min_{Q \in \mathfrak{M}} \max_{f \in \Gamma(s,Q,x)} v(f_2(s))$$

$$v_s^P = \max_{f \in \Gamma(s,P,x)} v(f_2(s))$$

and let the amount of vindication or repudiation from choosing $P \in \mathfrak{M}$ when the choice set is x and the state turns out to be s be

$$r_s(P) = \frac{v_s^P - \frac{v_s^+ + v_s^-}{2}}{v_s^+ - v_s^-}$$

Note that $r_s(P)$ takes values in $[-1/2, 1/2]$, so that positive values correspond to vindication. The following separable value function then captures direct utility from the second consumption component and vindication or repudiation of the choice of P for the second component:

$$v_s(f_2(s), P) = v(f_2(s)) + \beta r_s(P)$$

2. Domain

Let the consumption space Y be a compact metric space. The space of lotteries over Y is denoted by $\Delta(Y)$. Let S be a finite space of states, and $\mathscr{F}\big(\Delta(Y)\big)$ denote the space of all acts from S that realise a lottery over Y. Let $X := \mathscr{K}\Big(\mathscr{F}\big(\Delta(Y)\big)\Big)$ denote the space of all closed subsets of $\mathscr{F}\big(\Delta(Y)\big)$; a *menu* is an element $x \in X$. A *preference* is a binary relation $\succsim$ on X.

Of special interest is a subspace of the domain $L := \mathscr{F}(\Delta(Y_L))$, where $Y_L \subset Y$ is closed and convex. Thus, L consists only of singleton menus; we will assume (in terms of axioms and the representation) that the value of elements $\ell \in L$ is not uncertain.

In Section 4 we will consider the special case $Y = C \times W$ where C and W are compact metric spaces, and where W is also convex.[1] When the second component corresponds to continuation choice problems as in Example 1.1, then $W = \mathscr{K}(\mathscr{F}(\Delta(C)))$, namely, the space of second-stage consumption choice problems, and $L = \mathscr{F}(\Delta(C \times \mathscr{F}(\Delta(C))))$.

(1) In other words, W can be embedded linearly in a vector space.

3. Information Acquisition and Contemplation

In this section we present axioms on the preference $\succsim$ over X that are equivalent to a representation with information acquisition and contemplation.

3.1. Standard Properties

Our first axiom collects basic properties of $\succsim$ that are common in the menu-choice literature.

Axiom 1 (Basic Properties)**.**

(a) Order: $\succsim$ is non-trivial, complete, and transitive.

(b) Continuity: The sets $\{y : x \succsim y\}$ and $\{y : y \succsim x\}$ are closed for each $x \in \mathscr{F}(\Delta(C \times W))$.

(c) Lipschitz Continuity: There exist $\ell^\sharp, \ell_\sharp \in L$ and $N > 0$ such that for all $x, y \in X$ and $t \in (0,1)$ with $t \geqslant Nd(x,y)$, we have $(1-t)x + t\ell^\sharp \succ (1-t)y + t\ell_\sharp$.

(d) Monotonicity: $x \cup y \succsim x$ for all $x, y \in X$.

(e) Aversion to Randomization: If $x \sim y$, then $x \succsim \frac{1}{2}x + \frac{1}{2}y$ for all $x, y \in X$.

Items (a)–(d) are standard.[2] Item (e) is familiar from Ergin and Sarver (2010a) and de Oliveira et al. (2017) and relaxes Independence in order to accommodate unobserved information choice.

The next axiom captures the special role played by the subdomain L, for which the consumption value bears no uncertainty, and hence does not benefit from either information acquisition about the objective state $s \in S$ or from contemplation about the subjective taste.

Axiom 2 (L-independence)**.** For all $x, y \in X$, $t \in (0,1]$, and $\ell \in L$, $x \succ y$ implies $tx + (1-t)\ell \succ ty + (1-t)\ell$.

Axiom 2 is closely related to the C-Independence axiom in Gilboa and Schmeidler (1989), and is motivated in a similar fashion: Because consumption streams require no information acquisition or contemplation, mixing two menus with the same consumption stream should not alter the ranking between these menus.

(2) For a discussion of (c) see Dekel et al. (2007) and for (d) see Kreps (1979).

3.2. Representation of Information Acquisition and Contemplation

Recall that $\mathbf{C}(Y)$ is the space of all uniformly continuous functions on the compact metric space Y and for $\alpha \in \Delta(Y)$ and $\mathfrak{u} \in \mathbf{C}(Y)$, $\mathfrak{u}(\alpha) := \int_Y \mathfrak{u}(y)\,\mathrm{d}\alpha(y) =: \langle \alpha, \mathfrak{u} \rangle$; endowed with the supremum norm, $\mathbf{C}(Y)$ is a Banach space. Fix $\ell^\dagger \in L$, and define $\mathfrak{U}_s := \{\mathfrak{u}_s \in C(Y) : \mathfrak{u}_s(\ell_s^\dagger) = 0, \|\mathfrak{u}_s\|_\infty = 1\}$. Finally, define $\mathfrak{U} := \big\{(p_1\mathfrak{u}_1, \ldots, p_n\mathfrak{u}_n) : \mathfrak{u}_s \in \mathfrak{U}_s,\ p_s \geqslant 0, \sum_s p_s = 1\big\}$. The space $\mathfrak{U}$ will serve as our *subjective state space* below. It is useful to reconsider $\mathfrak{U}$ as $\mathfrak{U} := \big\{(p, \mathfrak{u}) : p := (p_1, \ldots, p_n) \in \Delta(S),\ \mathfrak{u} := (\mathfrak{u}_1, \ldots, \mathfrak{u}_n) \in \bigtimes_{s \in S} \mathfrak{U}_s\big\}$.

Theorem 1. *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).*

(b) *There exists a metric space of continuous functions $\mathfrak{U}$ (as defined above) and a (unique) minimal[3] 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*

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

represents $\succsim$.

The proof of Theorem 1 is in Appendix 5.

4. Information-Dependent Consumption Values

We now pose additional axioms that are plausible if contemplation about the taste is not independent of information acquisition about $s \in S$, but rather consumption values depend directly on that information choice. To that end, let C and W be compact metric spaces, where W is also convex, and $Z \subset W$ also closed and convex. Let the space of prizes now be $Y = C \times W$, and let $Y_L := C \times Z$, so that $L = \mathcal{F}(\Delta(C \times Z))$. L-Independence (Axiom 2) suggests that the value of consuming elements in L is independent of the choice of information.[5]

(3) $\mathfrak{M}$ is a *minimal* set of charges if any larger set of charges gives the same utility for every x.

(4) A charge is a finitely additive measure.

(5) Recall Example 1.1, where $W = \mathcal{K}(\mathcal{F}(\Delta(C)))$ consists of consumption choice problems for the future, and the continuation utility v_s depends on the choice of P through its effect on the

4.1. No Complementarities

For tractability, we assume that there are no complementarities between consumption dimensions C and W. This is satisfied, for example, if consumption utility is quasilinear in C. Formally, DM's value for a menu does not change when substituting act f with g as long as they induce, on each state s, the same marginal distributions over C and W. For any $f \in \mathcal{F}(\Delta(C \times W))$, we denote by $f_1(s)$ and $f_2(s)$ the marginals of $f(s)$ on C and W, respectively.

Axiom 3 (State-Contingent Indifference to Correlation)**.** For a finite menu x, if $f \in x$ and $g \in \mathcal{F}(\Delta(C \times W))$ are such that $g_1(s) = f_1(s)$ and $g_2(s) = f_2(s)$ for all $s \in S$, then $[(x \backslash \{f\}) \cup \{g\}] \sim x$.[6]

4.2. Indifference to Incentivized Contingent Commitment

Let $\ell_*, \ell^* \in L$ be the $\succsim$-worst and -best members of L, respectively. Suppose the worst element of Y corresponds to receiving, in every state s, the worst outcome c_s^- in C as well as a particular outcome in Z, which we denote by z_s^-, so that $\ell_* = (c_s^-, z_s^-) \in L$. Analogously, suppose $\ell^* = (c_s^+, z_s^+) \in L$ is the best element in Y. For instance, under our leading interpretation, the worst outcome in state s is consumption c_s^- paired with the act that delivers the worst outcome in C in every state in the continuation stage, so that indeed $\ell_* = (c_s^-, z_s^-) \in L$.

Suppose, further, that DM is offered a chance to replace a certain choice problem with another. DM's attitude towards such replacements may depend on his information choice, which is subjective, unobserved, and menu-dependent. That said, any strategy of choice from a menu gives rise to a consumption act. Therefore, any choice problem y should leave DM no worse off than receiving ℓ_*, and no better off than receiving ℓ^*. Since ℓ^* leaves DM strictly better off than ℓ_* in every state, the optimal choice from a menu $(1-t)x + t\ell^*$ should generate an outcome that is also strictly better than ℓ_*.

For each $I \subset S$, $f \in \mathcal{F}(\Delta(C \times W))$, $(c, w) \in C \times W$, and $\varepsilon \in [0, 1]$, define

---

(unmodeled) future information constraint. Since $Z = \mathcal{F}(\Delta(C))$ contains acts (singleton menus) that do not require any choice in the future, the value of elements in $L = \mathcal{F}(\Delta(C \times Z))$ should indeed be independent of the choice of P.

(6) Axiom 3 is closely related to Axiom 5 in Krishna and Sadowski (2014), where other related notions of separability are also mentioned. The important difference is that Axiom 3 requires indifference to correlation in *any* menu x, rather than just singletons, because different information may be optimal for different menus.

$f \oplus_{\varepsilon,I} (c,w) \in \mathscr{F}(\Delta(C \times W))$ by

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

That is, for any state $s \in I$, the act $f \oplus_{\varepsilon,I} (c,w)$ perturbs the outcome with (c,w).

Fix $c \in C$ and let $\ell_s := \ell_* \oplus_{1,s} (c,z) \in L$, and define the induced binary relation $\succsim_s$ on Z by $z \succsim_s z'$ if $\ell_s \succsim \ell'_s$.

Let $X^* := \{(1-t)\,x + t\ell^* : x \in X \text{ is finite}, t \in (0,1)\}$. For a mapping $e : x \to (0,1]$, let $x \oplus_{e,s} w := \big\{f \oplus_{e(f),s} (c_s^-, w) : f \in x\big\}$, which perturbs the continuation lottery in state s for any act f in x by giving weight $e(f)$ to (c_s^-, w). For $x \in X^*$ we then require $x \succ [x \oplus_{e,s} z_s^-]$ and $[x \oplus_{e,s} w] \succsim [x \oplus_{e,s} z_s^-]$ for all $s \in S$ and $w \in W$. This is part (a) of Axiom 4 below.

Part (b) investigates the conditions under which DM is actually *indifferent* to replacing continuation lotteries with the worst consumption outcome. The idea is that there should be a contingent plan that specifies which act DM will choose for each state, such that he will be indifferent between the original menu and one where he is penalized whenever his choice does not coincide with that plan.

To formalize this state contingent notion of strategic rationality, we define the set of contingent plans Ξ_x to be the collection of all functions $\xi : S \to x$. An *Incentivized Contingent Commitment* to $\xi \in \Xi_x$, is then the set

$$\mathscr{I}(\xi) = \big\{f \oplus_{1,I^c} (c_s^-, z_s^-) : f \in x \text{ and } I = \{s : f = \xi(s)\}\big\}$$

which replaces the outcome of f with the worst outcome (c_s^-, z_s^-) in any state where f should not be chosen according to ξ. Obviously $x \succsim \mathscr{I}(\xi)$ for all $\xi \in \Xi_x$. However, if for no $s \in S$ is it ever optimal to choose an act outside $\xi(s)$, then $x \sim \mathscr{I}(\xi)$ should hold.

Axiom 4 (Indifference to Incentivized Contingent Commitment)**.**

(a) If $x \in X^*$ and $e : x \to (0,1]$, then $x \succ [x \oplus_{e,s} z_s^-]$ and $[x \oplus_{e,s} w] \succsim [x \oplus_{e,s} z_s^-]$ for all $s \in S$ and $w \in W$.

(b) For all $x \in X$, there is $\xi \in \Xi_x$ such that $x \sim \mathscr{I}(\xi)$.

4.3. Concordant Independence

We aim to capture a situation where the choice of partition and the actual realization of the payoff-relevant state fully determine the value for the W component of consumption.

We say that x and y are *concordant* if the same information choice is optimal for both x and y. We argue that, if x and y are concordant, then both should be concordant with the convex combination $\frac{1}{2}x + \frac{1}{2}y$. While Independence may be violated when considering menus that lead to different optimal initial information choices, $\gtrsim|_{X'}$ should satisfy Independence if $X' \subset X$ consists only of concordant menus. That is, any violation of Independence is entirely due to a change in the choice of information. We now introduce our behavioral notion of concordance (Definition 4.1 below).

For each $J \in P$, let

$$f_J(s) = \begin{cases} (c_s^+, z_s^+) & s \in J \\ (c_s^-, z_s^-) & \text{otherwise} \end{cases}$$

and define the menu $x_1(P) := \{f_J : J \in P\}$. Then, $x_1(P) \sim \ell^*$ if, and only if, $x_1(P)$ is evaluated under a partition that is at least as fine as P.

For a choice problem x, we then have $\frac{1}{2}x + \frac{1}{2}x_1(P) \sim \frac{1}{2}x + \frac{1}{2}\ell^*$ if, and only if, some partition that is at least as fine as P is optimal for x. Thus, the same collection of partitions is optimal for two menus x and y, if for all $P \in \mathscr{P}$ we have $\frac{1}{2}x + \frac{1}{2}x_1(P) \sim \frac{1}{2}x + \frac{1}{2}\ell^*$ if, and only if, $\frac{1}{2}y + \frac{1}{2}x_1(P) \sim \frac{1}{2}y + \frac{1}{2}\ell^*$.[7]

Definition 4.1. Choice problems x and y are *concordant*, if $\frac{1}{2}x + \frac{1}{2}x_1(P) \sim \frac{1}{2}x + \frac{1}{2}\ell^*$ if and only if $\frac{1}{2}y + \frac{1}{2}x_1(P) \sim \frac{1}{2}y + \frac{1}{2}\ell^*$ for all $P \in \mathscr{P}$.

Axiom 5 (Concordant Independence)**.** If x and y are concordant, so are x and $\frac{1}{2}x + \frac{1}{2}y$. Furthermore, if $X' \subset X$ consists of pairwise concordant menus, then $\gtrsim|_{X'}$ satisfies Independence.[8]

4.4. Representation with Information Dependent Consumption Values

Theorem 2. *A binary relation $\gtrsim$ on X satisfies Axioms 1–5 if, and only if, there exists a function $V : X \to \mathbb{R}$ that represents $\gtrsim$ 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 $v_s(\cdot, P) \in \mathbf{C}(W)$ for each $s \in S$ and $P \in \mathfrak{M}_p^\sharp$, with the property that for all $P, P' \in \mathfrak{M}_p^\sharp$, $s \in S$, $v_s(\cdot, P)|_L = v_s(\cdot, P')|_L$, and $v_s(w, \cdot) \geqslant v_s(z_s^-, \cdot)$ for all $w \in W$.

(7) See Lemma 6.19 in Appendix 6.4 for a instantiation of this intuition.
(8) If $x, y, z, (1-t)\,x + tz, (1-t)\,y + tz \in X'$, $t \in (0,1)$, and $x \succ y$, then $(1-t)\,x + tz \succ (1-t)\,y + tz$.

The theorem suggests that the decision maker does not independently contemplate his taste, but rather that the choice of information about the state of the world directly affects his value for (the W component of) consumption. For instance, a positive outcome may be valued higher if it was generated by a well informed choice than by pure luck. As another example, the decision maker may directly experience the C component of consumption once the state is realized, but derives a value from the W component prior to the realized state and dependent directly on the information event $J \in P$ with $s \in J$.

5. Appendix: Proof of Theorem 1

5.1. Algebraic Representation

Proposition 5.1. 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).

(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 thus omitted. The remainder of this section proves that (a) implies (b). We shall first show that under our assumptions, every closed subset is indifferent to its closed convex hull.

Lemma 5.2. If $\succsim$ satisfies Basic Properties (Axiom 1), then for each $x \in \mathscr{K}(Z)$, $x \sim \operatorname{cch}(x)$.

Proof. First consider $x \in X$ that is finite and follow Ergin and Sarver (2010a, Lemma 2). Notice that $\operatorname{cch}(x) \succsim x$ by Monotonicity (Axiom 1(d)). Let $x^0 := x$, and for each $k \geqslant 1$, 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 $k \geqslant 1$. But notice that $d(x^k, \operatorname{cch}(x)) \to 0$ as $k \to \infty$. Therefore, by Continuity (Axiom 1(b)), it follows that $x \succsim \operatorname{cch}(x)$, which proves that $x \sim \operatorname{cch}(x)$ 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 $x_m \sim \operatorname{cch}(x_m)$. It is also easy to see that $d(\operatorname{cch}(x), \operatorname{cch}(x_m)) \to 0$ as $m \to \infty$. Continuity (Axiom 1(b)) now implies that $x \sim \operatorname{cch}(x)$, which proves the claim. □

In light of Lemma 5.2, in what follows, we may restrict attention to the space $\mathscr{K}_c(X)$.

Lemma 5.3. If $\succsim$ satisfies Order (Axiom 1(a)), Continuity (Axiom 1(b)), and L-Independence (Axiom 2), 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, $\zeta(\ell) \geqslant \zeta(\ell')$. Moreover, ζ is unique up to positive affine transformation.

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

Corollary 5.4. 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 5.5. Given the function $\zeta : L \to \mathbb{R}$ from lemma 5.3 above, there exists $V : X \to \mathbb{R}$ such that

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

Proof. Let ℓ^* be a $\succsim$-best and ℓ_* a $\succsim$-worst element of L. By Corollary 5.4, $\ell^* \succ \ell_*$. 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 $ax + (1-a)\ell_* \sim \ell$, by L-Independence (Axiom 2) 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

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

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 5.6. 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 $\varepsilon \leqslant (1-\bar{t})$ 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 $x \sim y$. □

The next Corollary follows immediately from L-Independence and Lemma 5.6.

Corollary 5.7. $x \succ y$ if, and only if, $tx + (1-t)\ell \succ ty + (1-t)\ell$ for all $t \in (0,1]$.

Lemma 5.8. $\ell \succ \ell'$ if, and only if, $tx + (1-t)\ell \succ tx + (1-t)\ell'$ for all $t \in [0,1)$.

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 5.7 repeatedly yields that $\ell \succ \ell'$ if, and only if, $t'\left[\alpha x + (1-\alpha)\ell_*\right] + (1-t')\ell \sim t'\overline{\ell} + (1-t')\ell \succ t'\overline{\ell} + (1-t')\ell' \sim t'\left[\alpha x + (1-\alpha)\ell_*\right] + (1-t')\ell'$ for all $t' \in (0,1)$. Again by Corollary 5.7, 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. □

Lemma 5.9. The function V defined in the proof of Lemma 5.5 has the following properties:

(a) V is monotone, ie, $V(x \cup y) \geqslant V(x)$ 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 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 5.8 yields $t\ell_* + (1-t)\ell \succsim \ell_*$, and then, by Corollary 5.7, $tx + (1-t)\ell \succ \ell_*$. By continuity, there are $\alpha \in (0,1)$ and $\overline{\ell}$, such that $\ell^* \succ \alpha\left(tx + (1-t)\ell\right) + (1-\alpha)\ell_* \sim \overline{\ell} \succ \ell_*$. 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 5.7 and the definition of V it is easy to verify that $V\left(tx + (1-t)\ell\right) = \frac{\gamma}{\alpha}$ (independent of whether $tx + (1-t)\ell \succsim \ell^*$ or not). Next, by Lemma 5.8, $tx + (1-t)\ell \sim tx + (1-t)\left(\beta\ell^* + (1-\beta)\ell_*\right)$. Then, by Corollary 5.7,

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

or

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

Because $x \succ \ell^*$, Corollary 5.7 and Lemma 5.8 further imply that $\alpha(1-t)(1-\beta) + (1-\alpha) \succ (1-\gamma)$ or $\gamma - \alpha(1-t)\beta \succ \alpha t > 0$. This implies that $\gamma > \alpha(1-t)\beta$.

Corollary 5.7 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 5.7,

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

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) We first show that V is midpoint convex, ie, $V\big(\frac{1}{2}x_1 + \frac{1}{2}x_2\big) \leqslant \frac{1}{2}V(x_1) + \frac{1}{2}V(x_2)$. 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 $V(\frac{1}{2}x_1 + \frac{1}{2}x_2) \leqslant \frac{1}{2}V(x_1) + \frac{1}{2}V(x_2)$. 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 $V(\bar{x}) \leqslant \frac{\lambda}{1+\lambda}V(x_1) + \frac{1}{1+\lambda}V(y)$ 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 $V(\frac{1}{2}x_1 + \frac{1}{2}x_2) \leqslant \frac{1}{2}V(x_1) + \frac{1}{2}V(x_2)$, 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)), $V(\bar{x}) \leqslant \frac{a}{1+a}V(x_2) + \frac{1}{1+a}V(y)$. We also have $\frac{a}{1+a}V(x_2) + \frac{1}{1+a}V(y) = \frac{a}{1+a}\big(V(x_1) + V(x_2)\big) + \frac{1-a}{1+a}V(\ell_*)$. Substituting in the value of $V(\bar{x})$ obtained above, we see that $V(\frac{1}{2}x_1 + \frac{1}{2}x_2) \leqslant \frac{1}{2}V(x_1) + \frac{1}{2}V(x_2)$, as claimed.
As noted above, V is continuous, and because it is midpoint convex, it is convex. $\square$

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

Lemma 5.10. If $\gtrsim$ 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 $\gtrsim$, 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 $0 < d(x, y) \leqslant \beta$ 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 $V(y) - V(x) < \dfrac{\alpha}{1-\alpha}\big[V(\ell^\sharp) - V(\ell_\sharp)\big]$. But notice that $\alpha/N \leqslant \beta$, 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 $(\lambda_{j+1} - \lambda_j)d(x, y) \leqslant \beta$ for all $j = 0, \ldots, J$. 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 $V(x_{j+1}) - V(x_j) \leqslant Kd(x_{j+1}, x_j) = K(\lambda_{j+1} - \lambda_j)d(x, y)$. Summing over j, we find $V(y) - V(x) \leqslant Kd(x, y)$. Interchanging the roles of x and y, it follows that $|V(x) - V(y)| \leqslant Kd(x, y)$, as claimed. The converse is as in Dekel et al. (2007) and is omitted. □

Because V is L-affine, ie, $V = \zeta$ on L, it follows from Lemma 5.3 that V is unique up to positive affine transformation. This proves (a) implies (b), which establishes Proposition 5.1.

5.2. Abstract Convex and Monotone Representation

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

Therefore, a function $H : \mathfrak{U} \to \mathbb{R}$ is a support function if its unique extension to $\operatorname{span}(\mathfrak{U})$ by positive homogeneity is sublinear and norm continuous.

Given a function $H : \mathfrak{U} \to \mathbb{R}$ whose extension to $\operatorname{span}(\mathfrak{U})$ by positive homogeneity is sublinear and norm continuous, we may define $x_H := \{f \in \operatorname{aff}\big(\mathcal{F}(\Delta(C \times W))\big) : \mathfrak{u}(f) \leqslant H(\mathfrak{u}) \text{ for all } \mathfrak{u} \in \mathfrak{U}\}$. Support functions enjoy the following duality: For any weak* compact, convex subset x of $\operatorname{aff}\big(\mathcal{F}(\Delta(C \times W))\big)$, $x_{H_x} = x$, 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, $H_x \leqslant H_{x'}$, (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) $H_{\operatorname{ch}(x \cup x')} = H_x \vee H_{x'}$. (By Lemma 5.14 of Aliprantis and Border (1999), $\operatorname{ch}(x \cup x')$ is compact because x and x' are compact, which ensures that $H_{\operatorname{ch}(x \cup x')}$ is well defined.) Finally, observe that for $\ell^\dagger := \ell_i^\dagger \times \cdots \times \ell_n^\dagger$, $H_{\ell^\dagger} = \mathbf{0}$.

Proposition 5.11. 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

$$V(x) = \max_{\mu \in \mathfrak{M}} \left[\int_{\mathfrak{U}} \max_{f \in x} \sum_s p_s \mathfrak{u}_s(f(s))\, \mathrm{d}\mu(p, \mathfrak{u}) \right] \tag{5.1}$$

where the set $\mathfrak{M} \subset ba_{\mathrm{n}}(\mathfrak{U})$ is weak* compact and $\int_{\mathfrak{U}} \max_{f \in x} \sum_s p_s \mathfrak{u}_s(f(s))\, \mathrm{d}\mu(p, \mathfrak{u})$ is independent of μ for all $x \in L$.[9] Moreover, for a dense set of points in X, there is a unique $\mu \in \mathfrak{M}$ that achieves the maximum in [5.1].

In Proposition 5.11 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 $V(x) > \max_{\mu \in \mathcal{N}} \big[\int_{\mathfrak{U}} \max_{f \in x} \sum_i p_i \mathfrak{u}_i(f(i))\, \mathrm{d}\mu(p, \mathfrak{u})\big]$.

Proof. By Lemma 5.2, for every $x \in X$, $V(x) = V(\operatorname{cch}(x))$. Therefore, we may restrict attention to convex menus.

Let $\Psi : \mathcal{K}_c(\mathcal{F}(\Delta(C \times W))) \to \mathbf{C}_b(\mathfrak{U})$ be the map that associates each compact, convex subset x of $\mathcal{F}(\Delta(C \times W))$ 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 $x, x' \in \mathcal{K}_c(\mathcal{F}(\Delta(C \times W)))$. Thus Ψ is an affine isometric embedding of $\mathcal{K}_c(\mathcal{F}(\Delta(C \times W)))$

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

in $\mathbf{C}_b(\mathfrak{U})$. Moreover, $\Psi(\{\ell^*\}) = \mathbf{0}$. In sum, $\Psi\big(\mathscr{K}_c(\mathscr{F}(\Delta(C \times W)))\big)$ is a compact and convex subset of $\mathbf{C}_b(\mathfrak{U})$ that contains the origin.

Let $\bar{V} : \Psi\big(\mathscr{K}_c(\mathscr{F}(\Delta(C \times W)))\big) \to \mathbb{R}$ 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 $\text{cone}(\Psi(\mathscr{K}_c(\mathscr{F}(\Delta(C \times W)))))$ by positive homogeneity, it follows by Proposition 7.4 that $\bar{V}$ (and hence V) has the desired representation. □

Proposition 5.12. Let $V : X$ be as in [**5.1**]. Then, the following are equivalent.

(a) V is monotone, in the sense that $x \subset x'$ implies $V(x) \leqslant V(x')$.

(b) Every charge $\mu \in \mathfrak{M}$ is *positive*, ie, $\mu(E) \geqslant 0$ for all (Borel) measurable $E \subset \mathfrak{U}$.

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 5.12) is monotone. We note that a similar statement is contained in the proof of Lemma 3.5 of Gilboa and Schmeidler (1989). □

This establishes Theorem 1. The following corollary follows immediately from Corollary 7.3 and Lemma 7.5.

Corollary 5.13. Let $V : X \to \mathbb{R}$ be as in [**5.1**]. Suppose $E \subset X$ is convex and $V|_E$ is linear. Then, there exists $\mu \in \mathfrak{M}$ such that $V(x) = \int_{\mathfrak{U}} \max_{f \in x} \sum_i p_i \mathfrak{u}_i(f(i))\, \mathrm{d}\mu(p, \mathfrak{u})$ for all $x \in E$.

6. Appendix: Proof of Theorem 2

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.

6.1. Partitional Representation

In this section, we consider the representation in [♦] of $\gtrsim$ 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 $\mathfrak{U} \times \Delta(S)$ 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 $\mathfrak{u} \in \mathbf{C}(C \times W)$.

Proposition 6.1. Consider a preference relation $\succsim$ on X, and suppose $V : X \to \mathbb{R}$ represents $\succsim$ and has the form in [$\blacklozenge$]. Then, (a) implies (b), where:

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

(b) The function V has the form

$$V(x) = \max_{(P,\mathfrak{u})\in\mathfrak{M}_p}\left[\sum_{J\in P}\left(\max_{f\in x}\sum_{s\in J}\pi_0(s \mid J)\,\mathfrak{u}_s(f(s))\right)\pi_0(J)\right] \tag{6.1}$$

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 $C \times W$ 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 $\alpha \in \Delta(C \times Z)$.

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 6.1 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 [$\blacklozenge$] 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 6.1.

6.1.1. *Nice Menus and their Density*

Recalling the notation introduced in Section 4.2, let $\tilde{\Xi}_x := \{\xi \in \Xi_x : \mathscr{I}(\xi) \sim x\}$. 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 6.2. 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. By defintion of ξ, $x \sim \mathscr{I}(\xi)$, and by repeated application of IICC(a) (Axiom 4(a)), we obtain $\{f_1, \ldots, f_m, \ell_*\} \succsim \mathscr{I}(\xi)$. It is also the case, by Monotonicity (Axiom 1(d), that $x \succsim \{f_1, \ldots, f_m\}$. Thus, to establish the claim, it suffices to show that $\{f_1, \ldots, f_m,\} \succsim \{f_1, \ldots, f_m, \ell_*\}$.

To see this, first let $x_0 := \{f_1, \ldots, f_m\}$ and $x'_0 := \{f_1, \ldots, f_m, \ell_*\}$, and suppose, by way of contradiction, that $x'_0 \succ x_0$.

Set $\varepsilon := d(x_0, x_0') > 0$ and let g^k be such that $d(x_0, x_0 \cup \{g^k\}) < \min[\varepsilon/2, 1/k]$, where k is an integer.[10] By Monotonicity (Axiom 1(d)), $x_0 \cup \{g^k\} \gtrsim x_0$.

By repeated application of IICC(a) (Axiom 4(a)), we obtain that $x_0 \cup \{g^k\} \gtrsim x_0'$. Thus, we have constructed a sequence of menus $(x_0 \cup \{g^k\})_k$ that converges to x_0 with the property that $x_0 \cup \{g^k\} \gtrsim x_0' \succ x_0$ for all k. This is impossible because $\gtrsim$ is Continuous (Axiom 1(b)) which establishes the claim. $\square$

Definition 6.3. 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 $x \succ x \backslash \{f\}$ 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 6.4. The space X_0 of nice menus 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 can be shown 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)), $y \gtrsim x$. 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 6.2) where $f_a' \in x$ for all $a \in \{1, \ldots, k\}$. Consider, now, $\mathcal{J}\,(\xi)$ and note that it can be generated inductively from $y_0 := \{f_1', \ldots, f_k'\}$ 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

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

(10) Such a g^k can always be found, for instance, by perturbing f_1 by an arbitrarily small amount.

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)), $y_i \gtrsim y_{i+1}$, 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, $x \gtrsim y_0$, and hence $x \succ y$, which contradicts the observation above that $y \gtrsim x$. Therefore, $m = j$. But then $y \gtrsim x$ 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. $\square$

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 *finite* 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 finite strategy is

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

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

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

where $\langle (p,\mathfrak{u}), f\rangle = \sum_s p_s \mathfrak{u}_s(f_i(s))$. Notice that if a finite 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 $(p,\mathfrak{u}) \in E_i$, $\langle (p,\mathfrak{u}), f_i\rangle \geqslant \langle (p,\mathfrak{u}), f\rangle$ for all $f \in x$.

In the sequel, ζ_x^μ denotes an finite optimal strategy when one exists. It is easy to see that for a finite x, an optimal strategy is always finite, 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 of $\mathfrak{U}$ for μ at x.

Definition 6.5. Let $\{f_1,\dots,f_m\}$ 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* $\{f_1,\dots,f_m\}$ if $(p,\mathfrak{u}) \in E_i$ implies $\langle (p,\mathfrak{u}), f_i\rangle \geqslant \langle (p,\mathfrak{u}), f_j\rangle$ for all $j = 1,\dots,m$.

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

For each $\mu \in \mathfrak{M}$, let $V(x,\mu) := \int_{\mathfrak{U}} \max_{f\in x} \sum_s p_s \mathfrak{u}_s(f(s))\,\mathrm{d}\mu(p,\mathfrak{u})$ 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 6.6. Let $x \in X$ be finite and suppose $\{f_1,\dots,f_m\}$ is a set of generators of x. Then, $\mu \in \Upsilon(\{f_1,\dots,f_m\})$ implies $\mu \in \Upsilon(x)$.

Proof. Consider the following string of inequalities:

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

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

Lemma 6.7. 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 of $\mathfrak{U}$ for μ at x, then it is also an optimal partition of $\mathfrak{U}$ for μ at $(1-\varepsilon)x + \varepsilon\ell$.

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

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

where the inequality uses the fact that $V(x,\mu) \geqslant V(x,\mu')$ 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 finite optimal strategy with (E_i) as the optimal partition of $\mathfrak{U}$. Then,

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

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,

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

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 finite optimal strategy at the menu x given the optimal $\mu \in \mathfrak{M}$ and completes the proof. □

6.1.2. *From Finitely Additive Measures to Partitional Systems*

A collection of probability measures $\{p_1, \dots, p_k\}$ 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. In this Section, we show that IICC also implies that the abstract measures considered above can be replaced by a partitional system.

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

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

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

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

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

Definition 6.8. 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 $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s} = \mathbf{0}$, ie, if $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}$ is trivial.
- A measure μ is *Type Ib* on E_i in state s if $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}$ is non-trivial, $\bar{p}_i(s)\bar{\mathfrak{u}}_s$ is constant on $\Delta(C \times Z)$, and ℓ_* maximizes $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}$ on $\Delta(C \times W)$.
- A measure μ is *Type IIa* on E_i in state s if $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}$ is non-trivial and not constant on $\Delta(C \times Z)$.
- A measure μ is *Type IIb* on E_i in state s if $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}$ is non-trivial, constant on $\Delta(C \times Z)$, and there exists $\alpha \in \Delta(C \times W)$ such that $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(\alpha) > \bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(\beta)$ for some (and hence all) $\beta \in \Delta(C \times Z)$.

It is easy to see that the above taxonomy of measures is both mutually exclusive and exhaustive. Analogous to the definition in Section 3.1 above (and abusing notation), for any $\alpha \in \Delta(C \times W)$ 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 6.9. Let x be a finite menu, $\mu \in \Upsilon(x)$, and suppose there is a finite optimal strategy ζ_x^μ with (E_i) as the optimal partition of $\mathfrak{U}$, 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 $\alpha \in \Delta(C \times W)$ such that

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

Then, the menu $z := x \backslash \{f_i\} \cup \{f_i \oplus_{\varepsilon,s} \alpha\}$ 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

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

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

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

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

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

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

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

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

Let $\mathfrak{M}_0 := \{\Upsilon(\{f_1,\dots,f_m\}) : \{f_1,\dots,f_m\} \text{ generates } x \text{ for some } x \in X\}$. It follows from Lemma 6.6 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 6.10. 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 a finite optimal 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 $\xi \in \Xi_x$ 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 $\Delta(C \times W)$. 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 $\Delta(C \times W)$, 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 $x_{i,\varepsilon} := \{f_1, \ldots, (1-\varepsilon)f_i + \varepsilon f_i^*, \ldots, f_m\}$. By Lemma 6.9, $V(x_{i,\varepsilon}) > V(x)$ for all $\varepsilon > 0$. Notice also that $x_{i,\varepsilon} \to x$ as $\varepsilon \to 0$.

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

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

Notice that $W(0) = V(x)$ and that because $\Xi_{i,\varepsilon}$ 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 6.2, it follows that $V(x_{i,\varepsilon}) \geqslant V(\mathcal{J}(\xi))$ 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, $W(\varepsilon) \geqslant V(x_{i,\varepsilon})$. 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 $s \in J_i^c$, 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 6.9) 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 $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(\cdot)$ on $\Delta(C \times W)$. If μ is of Type IIa on E_i in state $s \in J_i^c$, then $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(\cdot)$ is not constant on $\Delta(C \times Z)$. If μ is of Type IIb on E_i in state $s \in J_i^c$, then $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(\cdot)$ is constant on $\Delta(C \times Z)$. However, in either case, there exists $\ell \in L$ such that $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(f_i(s)) > \bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(\ell(s))$. (Such an ℓ exists because $f_i(s)$ is a maximizer of $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(\cdot)$ and by hypothesis that μ is of Type II, there exists some $\beta \in \Delta(C \times Z)$ that is *not* a maximizer.)

Consider the menu $\frac{1}{2}x + \frac{1}{2}\ell$. By Lemma 6.7, 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 6.7 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 $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}$ on $\Delta(C \times W)$ (although $f_i(s)$ is), which means that we now satisfy the hypotheses of Case 1. Lemma 6.7 ensures that $\Upsilon(\frac{1}{2}x + \frac{1}{2}\ell) \cap \Upsilon(x) \neq \varnothing$ and that a finit 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. $\square$

Let x be nice and let $\mu \in \Upsilon_0(x)$. Let $(E_i^{\mu,x})$ 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

$$\gamma^i_{\mu,x} := \sum_{s \in J_i} \bar{p}_i(s)$$

[♣]
$$p_i(s) := \begin{cases} \bar{p}_i(s)/\gamma^i_{\mu,x} & \text{if } s \in J_i \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{\mathfrak{u}}_s := \gamma^i_{\mu,x}\bar{\mathfrak{u}}_{i,s} \qquad \text{where } i \text{ is such that } s \in J_i$$

and let

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

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 $\{f_1, \ldots, f_k\} \sim \{f_1, \ldots, f_k\}\backslash\{f_i\}$. 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

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

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

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

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 $\mathfrak{U}_{s,\ell^\dagger(s)}$ 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 $\{p_1, \ldots, p_k\}$, 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 measures in $\hat{\mathfrak{M}}_p$. Towards this end, let us define $\hat{V} : X_0 \to \mathbb{R}$ as

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

Proposition 6.11. 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 $\{f_1, \ldots, f_k\}$ the unique generator set of x. Let us first prove that $V(x) \leqslant \hat{V}(x)$. Let $(E_i^{\mu,x})$ be an optimal partition of $\mathfrak{U}$ for μ at x, and let $\hat{\mu} = \mathfrak{D}(\mu, x, (E_i^{\mu,x}))$. Then,

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

Lemma 6.10 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, $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(f(s)) \leqslant 0 = \bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(\ell^\dagger(s))$ for all $s \in J_i^c$. Therefore, it must be that

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

We now prove that $\hat{V}(x) \leqslant V(x)$ 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 $(p_i, \mathfrak{u})$.

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 $(p_i, \mathfrak{u})$). The value of this strategy, $V(x,\mu,\zeta^\mu)$, is given by

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

It follows from Lemma 6.10 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 $\bar{p}_i(s)\bar{\mathfrak{u}}_{i,s}(f_i(s)) \leqslant 0$. 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 $V(x') \geqslant \hat{V}(x,\hat{\mu}) > V(x)$, 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 $x \gtrsim x'$. 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

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

where $\hat{\mu} = \mathfrak{D}(\mu, x, (E_i^{\mu,x}))$, $\mu \in \Upsilon_0(x)$, and $(E_i^{\mu,x})$ is a finite optimal 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 6.4, 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 $\gtrsim$ is represented by a function $V : X \to \mathbb{R}$ that has the form

$$V(x) = \max_{\mu \in \mathfrak{M}} V(x,\mu) \tag{6.2}$$

where

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

Lemma 6.12. 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 $(p_1,\mathfrak{u}),\ldots,(p_k,\mathfrak{u})$, and let $\|\mu\|_1$ be the total weight of μ. (That is, $\|\mu\|_1 := \sum_i \mu((p_i,\mathfrak{u}))$.) For any $s \in S$, define $\hat{\mathfrak{u}}_s := \|\mu\|_1 \mathfrak{u}_s$, and for any $p \in \Delta(S)$, let $\hat{\mu}(p,\hat{\mathfrak{u}}) := \mu(p,\mathfrak{u})/\|\mu\|_1$ where $\hat{\mathfrak{u}} = (\hat{\mathfrak{u}}_s)_{s\in S}$. It is easy to see that $\hat{\mu}$ so defined is elementary and is also a probability measure.

Moreover, we have

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

which establishes the claim. $\square$

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

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

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

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

which completes the proof. $\square$

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 $J_i := \text{supp}(p_i)$.

Lemma 6.14. 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 μ^* 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 6.13, 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

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

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

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

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

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

which completes the proof. $\square$

We are now in a position prove Proposition 6.1.

6.1.3. *Proof of Proposition 6.1*

Proof of Proposition 6.1. We shall prove (a) implies (b). We have shown that given the representation [♦] in Theorem 1 and IICC (Axiom 4), V has the form in [**6.2**], where every $\mu \in \mathfrak{M}$ is an elementary (positive, but finite) measure. Lemma 6.12 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

$J_i = \text{supp}(p_i)$, and notice that (J_i) is a partition of S. Lemma 6.14 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 [6.2] — 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 $\alpha \in \Delta(C\times Z)$, 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 $\alpha \in \Delta(C \times Z)$, it follows that $\mathfrak{u}_s$ and $\mathfrak{u}'_s$ are identical on $C \times Z$ for all $(P, \mathfrak{u}), (P', \mathfrak{u}') \in \mathfrak{M}'$. This proves that (a) implies (b). □

6.2. Separable Representation

We now investigate the implication of imposing State-Contingent Indifference to Correlation (henceforth SCIC, Axiom 3).

Proposition 6.15. Let V be as in [6.1] 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 $v_s \in \mathbf{C}(W)$ for each s such that $(P, (v_s)), (P', (v'_s)) \in \mathfrak{M}''_p$ implies $v_s|_Z = v'_s|_Z$ for all $s \in S$, and V can be written as

[6.3]
$$V(x) = \max_{(P,(v_s)) \in \mathfrak{M}''_p} \sum_{J \in P} \pi_0(J) \max_{f \in x} \sum_{s} \pi_0(s \mid J) \left[u_s(f_1(s)) + v_s(f_2(s)) \right]$$

Before we prove the proposition, we introduce some notation and prove a lemma. Suppose $V : X \to \mathbb{R}$ represents $\succsim$ and takes the form [6.1]. For each $(P, \mathfrak{u})$, define

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

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

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

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

and observe that $\mathfrak{u}'_s : \Delta(C \times W) \to \mathbb{R}$ is continuous and linear[11] so that $\mathfrak{u}'_s \in \mathbf{C}(C \times W)$. Moreover, for all $(P, \mathfrak{u}'), (\hat{P}, \hat{\mathfrak{u}}') \in \mathfrak{M}_p$, $\mathfrak{u}'_s|_{C \times Z} = \hat{\mathfrak{u}}'_s|_{C \times Z}$. 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

[6.4]
$$V'(x) := \max_{(P, \mathfrak{u}') \in \mathfrak{M}'_p} V(x, (P, \mathfrak{u}'))$$

(11) 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.

Observe that V' is monotone, ie, $x \subset x'$ implies $V'(x) \leqslant V'(x')$. This follows immediately from the form of V' in [6.4]. We claim that V' also represents $\succsim$.

Lemma 6.16. Let V and V' be defined as in [6.1] and [6.4] 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 $x \in X$, $V'(x) \leqslant V(x)$. 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 $V(x, (P, \mathfrak{u}')) \leqslant V(x, (P, \mathfrak{u})) \leqslant V(x)$, where the first inequality follows from the definition of $\mathfrak{u}'_s$, which entails that for each $\alpha \in \Delta(C \times W)$, $\mathfrak{u}'_s(\alpha) \leqslant \mathfrak{u}_s(\alpha)$.

We shall now show that for all finite $x \in X$, $V(x) \leqslant V'(x)$. 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.

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.

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

which completes the proof. □

Proof of Proposition 6.15. It is easy to see that (b) implies (a). We now show that (a) implies (b).

Lemma 6.16 implies we can replace V in [**6.1**] by V' in [**6.4**]. Moreover, from the definition of V in [**6.1**], $\mathfrak{u}_s(\alpha) = \mathfrak{u}'_s(\alpha)$ for all $(P,\mathfrak{u}),(P',\mathfrak{u}') \in \mathfrak{M}'_p$ and for all $\alpha \in \Delta(C \times Z)$.

For any $\alpha \in \Delta(C \times W)$ with marginals α_1 and α_2, let $\alpha_1 \otimes \alpha_2 \in \Delta(C) \times \Delta(W)$ denote the product lottery with the same marginals. Recall that the lottery $\ell^\dagger \in L$ 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 $\alpha \in \Delta(C \times Z)$ because $\alpha_1 \otimes \ell_2^\dagger(s) \in \Delta(C \times Z)$); 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 $v_s(\cdot) \in \mathbf{C}(W)$. 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 [**6.4**] and invoking Lemma 6.16 gives us [**6.3**], as desired. □

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

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

6.3. Representation with Deterministic Continuation Utilities

Thus far, we have seen that $\succsim$ has a representation as in [**6.3**]. We now impose Concordant Independence (Axiom 5).

Proposition 6.17. Let V be as in [**6.3**] and suppose V represents $\succsim$. Then, the following are equivalent.

(a) $\succsim$ satisfies Concordant Independence (Axiom 5).

(b) V can be written as

$$[\mathbf{6.5}] \qquad V(x) = \max_{P \in \mathfrak{M}^\sharp_p} \sum_{J \in P} \left[\max_{f \in x} \sum_s \pi_0(s \mid J) \big[u_s(f_1(s)) + v_s(f_2(s), P)\big] \pi_0(J) \right]$$

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

For a fixed P in the representation in [**6.3**], let X'_P and $\hat{X}_P$ be defined as follows:

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

Recall the choice problem $x_1(P)$ defined above. For each $J \in P$, let

$$f_J(s) = \begin{cases} (c^+_s, z^+_s) & s \in J \\ (c^-_s, z^-_s) & \text{otherwise} \end{cases}$$

and define the menu $x_1(P) := \{f_J : J \in P\}$. That is, for any partition P, $x_1(P)$ is a problem where the choice of P is optimal.

Lemma 6.18. 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, and only if, P is not finer than Q.

Proof. We begin by establishing three claims.

(i) In the representation [6.3], $v_s(z_s^+) > v_s(z_s^-)$ for all $s \in S$.

(ii) $V(x_1(Q)) \leqslant V(\ell^*)$ 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 a repeated application of IICC (Axiom 4), $\big[\ell_* \oplus_{(1,s)} (c_s^+, w_s^+)\big] \succ \big[\ell_* \oplus_{(1,s)} (c_s^-, w_s^-)\big] = \ell_*$ for all $s \in S$. Because we have $v_s(z) = v'_s(z)$ for all $z \in Z$ and $(P, (v_s)), (P', (v'_s)) \in \mathfrak{M}''_p$ in [6.3], $v_s(z_s^+) > v_s(z_s^-)$ follows for all $s \in S$.

Given claim (i), and because $u_s(c_s^+) \geqslant u_s(c_s^-)$ for all s, claim (ii) follows by evaluating V in [6.3] at $x_1(Q)$.

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

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

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

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

Then, by claim (i) and because $u_s(c_s^+) \geqslant u_s(c_s^-)$ 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

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

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

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

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 $V(x_1(P),(Q,(v'_s))) \leqslant V(\ell^*)$ (claim (ii) above). This implies $(1-\lambda)x + \lambda x_1(P) \in X'_P$. Moreover, it now follows immediately that $V\left((1-\lambda)\,x + \lambda x_1(P)\right) = V\left((1-\lambda)\,x + \lambda \ell^*\right)$.

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,

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

It is immediate that $V((1-\lambda)x + \lambda x_1(Q)) = V((1-\lambda)x + \lambda\ell^*)$ if Q coarser than P, which completes the proof. $\square$

Lemma 6.19. X'_P is convex and consists of concordant choice problems.

Proof. Consider $x, y \in X'_P$. By Lemma 6.18, $\frac{1}{2}x + \frac{1}{2}x_1(P) \in X'_P$ and $V(\frac{1}{2}x + \frac{1}{2}x_1(Q)) = V(\frac{1}{2}x + \frac{1}{2}\ell^*)$ if, and only if, Q is coarser than P, and the same is true for y. Hence, $\frac{1}{2}x + \frac{1}{2}x_1(Q) \sim \frac{1}{2}x + \frac{1}{2}\ell^*$ if, and only if, $\frac{1}{2}y + \frac{1}{2}x_1(Q) \sim \frac{1}{2}y + \frac{1}{2}\ell^*$ for all $Q \in \mathscr{P}$, which establishes that x and y are concordant (as in Definition 4.1). By Concordant Independence (Axiom 5), x, y, and $\frac{1}{2}x + \frac{1}{2}y$ are concordant. Concordant Independence (Axiom 5) also implies that V is linear on a concordant set of menus, so that $V(\frac{1}{2}x + \frac{1}{2}y) = \frac{1}{2}V(x) + \frac{1}{2}V(y)$.

Now suppose $(Q,(v'_s)) \in \mathfrak{M}''_p$ is optimal for $\frac{1}{2}x + \frac{1}{2}y$. By the representation in [**6.4**], $V(\frac{1}{2}x + \frac{1}{2}y) = V(\frac{1}{2}x + \frac{1}{2}y;(Q,(v'_s))) = \frac{1}{2}V(x;(Q,(v'_s))) + \frac{1}{2}V(y;(Q,(v'_s)))$. Since $V(x) \geqslant V(x;(Q,(v'_s)))$ and $V(y) \geqslant V(y;(Q,(v'_s)))$ for all $Q \in \mathscr{P}$, it must be that $V(x;(Q,(v'_s))) = V(x)$. Because $x \in X'_P$, $Q = P$, ie, $\frac{1}{2}x + \frac{1}{2}y \in X'_P$.

Standard continuity 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. $\square$

Lemma 6.20. 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 $n \geqslant 2$ (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 6.18, we find that $V((1-\lambda)x+\lambda x_1(P_1), (P_1, (v_s^1))) = V((1-\lambda)x+\lambda\ell^*) > V((1-\lambda)x+\lambda x_1(P_1), (P_i, (v_s^i)))$ for some v_s^1 with $(P, (v_s^1)) \in \mathfrak{M}''_p$ and all $(P_i, (v_s^i)) \in \mathfrak{M}''_p$ 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 6.21. 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 6.19.) If P is not finer than Q, by Lemma 6.18, $(1-\lambda)y + \lambda x_1(Q) \succ (1-\lambda)x + \lambda x_1(Q)$, 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 6.19. ☐

Notice that replacing $\mathfrak{M}''_p$ with its weak* closure (in the event that it is not weak* compact) in [**6.3**] does not affect the representation. Therefore, we shall now assume that $\mathfrak{M}''_p$ is weak*-compact.

Lemma 6.22. Let $x \in \mathrm{cl}(X'_P)$. Then, there exists (v_s) such that $(P, (v_s)) \in \mathfrak{M}''_p$ is optimal for all $y \in \mathrm{cl}(X'_P)$.

Proof. By Lemma 6.21, $Y_x \subset X'_P$ which, by Lemma 6.19, is convex. By Concordant Independence, $\succsim|_{X'_P}$ satisfies Independence. That is, $V|_{X'_P}$ is linear. It follows from Corollary 7.3 and Lemma 7.5 below 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 $x \in \mathrm{cl}(X'_P)$. ☐

It follows that we can replace the set $\mathfrak{M}''_p$ by a finite collection $\{(P_1, (v_s^1)), \ldots, (P_n, (v_s^n))\} = \mathfrak{M}^\sharp_p$ as in [**6.5**]. Thus, we have shown that (a) implies (b) in Proposition 6.17. That (b) implies (a) is clear.

With the representation in [**6.5**] in hand, for any partition $P \in \mathfrak{M}^\sharp_p$, we define

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

which is the utility of a menu, conditional on an initial information choice.

6.4. Proof of Theorem 2

Given the representation in [6.5], we complete the proof that the Axioms 1–5 imply the representation in Theorem 2 by showing that $v_s(w,\cdot) \geqslant v_s(z_s^-,\cdot)$ for all $w \in W$.

Notice that $\mathfrak{M}_p^\sharp$ in [6.5] is finite and can be taken to be minimal (in the sense that if $\mathfrak{N}_p^\sharp$ is another set that represents V as in [6.5], then $\mathfrak{M}_p^\sharp \subset \mathfrak{N}_p^\sharp$) without affecting the representation. Recall that $X^* := \{(1-t)x + t\ell^* : x \in X \text{ is finite}, t \in (0,1)\}$.

Lemma 6.23. Let $\succsim$ have a representation as in [6.5]. For all $P \in \mathfrak{M}_p^\sharp$, there exists a finite $x \in X'_P \cap X^*$ that can be written as $x = \frac{1}{2}x' + \frac{1}{2}x_1(P)$ for some $x' \in X$.

Proof. The finiteness and minimality of $\mathfrak{M}_p^\sharp$ in [6.5] implies that for any $P \in \mathfrak{M}_p^\sharp$, there exists an open set $O \subset X'_P$. Because the space X^* is dense in X, there exists $x' \in O \cap X^*$. It follows immediately from the representation in [6.5] that $x := \frac{1}{2}x' + \frac{1}{2}x_1(P) \in X'_P \cap X^*$, as claimed. $\square$

Lemma 6.24. Let $\succsim$ have a representation as in [6.5]. For all $P \in \mathfrak{M}_p^\sharp$, $v_s(w,P) \geqslant v_s(z_s^-,P)$.

Proof. Suppose instead that $v_s(w,P) < v_s(z_s^-,P)$. Consider $x \in X'_P \cap X^*$ which exists by Lemma 6.23. Then, for $\varepsilon > 0$ small enough such that P remains optimal, $[x \oplus_{\varepsilon,s} z_s^-] \succ [x \oplus_{\varepsilon,s} w]$. To see this, suppose f is such that $f \oplus_{\varepsilon,s} (c_s^-, w)$ is chosen optimally from the menu $x \oplus_{\varepsilon,s} w$. Then, $v_s(w,P) < v_s(z_s^-,P)$ implies

$$
\begin{aligned}
&(1-\varepsilon)[u_s(f_1(s)) + v_s(f_2(s),P)] + \varepsilon\left[u_s(c_s^-) + v_s(w,P)\right] \\
&\qquad < (1-\varepsilon)\left[u_s(f_1(s)) + v_s(f_2(s),P)\right] + \varepsilon\left[u_s(c_s^-) + v_s(z_s^-,P)\right]
\end{aligned}
$$

This implies $V(x \oplus_{\varepsilon,s} z_s^-) > V(x \oplus_{\varepsilon,s} w)$. But this contradicts part (a) of IICC (Axiom 4), which requires that $[x \oplus_{\varepsilon,s} w] \succsim [x \oplus_{\varepsilon,s} z_s^-]$ for all $w \in W$. $\square$

That the representation in Theorem 2 implies the Axioms is straightforward.

7. 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 $\partial f(x) := \{x^* \in X^* : \langle y - x, x^* \rangle \leqslant f(y) - f(x) \text{ for all } y \in C\}$. A necessary and sufficient condition for the existence of a subdifferential at $x \in C$ is that there exists $K \geqslant 0$ such that for all $y \in X$, $f(x) - f(y) \leqslant K \|y - x\|$. To see this, recall that the set $\text{epi}(f) := \{(x, t) \in X \times \mathbb{R} : t \geqslant f(x)\}$, 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 $\|x^*\| \leqslant K$. 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 7.1 (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 $\|x^*\| \leqslant K$.

In what follows, we will denote by $\partial_K f(x) := \{x^* \in \partial f(x) : \|x\| \leqslant K\}$. 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 $\varphi \leqslant f$ for all $y \in C$ if, and only if, $\langle y, x^* \rangle - a \leqslant f(y)$, 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 $\varphi(y; x^*) = f(x) + \langle y - x, x^* \rangle \leqslant f(y)$ 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 $a \geqslant a(x^*)$, $\langle x, x^* \rangle - a < \langle x, x^* \rangle - a(x^*)$, so $a(x^*) = \inf[a \in \mathbb{R} : f(x) \geqslant \langle x, x^* \rangle - a] = \sup[x \in C : \langle x, x^* \rangle - f(x)]$. 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

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

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

Proposition 7.2. 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 $y \in C$, $\langle y - x, x^* \rangle \leqslant f(y) - f(x^*)$. 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 $\langle y - \lambda x \rangle = \lambda \langle z - x^* \rangle \leqslant \lambda \big[f(z) - f(x) \big] = f(y) - f(\lambda x)$, 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 $f(0) = f(2 \times 0) = 2f(0)$ which implies $f(0) = 0$.) Therefore, $f^\star(x^*) \geqslant \langle 0, x^* \rangle - f(0) = 0$. 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 7.3. Let $C \subset X$ be a convex cone, and $f \in \mathbb{R}^C$ be convex, Lipschitz, and positively homogeneous. 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 $\|x^*\| \leqslant K$, where K is the Lipschitz constant of f. We have also established that $x^* \in \partial f(\lambda x)$ for all $\lambda \geqslant 0$. Therefore, $f(y) \geqslant \langle y, x^* \rangle$ for all $y \in C$. Letting $\mathfrak{M} = \mathrm{cl}(\{x^* \in \partial f(x) : x \in C, \|x^*\| \leqslant K\})$ (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 $\mathrm{ch}(\{x\} \cup A)$. 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 7.4. 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 7.1, there exists $y^* \in \partial_K f(y)$. Recall the affine function $\varphi(\cdot, y^*) X :\rightarrow \mathbb{R}$ given by

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

The affine function φ satisfies the following two properties:

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

The first requirement implies that $f(y_i) \geqslant \varphi(y_i; y^*)$ for all $i = 1, \ldots, n$. Summing up and dividing by n, we see that $\frac{1}{n} \sum_i f(y_i) \geqslant \frac{1}{n} \sum_i \varphi(y_i; y^*)$. 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 $\{x^* \in X^* : \|x^*\| \leqslant K\}$.[12] Thus, $\left(\partial_K f(y)\right)_{y \in E_x}$ is a collection of closed subsets of the compact set $\{x^* \in X^* : \|x^*\| \leqslant K\}$. 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 $\left(\partial_K f(y)\right)_{y \in E_x}$ has the finite intersection property. The compactness of $\{x^* \in X^* :$

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

$\|x^*\| \leqslant K\}$ 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 7.5. 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 $x_1, \dots, x_n \in C_0$ 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

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

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. □

References

Aliprantis, Charalambos, and Kim C. Border. 1999. *Infinite Dimensional Analysis: A Hitchhiker's Guide, 2/e.* Springer, New York. (Cit. on pp. 14, 15, 42).

de Oliveira, Henrique, Tommaso Denti, Maximilian Mihm, and M. Kemal Ozbek. 2017. "Rationally Inattentive Preferences and Hidden Information Costs". *Theoretical Economics* 12:621–654. (Cit. on pp. 2, 4).

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 pp. 4, 14).

Dillenberger, David, R. Vijay Krishna, and Philipp Sadowski. 2020. *Subjective Information Choice Processes*. Tech. rep. Duke University. (Cit. on p. 2).

Ekeland, Ivar, and Thomas Turnbull. 1983. *Infinite-Dimensional Optimization and Convexity*. Chicago: University of Chicago Press. (Cit. on pp. 39, 40).

Ergin, Haluk, and Todd Sarver. 2010a. "A Unique Costly Contemplation Representation". *Econometrica* 78 (4): 1285–1339. (Cit. on pp. 2, 4, 9).

– . 2010b. "Supplement to 'A Unique Costly Contemplation Representation'". *Econometrica Supplemental Material* 78. (Cit. on p. 16).

Gale, David. 1967. "A Geometric Duality Theorem with Economic Applications". *Review of Economic Studies* 34 (1): 19–24. (Cit. on p. 40).

Gilboa, Itzhak, and David Schmeidler. 1989. "Maxmin Expected Utility with a Non-Unique Prior". *Journal of Mathematical Economics* 18:141–153. (Cit. on pp. 4, 16).

Kreps, David M. 1979. "A Representation Theorem for 'Preference for Flexibility'". *Econometrica* 47 (3): 565–577. (Cit. on p. 4).

Krishna, R. Vijay, and Philipp Sadowski. 2014. "Dynamic Preference for Flexibility". *Econometrica* 82 (2): 655–703. (Cit. on p. 6).

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