Supplementary Appendix to "Constrained Optimal Discounting" *

Jawwad Noor   Norio Takeoka

May 8, 2020

Abstract

This supplementary appendix to Noor and Takeoka [3] provides an axiomatization for the General Constrained Costly Empathy (General CCE) representation with negative payoffs.

1 Primitives

There are $T+1<\infty$ periods, starting with period 0. The space C of outcomes is assumed to be $C=\mathbb{R}$.[1] Let Δ denote the set of simple lotteries over C, with generic elements $p, q, ...$ We will refer to p as consumption. Consider the space of consumption streams $X=\Delta^{T+1}$, endowed with the product topology. A typical element in X is denoted by $x=(x_0, x_1, \cdots, x_T)$. The primitive of our model is a preference $\succsim$ over X.

Let $\Delta_0 \subset X$ denote the set of streams $x=(p, 0, \cdots, 0)$ that offer consumption p immediately and 0 in every subsequent period. Abusing notation, we often use p to denote both a lottery $p \in \Delta$ and a stream $(p, 0, \cdots, 0) \in \Delta_0$. Thus, 0 also denotes the stream $(0, \cdots, 0)$. An element of Δ that is a mixture between two consumption alternatives $p, q \in \Delta$ is denoted $\alpha \circ p+(1-\alpha) \circ q$ for any $\alpha \in [0,1]$. The same mixture is also regarded as $\alpha \circ p+(1-\alpha) \circ q \in \Delta_0$.

Denote by p^t the stream that pays $p \in \Delta$ at time t and 0 in all other periods. Such a stream is called a *dated reward*.

For notational convenience, for all streams x, $y \in X$ and $S \subset \{0,1, \cdots, T\}$, let xSy denote the stream that pays according to x at $t \in S$ and according to y otherwise. In particular, if $S=\{t\}$, the stream is denoted by $x\{t\}y$.

*Noor (the corresponding author) is a the Dept of Economics, Boston University, 270 Bay State Road, Boston MA 02215. Email: jnoor@bu.edu. Takeoka is at the Dept of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan. Email: norio.takeoka@r.hit-u.ac.jp.

[1] This can be replaced with an abstract metric space for the main results of the paper with a suitable element that can be interpreted as "zero", but we adopt real consumption, which is applicable to the MRS approach considered by Noor and Takeoka[2].

Say that a stream x is *positive* if $x_t \succsim 0$ for all t, and it is *negative* if $x_t \precsim 0$ for all t. Let X_+ denote the set of positive streams, that is,

$$X_+ := \{x \in X \mid x_t \succsim 0, \ \forall t\}.$$

Note that via the identification between p and $(p, 0, .., 0)$, it is meaningful to say that $p \in \Delta$ is positive or negative.

2 The General CCE Representation

2.1 Functional Form

Say that a tuple $(u, \{\varphi_t\}, K)$ is *basic* if

(a) $u : \Delta \to \mathbb{R}$ is continuous and mixture linear with (a) $u(0) = 0$ and (b) $u(\Delta) = \mathbb{R}$,

(b) $\varphi_t : [0, 1] \to \mathbb{R}_+ \cup \{\infty\}$ is an increasing convex function that is

   (i) strictly increasing, strictly convex and differentiable on $\{d : 0 < \varphi_t(d) < \infty\}$, and

   (ii) satisfies $\varphi_t(0) = 0$, $\varphi_t'(0) = 0$ and $\varphi_t \leq \varphi_{t+1}$ for all $t < T$,

(c) $K : X \setminus \Delta_0 \to \mathbb{R}_{++} \cup \{\infty\}$ is either $K_x = \infty$ for all x or a continuous function with the following properties:

   (i) $K_{\lambda x} = K_x$ for all x and $\lambda > 0$,

   (ii) $K_x = K_y$ if $|u(x_t)| = |u(y_t)|$ for all $t \geq 1$,

   (iii) $K_{p^t} = \varphi_t(\overline{d}_t)$ for all $p \in \Delta$ and $t \geq 1$, where $\overline{d}_t$ is a supremum of the effective domain $\mathrm{eff}(\varphi_t) := \{d_t \in [0, 1] : \varphi_t(d_t) < \infty\}$,

   (iv) if $K_x < \infty$, $K_x \leq K_{xS0}$ for all $S \subset \{1, \cdots, T\}$ with $x_t \succ 0$ for some $t \in S$.

Compared to the regular tuple that defines a CCE representation of Noor and Takeoka [3], the cognitive cost function φ_t here is a more general convex function, and moreover, it is possible that $\varphi_t(d) = 0$ for all d in some interval $[0, \underline{d}_t]$. Intuitively, there can be a base-line degree of selflessness (corresponding to a discount function $\underline{d}_t$) that the agent can access costlessly, that is, $\varphi_t(\underline{d}_t) = 0$ for each t.

In Noor and Takeoka [3], the capacity constraint is some constant number $K > 0$. The capacity constraint K as given above is more general in that it can now change with the stream. Property (c)(i) states that it is homogeneous of degree 0. This can be viewed as saying that it depends on the normalized distribution consumption across time. Property (c)(ii) states that K depends only on the absolute value of utility streams. Property (c)(iii) states that the empathy constraints for dated rewards attain the cost for the maximum level of discount factor at each period. Property (c)(iv) requires that a stream x is associated with weakly less capacity than any of its component rewards.

Property (c) implies that for all x with $x_t \not\sim 0$ for some t,

$$K_x \leq K_{x\{t\}0} = K_{p^t} = \varphi_t(\overline{d}_t), \tag{1}$$

which means that the empathy constraint K_x is bounded by the empathy constraints for dated rewards. Another implication of (c) is that if $K_x = \infty$ for some x, $K_{p^t} = \infty$ for all t.

For each x, the optimal discount function D_x is chosen subject to two constraints. The first is the *capacity (or empathy) constraint*:

$$\varphi(D) \leq K_x,$$

where $\varphi(D) = \sum_{t\geq 1} \varphi_t(D(t))$. The second constraint,

$$D(t) \leq \overline{d}_t, \text{ for all } t \geq 1,$$

is called the *boundary constraint*.

Define the representation as follows:

Definition 1 (General CCE Representation) *A General Constrained Costly Empathy (CCE) representation is a basic tuple* $(u, \{\varphi_t\}, K)$ *such that* $\succsim$ *is represented by the function* $U : X \to \mathbb{R}$ *defined by*

$$U(x) = u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t), \qquad x \in X,$$

$$\textit{s.t. } D_x = \arg\max_{D\in[0,\overline{d}_t]^T} \{\sum_{t\geq 1} D(t)\,|u(x_t)| - \varphi_t(D(t))\} \textit{ subject to } \varphi(D) \leq K_x.$$

For each stream x, an optimal discount function D_x is determined by maximizing the discounted utilities minus aggregated costs for the discount function subject to the empathy and boundary constraints. By condition (1), for all t,

$$\varphi(D) \leq K_x \leq K_{p^t} = \varphi_t(\overline{d}_t).$$

Therefore, if D satisfies the empathy constraint, it also satisfies the boundary constraint, that is, the boundary constraint is redundant. Consequently, an optimal discount function for the General CCE representation is determined as

$$D_x = \arg\max_{D\in\mathbb{R}^T_+} \{\sum_{t\geq 1} D(t)\,|u(x_t)| - \varphi_t(D(t))\} \text{ subject to } \varphi(D) \leq K_x.$$

2.2 Properties of Optimal Discount Functions

For any stream $\overline{x} \in X \setminus \Delta_0$, consider the ray passing through $\overline{x}$:

$$X_{\overline{x}} = \{x \in X \,|\, x = \alpha\overline{x},\ \exists\alpha > 0\}.$$

By property (c) of the General CCE representation, K_x is constant on $X_{\overline{x}}$. Denote $K = K_x$ for some (any) $x \in X_{\overline{x}}$.

Proposition 1 *(1) For any $x \in X_{\overline{x}}$, if $\varphi(D_x) < K$, D_x is strictly increasing, and is obtained explicitly as*

$$D_x(t) = (\varphi_t')^{-1}(|u(x_t)|).$$

(2) For any $x, y \in X_{\overline{x}}$, if $\varphi(D_x) = \varphi(D_y) = K$,

$$D_x = D_y.$$

Moreover, $D_x(t)$ depends on the capacity cap K and the whole stream x, not only on the payoff at t.
(3) D_x is weakly increasing on $X_{\overline{x}}$: D_x is strictly increasing if $\varphi(D_x) < K$, and is constant if $\varphi(D_x) = K$.

From part (1), an optimal discount function $D_x(t)$ depends only on the payoff in period t if the capacity constraint is not binding. By substituting it into the General CCE representation, $U(x)$ is written as

$$U(x) = u(x_0) + \sum_{t\geq 1}(\varphi_t')^{-1}(|u(x_t)|)u(x_t).$$

Thus, $U(x)$ is additively separable if the capacity constraint is not binding.. According to increasing in payoffs, $D_x(t)$ grows unless the capacity constraint is binding. Once the constraint hits, $D_x(t)$ stops growing. Afterwards, on the same ray, $D_x(t)$ is constant but depends on the whole stream x. The representation $U(x) = u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t)$ is not additively separable.

3 Behavioral Foundation

Consider a binary relation $\succsim$ over the space of consumption streams $X = \Delta^{T+1}$ as defined in Section 1.

3.1 Basic Axioms

Axiom 1 (Weak Regularity) *(a) (Order). $\succsim$ is complete and transitive.*
(b) (Continuity). For all $x \in X$, $\{y \in X : y \succsim x\}$ and $\{y \in X : x \succsim y\}$ are closed.

(c) (Impatience). For any $p \in \Delta$ and $t < t'$,

$$(p)^t \succsim (p)^{t'}.$$

(d) (C-Monotonicity): for all $c, c' \in C$,

$$c \geq c' \Longleftrightarrow c \succsim c'.$$

(e) (Risk Preference). For any $p, p', p'' \in \Delta$ and $\alpha \in (0, 1]$,

$$p \succ p' \Longrightarrow \alpha \circ p + (1 - \alpha) \circ p'' \succ \alpha \circ p' + (1 - \alpha) \circ p''.$$

(f) (Present Equivalents). For any stream x there exist $c, c' \in C$ s.t.

$$c \succsim x \succsim c'.$$

Order and Continuity are standard. Impatience requires that positive outcomes are weakly preferred sooner rather than later. C-Monotonicity states that more consumption is better than less. Present Equivalents states that for any stream, there are immediate consumption levels that are better and worse than x. Given Order and Continuity, this ensures that each stream x has a present equivalent $c_x \in C$. Notably, each x has a unique present equivalent c_x (by C-Monotonicity, $x \sim c_x > c_y \sim y$ implies $c_x \succ c_y$ and therefore $x \succ y$). Risk Preference imposes vNM Independence only on immediate consumption.

3.2 Axioms for Positive Streams

As argued by Noor and Takeoka [2, 3], the magnitude effect can be identified via violation of Homotheticity: For any positive streams $x \in X_+$ and any $\alpha \in (0, 1)$,

$$c_x \sim x \Longrightarrow \alpha \circ c_x \sim \alpha x.$$

Axiom 2 (Weak Homotheticity) *For any $x \in X_+$ and any $\alpha \in (0, 1)$,*

$$c_x \sim x \Longrightarrow \alpha \circ c_x \succsim \alpha x.$$

Say that a stream $x \in X_+$ is *ℓ-Magnitude Sensitive* if the agent's impatience strictly reduces whenever the stream is made less desirable.

Definition 2 (ℓ-Magnitude-Sensitivity) *A stream $x \in X_+$ is ℓ-Magnitude Sensitive if*

$$c_x \sim x \Longrightarrow \alpha \circ c_x \succ \alpha x \text{ for all } \alpha \in (0, 1).$$

The set of all ℓ-Magnitude Sensitive streams is denoted by $X_\ell \subset X_+$.

By Risk Preference, it is clear that immediate rewards are not ℓ-Magnitude Sensitive. That is, $\Delta_0 \cap X_\ell = \emptyset$.

3.2.1 Structures on X_ℓ

One can imagine that if an agent establishes empathy for self t then she may costlessly empathize with adjacent selves $t-1$ and $t+1$. However, in order to retain some additive separability in our model, we rule this out, or alternatively, we require that the duration of a period is sufficiently "long" that such intertemporal complementarities disappear.

Axiom 3 (X_ℓ-Separability) *For all $x \in X_\ell$ and all t,*

$$\frac{1}{2} \circ c_{x\{t\}0} + \frac{1}{2} \circ c_{0\{t\}x} \sim \frac{1}{2} \circ c_x + \frac{1}{2} \circ c_0.$$

Given that preferences over immediate consumption satisfy Independence and that the axiom considers lotteries over present equivalents, the axiom can be interpreted in a standard way via its analogy with the usual Independence condition applied for a hypothetical preference that is defined over lotteries over streams.[2]

Axiom 4 (Strong X_ℓ-Regularity) *For all $x \in X_+ \setminus \Delta_0$, the following hold.*
(i) if $x \notin X_\ell$ then $\alpha x \in X_\ell$ for some $\alpha \in (0,1]$.
(ii) if $x \in X_\ell$ then $\alpha x \in X_\ell$ for all $\alpha \in (0,1)$.

Consider the ray $\{\alpha x \,|\, \alpha \in (0,1]\}$ that contains all the mixtures that lie between x and 0. By Weak Homotheticity, the agent's impatience must be weakly increasing as we go down this ray from x to 0. Strong X_ℓ-Regularity requires that impatience is in fact strictly increasing as we go down the ray, except possibly for being constant near x. Specifically, Strong X_ℓ-Regularity (i) requires that X_ℓ should always intersect with this ray. That is, there always exists some $\alpha \in (0,1]$ for which αx exhibits ℓ-Magnitude Sensitivity.[3] Strong X_ℓ-Regularity (ii) requires in addition that if x exhibits an ℓ-Magnitude Sensitivity then so must every stream in the ray $\{\alpha x \,|\, \alpha \in (0,1]\}$.

Moreover, we impose four axioms for streams in X_ℓ.

Axiom 5 (X_ℓ-Time-Invariance) *For all $p, \widehat{p} \in \Delta$ and t, if $p^t, \widehat{p}^t \in X_\ell$, then*

$$p \succsim \widehat{p} \iff p^t \succsim \widehat{p}^t.$$

Axiom 6 (Time-0 Irrelevance) *For any $x \in X_+$ and any $p \in \Delta_0$,*

$$x \in X_\ell \implies p\{0\}x \in X_\ell.$$

[2] To illustrate, consider:

$$\frac{1}{2} \circ (0, c', 0) + \frac{1}{2} \circ (c, 0, c'') \sim \frac{1}{2} \circ (c, c', c'') + \frac{1}{2} \circ (0, 0, 0).$$

This says that the agent only cares about the distribution of consumption across periods, and not the possible correlation across periods.

[3] That is, $c_{\alpha x} \sim \alpha x \Longrightarrow \beta \circ c_{\alpha x} \succ \beta \alpha x$ for all $\beta \in (0,1)$.

Axiom 7 (X_ℓ-Dominance) *For any $x \in X_+$ and any $S \subset \{1, \cdots, T\}$ such that $x_t \succ 0$ for some $t \in S$,*

$$x \in X_\ell \implies xS0 \in X_\ell.$$

Axiom 8 (X_ℓ-Continuity) *X_ℓ is closed in $X_+ \setminus \Delta_0$.*

The first axiom requires that rankings over dated rewards in period t are independent of t for ℓ-magnitude sensitive streams. The second requires that ℓ-magnitude sensitivity of a stream x does not rely on x_0 in any way. The third states that if there is an ℓ-magnitude sensitive stream x paying positive outcomes at some periods within S, then the stream that is identical on S and paying nothing elsewhere is also ℓ-magnitude sensitive. The fourth states that the limit of a sequence of ℓ-magnitude sensitive streams is ℓ-magnitude sensitive if the limit is not an immediate reward.

3.3 Representation Theorem for Positive Streams

The representation theorem for positive streams only is provided in Noor and Takeoka [3].

Theorem 1 *A preference $\succsim$ on X_+ satisfies Weak Regularity, Weak Homotheticity, X_ℓ-Separability, Strong X_ℓ-Regularity, X_ℓ-Time-Invariance, Time-0 Irrelevance, X_ℓ-Dominance, and X_ℓ-Continuity if and only if it admits a General CCE representation.*

The positive General CCE representation has strong uniqueness properties.

Theorem 2 *If there are two positive General CE representations $(u^i, \{\varphi_t^i\}, K^i)$, $i = 1, 2$ of the same preference $\succsim$, then there exists $\alpha > 0$ such that (i) $u^2 = \alpha u^1$, (ii) $\varphi_t^2 = \alpha \varphi_t^1$, and (iii) $K^2 = \alpha K^1$.*

3.4 Monotonicity

An axiom that has been conspicuously missing is Monotonicity, that is, the condition that a stream that yields more preferred consumption in each period than another must also be preferred. Formally:

Axiom 9 (Monotonicity) *For any $x, y \in X$,*

$$(x_t, 0, .., 0) \succsim (y_t, 0, .., 0) \text{ for all } t \implies x \succsim y.$$

Moreover, if $(x_t, 0, .., 0) \succ (y_t, 0, .., 0)$ for some t, then $x \succ y$.

As shown in Noor and Takeoka [3], the CCE representation satisfies Monotonicity. However, the General CCE representation can potentially violate Monotonicity. For example, let $x, y \in \mathbb{R}_+^{T+1}$. If $y = x - (\varepsilon, \cdots, \varepsilon)$ for some $\varepsilon > 0$ but there is less cognitive resource available for x, ie, $K_x < K_y$, then it may be that the agent is unable to appreciate x

as much as he could with higher cognitive resources, causing him to exhibit $y \succ x$. The intuition is similar to the case where an agent may disprefer a stochastically dominating lottery because it is too complex to be recognized as such.

Nevertheless there is a subdomain where Monotonicity holds in the general model:

Proposition 2 *Suppose $\succsim$ admits a General CCE representation on X_+. Then, Monotonicity holds on X_ℓ and along rays.*

Proof. The first claim is implied by additive separability of the representation on X_ℓ. For the second part, take any $\alpha \in (0,1)$ and any $x \in X_+$. Let D_x and $D_{\alpha x}$ be optimal discount functions for x and αx, respectively. Note that since $D_x \cdot u(x) \geq D_{\alpha x} \cdot u(x)$, it follows that

$$U(x) = D_x \cdot u(x) \geq D_{\alpha x} \cdot u(x) > \alpha D_{\alpha x} \cdot u(x) = D_{\alpha x} \cdot u(\alpha x) = U(\alpha x),$$

establishing Monotonicity along any ray. ■

3.5 Separability

X_ℓ-Separability requires the conclusion of the axiom to hold only on X_ℓ. A natural question is what restrictions on our model are imposed by requiring Separability on all of X. This is answered by Noor and Takeoka [2, Theorem 3].

Theorem 3 *A preference $\succsim$ on X_+ satisfies Weak Regularity, Monotonicity, Weak Homotheticity, Separability and if and only if it admits a CE representation, that is, a basic tuple $(u, \{\varphi\}, K)$ with $K = \infty$ represents $\succsim$.*

The cost functions in Noor and Takeoka [2] may have kinks, which generate weakly increasing optimal discount functions, that are more general than part (3) of Proposition 1. See Noor and Takeoka [2, Theorem 3] for more details.

3.6 Convexity

In this subsection, we show that a positive General CCE representation must be star-shaped, $\alpha U(x) \geq U(\alpha x)$, which is a property weaker than convexity. It is easy to generate examples (for instance, when there are only two periods) where our model violates convexity and thus goes beyond models of convex preferences in the literature (such as Maccheroni et al [1]).

Proposition 3 *A General CCE representation U is star-shaped on X_+: for all positive x and $\alpha \in [0,1]$, $\alpha U(x) \geq U(\alpha x)$.*

Proof. By Proposition 1 (3), for all positive stream x and $\alpha \in (0,1)$, $D_x \geq D_{\alpha x}$. By linearity of u,

$$\begin{aligned} &\sum_{t\geq 1} D_x(t)u(x_t) \geq \sum_{t\geq 1} D_{\alpha x}(t)u(x_t) \\ \Longrightarrow\ &\alpha(u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t)) \geq u(\alpha \circ x_0) + \sum_{t\geq 1} D_{\alpha x}(t)u(\alpha \circ x_t) \\ \Longrightarrow\ &\alpha U(x) \geq U(\alpha x). \end{aligned}$$

■

Under a certain condition on K, a General CCE representation satisfies convexity for non-ℓ-magnitude sensitive streams.

Proposition 4 *Suppose $x, y, \alpha x + (1-\alpha)y \notin X_\ell$ for some $\alpha \in (0,1)$. If $K_{\alpha x+(1-\alpha)y} \leq \min[K_x, K_y]$, then a General CCE representation on X_+ satisfies*

$$U(\alpha x + (1-\alpha)x) \leq \alpha U(x) + (1-\alpha)U(y).$$

Note that the condition always holds if K is constant. Thus, the CCE representation of Noor and Takeoka [3] is convex on the subdomain of non-ℓ-magnitude sensitive streams.

4 Extension to Negative Outcomes

We extend the model to accommodate negative outcomes. We say that $p^* \in \Delta$ is an *absolute value* of $p \in \Delta$ if:[4] $p^* \sim p$ when p is positive , or p^* satisfies

$$\frac{1}{2} \circ p + \frac{1}{2} \circ p^* \sim 0$$

when p is negative. That is, the 50-50 lottery over p and p^* is as good as receiving 0.

For any stream x, define x^* by the stream that replaces each outcome x_t with an absolute value $(x_t)^*$, that is,

$$x_t^* = (x_t)^* \text{ for all } t.$$

Note that the absolute value p^* is not unique, since anything in its indifference class will also be an absolute value for p.

Our earlier axioms below were formulated for positive streams only. We define a *negative* stream as one that delivers negative outcomes and we impose a Symmetry axiom so that axioms on positive streams translate into restrictions on negative ones.

[4] For simplicity, we use this terminology rather than the more accurate one that p^* *has the same absolute value as* p. This terminology anticipates the fact that in terms of the representation p, p^* will satisfy $u(p^*) = |u(p)|$.

Axiom 10 (Symmetry) *If x is a negative stream then*

$$(c_x)^* = c_{x^*}.$$

Consider a negative stream x and its present equivalent c_x. The axiom states that the absolute value $(c_x)^*$ of this present equivalent is the same as the present equivalent c_{x^*} of the stream's absolute value x^*.

As noted earlier, present equivalents carry information about the agent's assessment of the outcomes and his impatience towards them. So the axiom suggests that the agent's impatience towards two streams is identical when the streams give outcomes that have identical absolute values. This suggests that if impatience is not constant, it can only change with the absolute value of outcomes.

Finally, we re-define the set of ℓ-magnitude senseitive streams X_ℓ^*:

$$X_\ell^* = \{x \in X : \alpha c_{x^*} \succ \alpha x^* \text{ for all } \alpha \in (0,1)\}.$$

That is, an arbitrary stream x is ℓ-magnitude sensitive if its absolute value x^* is such that small changes in scale leads to changes in impatience (as with Symmetry, this presumes that impatience depends only on the absolute value of the stakes). Note that

$$x \in X_\ell^* \iff x^* \in X_\ell, \; \forall \; x.$$

Axiom 11 (X_ℓ^*-Separability) *For all $x \in X_\ell^*$ and all t,*

$$\frac{1}{2} \circ c_{x\{t\}0} + \frac{1}{2} \circ c_{0\{t\}x} \sim \frac{1}{2} \circ c_x + \frac{1}{2} \circ c_0.$$

Conspicuously missing is an axiom describing streams that offer both positive and negative consumption. Because of X_ℓ^*-Separability, it turns out to be unnecessary to formulate such an axiom for ℓ-magnitude sensitive streams, since we can restrict attention to dated rewards (streams p^t that pay p at t and 0 otherwise) for much of our analysis, and such streams are either positive or negative.

Finally, we impose the following axiom:

Axiom 12 (Symmetric Hometheticity) *For any x which is neither positive nor negative and any $\alpha \in (0,1)$,*

$$c_x \sim x, \; c_{x^*} \sim x^*, \; \textit{and} \; \; \alpha \circ c_{x^*} \sim \alpha x^* \implies \alpha \circ c_x \sim \alpha x.$$

As shown in Lemmas 7 and 8 in Appendix B, Weak Homotheticity and Strong X_ℓ-Regularity jointly imply the same property for positive streams. Together with Symmetry, the same also holds for negative streams. The above axiom requires the same property for the other streams.

We establish an axiomatic foundation for a General CCE representation $(u, \{\varphi_t\}, K)$.

Theorem 4 *A non-degenerate preference $\succsim$ on X satisfies Weak Regularity, Weak Homotheticity, X_ℓ^*-Separability, Strong X_ℓ-Regularity, X_ℓ-Time-Invariance, Time-0 Irrelevance, X_ℓ-Dominance, X_ℓ-Continuity, Symmetry, and Symmetric Homotheticity if and only if it admits a General CCE representation.*

Moreover, if there are two General CCE representations $(u^i, \{\varphi_t^i\}, K^i)$, $i = 1, 2$ of the same preference $\succsim$, then there exists $\alpha > 0$ such that (i) $u^2 = \alpha u^1$, (ii) $\varphi_t^2 = \alpha\varphi_t^1$, and (iii) $K^2 = \alpha K^1$.

4.1 Special Case 1: CCE Representations with Negative Outcomes

Say that a tuple $(u, \{\varphi_t\}_{t\geq 1}, K)$ is regular if
(i) $u : \Delta \to \mathbb{R}$ is continuous and mixture linear with increasing vNM utility index $u : C \to \mathbb{R}$ satisfying (a) $u(0) = 0$ and (b) unboundedness: $u(C) = \mathbb{R}$.
(ii) for each $t \geq 1$, a cost function $\varphi_t : [0, 1] \to \mathbb{R}_+$ takes the form

$$\varphi_t(d) = a_t \cdot d^m,$$

where $m > 1$, and $a_t > 0$ is increasing in t.
(iii) $0 < K \leq a_1$.

Definition 3 (CCE Representation) *A Constrained Costly Empathy (CCE) representation is a regular tuple $(u, \{\varphi_t\}, K)$ such that $\succsim$ is represented by the function $U : X \to \mathbb{R}$ defined by*

$$U(x) = u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t), \qquad x \in X,$$

$$\text{s.t. } D_x = \arg\max_{D\in[0,\bar{d}_t]^T} \{\sum_{t\geq 1} D(t)\,|u(x_t)| - \varphi_t(D(t))\} \text{ subject to } \varphi(D) \leq K.$$

Noor and Takeoka [3] axiomatize the CCE representation which is defined over X_+.

The next axiom is introduced by Noor and Takeoka [2] in order to characterize the homogeneous CE representation.

Axiom 13 (X_ℓ-Homogeneity) *For any $x, y \in X_\ell$ s.t. $x_0 \sim y_0 \sim 0$, their present equivalents $c_x \sim x$ and $c_y \sim y$, and any $\alpha, \beta \in (0, 1)$,*

$$\beta \circ c_x \sim \alpha x \Longrightarrow \beta \circ c_y \sim \alpha y.$$

For any $x \in X$, $0\{0\}x$ denotes the stream that pays 0 in period 0 and pays according to x from period 1 onward. That is, $0\{0\}x = (0, x_1, \cdots, x_T)$. Intuitively, $0\{0\}x$ is interpreted as the future payoffs obtained from x. The following axiom is introduced by Noor and Takeoka [3].

Axiom 14 (X_ℓ-Monotonicity) *For all $x \in X_+ \setminus \Delta_0$, the following hold.*
(i) if $x \notin X_\ell$, then $y \in X_\ell$ for some $y \in X_+ \setminus \Delta_0$ with $0\{0\}x \succsim 0\{0\}y$.
(ii) if $x \in X_\ell$, then $y \in X_\ell$ for all $y \in X_+ \setminus \Delta_0$ with $0\{0\}x \succsim 0\{0\}y$.

Theorem 5 *A non-degenerate preference $\succsim$ on X satisfies Weak Regularity, Monotonicity, Weak Homotheticity, X_ℓ^*-Separability, X_ℓ-Homogeneity, X_ℓ-Monotonicity, Symmetry, and Symmetric Homotheticity if and only if it admits a CCE representation.*

4.2 Special Case 2: CE Representations with Negative Outcomes

Definition 4 (CE Representation) *A Costly Empathy (CE) representation is a basic tuple $(u, \{\varphi_t\})$ such that $\succsim$ is represented by the function $U : X \to \mathbb{R}$ defined by*

$$U(x) = u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t), \qquad x \in X,$$

$$s.t.\ D_x = \arg\max_{D\in[0,\overline{d}_t]^T} \{\sum_{t\geq 1} D(t)\,|u(x_t)| - \varphi_t(D(t))\}.$$

Noor and Takeoka [2] axiomatize the CE representation which is defined over X_+.

Theorem 6 *A non-degenerate preference $\succsim$ on X satisfies Weak Regularity, Monotonicity, Weak Homotheticity, Separability, and Symmetry if and only if it admits a CE representation.*

Proof. By Lemma 1, there exists an expected utility representation u on lotteries. There exists a representation U on X which is an extension of u. Since $\succsim$ satisfies Separability on the whole domain X, the proof of Lemma 4 goes through without X_ℓ^*-Dominance. The lemma ensures that $U : X \to \mathbb{R}$ admits an additively separable form where each component function U_t is defined on the whole Δ. By replacing X_ℓ^*-Time-Invariance with Monotonicity, together with Lemma 2, Lemma 5 implies that $U : X \to \mathbb{R}$ admits a GDU representation such that the discount function depends only on the absolute value of payoffs. By a similar argument, we can show that the discount function is weakly increasing in positive payoffs by Weak Homotheticity. The rest of the proof is the same as in Theorem 3 of Noor and Takeoka [2]. ■

A Appendix: Proof of Proposition 1

By property (c) of the General CCE representation, $K_x = K_{x^*} = K_{|u(x)|}$. Since the optimization problem depends only on the absolute value of $(u(x_t))_{t\geq 1}$, it is enough to show the statement for positive streams. Moreover, by property (c), $K_{\overline{x}} = K_x$ if $\overline{x}, x$ belong to the same ray. Thus, denote $K = K_x$ for some (any) x on the ray.

For a positive stream x on the ray, its optimal discount factor D_x is characterized by the FOC of the following Lagrangian:

$$\mathcal{L} = \sum_{t\geq 1} D(t)u(x_t) - \sum_{t\geq 1} \varphi_t(D(t)) + \lambda(K - \sum_{t\geq 1} \varphi_t(D(t))),$$

where $\lambda \geq 0$ is a Lagrange multiplier.

If the capacity constraint is not binding, $\lambda = 0$. By differentiating $\mathcal{L}$ with respect to $D(t)$, an optimal D_x satisfies

$$u(x_t) = \varphi_t'(D_x(t)),\ \forall t \geq 1 \text{ with } u(x_t) > 0.$$

We have a closed-form solution such as

$$D_x(t) = D_{u(x_t)}(t) = (\varphi_t')^{-1}(u(x_t)).$$

In particular, $D_x(t)$ depends only on $u(x_t)$.

If the capacity constraint is binding, by differentiating $\mathcal{L}$ with respect to $D(t)$, an optimal D_x satisfies

$$u(x_t) = (1+\lambda)\varphi_t'(D_x(t)),\ \forall t \geq 1 \text{ with } u(x_t) > 0, \tag{2}$$

$$\sum_{t\geq 1} \varphi_t(D_x(t)) = K \tag{3}$$

for some $\lambda \geq 0$. From (2), $D_x(t) = (\varphi_t')^{-1}(u(x_t)/(1+\lambda))$. By substituting it into (3),

$$\sum_{t\geq 1} \varphi_t \left((\varphi_t')^{-1} \left(\frac{u(x_t)}{1+\lambda} \right) \right) = K.$$

This equation is solved for λ, which is denoted by $\lambda(x, K)$. Then, we obtain

$$D_x(t) = (\varphi_t')^{-1} \left(\frac{u(x_t)}{1+\lambda(x,K)} \right).$$

Note that $D_x(t)$ depends on the whole stream x, not only on the payoff at t, as well as the capacity cap K.

Take any other y on the ray such that the capacity constraint is binding at its optimal D_y. From the FOC,

$$u(y_t) = (1+\lambda')\varphi_t'(D_y(t)),\ \forall t \geq 1 \text{ with } u(x_t) > 0,$$

$$\sum_{t\geq 1} \varphi_t(D_y(t)) = K$$

for some positive λ'. Since x and y belong to the same ray, $y = \alpha x$ for some $\alpha > 0$. Since u is linear, by setting $1 + \lambda' = \alpha(1+\lambda)$, D_y satisfies the FOC of (2). Thus, $D_y = D_x$.

(3) It is implied from (1) and (2).

B Appendix: Proof of the General CCE Representation

The first four subsections establish sufficiency of the axioms. The fifth subsection establishes necessity. The last subsection shows uniqueness. For a positive General CCE representation, it is enough to follow the arguments of the first three subsections with replacing X_ℓ^* by X_ℓ.

B.1 Preliminaries

Lemma 1 *The preference $\succsim|_{\Delta_0}$ is represented by a utility function $u : \Delta \to \mathbb{R}$ with $u(0) = 0$ which is continuous, mixture linear, and homogeneous (that is, $u(\alpha p) = \alpha u(p)$ for all $\alpha \geq 0$.) Moreover, the preference $\succsim$ on X is represented by a continuous utility function $U : X \to \mathbb{R}$ such that $U(p) = u(p)$ for all $p \in \Delta_0$.*

Proof. By Weak Regularity, $\succsim|_{\Delta_0}$ satisfies the vNM axioms. There exists a continuous mixture linear function $u : \Delta \to \mathbb{R}$ which represents $\succsim|_{\Delta_0}$ and which can be chosen so that $u(0) = 0$.

Establish homogeneity of u next. If $\alpha \in [0, 1]$, by mixture linearity of u, together with identifying αp with $\alpha p + (1 - \alpha)0$,

$$u(\alpha p) = u(\alpha p + (1 - \alpha)0) = \alpha u(p) + (1 - \alpha)u(0) = \alpha u(p).$$

If $\alpha > 1$, we identify αp with $p' \in \Delta$ satisfying $p = \frac{1}{\alpha}p' + \frac{\alpha-1}{\alpha}0$. Then, mixture linearity of u implies that $u(p) = \frac{1}{\alpha}u(p')$, that is, $u(\alpha p) = u(p') = \alpha u(p)$, as desired.

For any $x \in X$, the Present Equivalents axiom ensures that there exists $c_x \in C$ such that $c_x \sim x$. Define $U(x) = u(c_x)$. By construction, U represents $\succsim$. Moreover, for all $p \in \Delta$, $U(p) = u(p)$. In particular, we have $U(0) = u(0) = 0$.

To show the continuity of U, take any sequence $x^n \to \widehat{x}$. There exists a corresponding present equivalent $c_{x^n} \sim x^n$. Since $U(x^n) = u(c_{x^n})$ and u is continuous, we want to show that $c_{x^n} \to c_{\widehat{x}}$.

Claim 1 *The present equivalent is continuous, that is, if $x^n \to x$, then $c_{x^n} \to c_{\widehat{x}}$.*

Proof. Take any $\overline{c}$ and $\underline{c}$ such that $\overline{c} > c_{\widehat{x}} > \underline{c}$. Let $W = \{x \in X \,|\, \overline{c} \succ x \succ \underline{c}\}$. Since $x^n \to \widehat{x} \sim c_{\widehat{x}}$, by Continuity, we can assume $x^n \in W$ for all n without loss of generality.

Seeking a contradiction, suppose $c_{x^n} \not\to c_{\widehat{x}}$. Then, there exists a neighborhood of $c_{\widehat{x}}$, denoted by $B(c_{\widehat{x}})$, such that $c_{x^m} \notin B(c_{\widehat{x}})$ for infinitely many m. Let $\{x^m\}$ denote the corresponding subsequence of $\{x^n\}$. Since $x^n \to \widehat{x}$, $\{x^m\}$ also converges to $\widehat{x}$. Without loss of generality, we can assume $x^m \in W$, that is, $\overline{c} \succ x^m \sim c_{x^m} \succ \underline{c}$. By C-Monotonicity, $\overline{c} > c_{x^m} > \underline{c}$. Thus, $\{c_{x^m}\}$ belongs to a compact interval $[\underline{c}, \overline{c}]$, and hence, there exists a convergent subsequence $\{c_{x^\ell}\}$ with a limit $\widetilde{c} \neq c_{\widehat{x}}$. On the other hand, since $x^\ell \to \widehat{x}$ and $x^\ell \sim c_{x^\ell}$, Continuity implies $\widehat{x} \sim \widetilde{c}$. Since $c_{\widehat{x}}$ is unique, $c_{\widehat{x}} = \widetilde{c}$, which is a contradiction. ■

The symmetric argument can be applied for the case that $0 \succsim \overline{x}, x^n$ for all n. Finally, suppose that $\overline{x} \sim 0$. If $x^n \sim 0$ for some n, $U(x^n) = 0 = U(\overline{x})$ for such n. Thus, we can assume without loss of generality that $x^n \not\sim 0$ for all n. For the subsequence $\{x^m\}$ of $\{x^n\}$ satisfying $x^m \succ 0$, we have $x^m \to \overline{x}$. By the above argument, $U(x^m) \to U(\overline{x})$. Similarly, for the subsequence $\{x^m\}$ of $\{x^n\}$ satisfying $0 \succ x^m$, we have $U(x^m) \to U(\overline{x})$. Therefore, $U(x^n) \to U(\overline{x})$, as desired. ■

We show several properties of an absolute value of streams.

Lemma 2 *(1) For all negative outcomes $p \in \Delta$, $u(p) = -u(p^*)$.*
(2) For any dated reward p^t with a negative outcome, $(c_{p^t})^* \sim (p^*)^t$.
(3) For all $x \in X$, if x^* is an absolute value of x, αx^* is an absolute value of αx for all $\alpha > 0$.

Proof. (1) If $0 \succ p$, by definition, its absolute value $p^* \in \Delta$ satisfies

$$\frac{1}{2} \circ p + \frac{1}{2} \circ p^* \sim 0.$$

Since u is mixture linear, $u(p) = -u(p^*)$.

(2) Since the dated reward p^t with this negative outcome is a negative stream, by Symmetry, $(c_{p^t})^* = c_{(p^t)^*}$. Since $(p^t)^* = (p^*)^t$ by definition, we have a desired result.

(3) By part (1), for any negative outcome $p \in \Delta$, its absolute value $p^* \in \Delta$ satisfies $u(p) = -u(p^*)$. Since u is homogeneous, for all negative outcomes $p \in \Delta$ and $\alpha > 0$, we have

$$u(\alpha p) = \alpha u(p) = -\alpha u(p^*) = -u(\alpha p^*),$$

that is, αp^* is an absolute value of αp. Thus, the claim holds by definition. ■

Axiom 15 (X_ℓ^*-Time-Invariance) *For all $p, \widehat{p} \in \Delta$ and t, if $p^t, \widehat{p}^t \in X_\ell^*$, then*

$$p \succsim \widehat{p} \iff p^t \succsim \widehat{p}^t.$$

Axiom 16 (X_ℓ^*-Dominance) *For all x and all $S \subset \{1, \cdots, T\}$ such that $x_t \not\sim 0$ for some $t \in S$,*

$$x \in X_\ell^* \implies xS0 \in X_\ell^*.$$

Lemma 3 *$\succsim$ satisfies X_ℓ^*-Time-Invariance and X_ℓ^*-Dominance.*

Proof. By X_ℓ-Time-Invariance, the statement holds for all positive outcomes. If $p \succsim 0 \succsim \widehat{p}$, p^t is a positive stream and $\widehat{p}^t$ is a negative stream. So, we have $p^t \succsim 0 \succsim \widehat{p}^t$. Thus, for negative outcomes $p, \widehat{p}$ with $p \succsim \widehat{p}$, assume $p^t, \widehat{p}^t \in X_\ell^*$. By definition of X_ℓ^*, $(p^t)^* = (p^*)^t \in X_\ell$ and $(\widehat{p}^t)^* = (\widehat{p}^*)^t \in X_\ell$. Moreover, by Lemma 2 (1), $\widehat{p}^* \succsim p^*$. X_ℓ-Time-Invariance implies that $(\widehat{p}^*)^t \succsim (p^*)^t$. By Lemma 2 (2), $(c_{\widehat{p}^t})^* \succsim (c_{p^t})^*$. Since $c_{\widehat{p}^t}$ and c_{p^t} are negative outcomes, again by Lemma 2 (1), $c_{p^t} \succsim c_{\widehat{p}^t}$, or equivalently, $p^t \succsim \widehat{p}^t$, as desired.

Note that $x \in X_\ell^*$ if and only if $x^* \in X_\ell$. Since x^* is a positive stream, X_ℓ-Dominance implies that $x^*S0 \in X_\ell$. Thus, we have $xS0 \in X_\ell^*$. ■

B.2 Representation on X_ℓ^*

For each $t \geq 1$, let $\Delta_t = \{p \in \Delta \,|\, p^t \in X_\ell^*\}$.

Lemma 4 *On the subdomain $X_\ell^* \cup \Delta_0 \subset X$, U can be written as an additively separable utility form, i.e. $U : X_\ell^* \cup \Delta_0 \to \mathbb{R}$ s.t. for all $x \in X_\ell^* \cup \Delta_0$,*

$$U(x) = u(x_0) + \sum_{t \geq 1} U_t(x_t),$$

where u is given as in Lemma 1 and $U_t : \Delta_t \to \mathbb{R}$ are continuous with $U_t(0) = 0$ for each t. Moreover, u is unbounded.

Proof. Take any $x \in X_\ell^*$, which is denoted by $x = (x_0, x_1, \cdots, x_T)$. There exists some $t > 0$ with $x_t \not\sim 0$. We start with the case where there are two $x_t, x_s \not\sim 0$. By notational convenience, denote such a stream by $(x_t, x_s, 0, \cdots, 0)$. By X_ℓ^*-Separability,

$$\frac{1}{2} \circ c_{(0,x_s,0,\cdots,0)} + \frac{1}{2} \circ c_{(x_t,0,\cdots,0)} \sim \frac{1}{2} \circ c_{(x_t,x_s,0,\cdots,0)} + \frac{1}{2} \circ 0.$$

Since u is mixture linear,

$$\begin{aligned} & u(c_{(0,x_s,0,\cdots,0)}) + u(c_{(x_t,0,\cdots,0)}) = u(c_{(x_t,x_s,0,\cdots,0)}) + u(0) \\ \iff & U(0, x_s, 0, \cdots, 0) + U(x_t, 0, \cdots, 0) = U(x_t, x_s, 0, \cdots, 0). \end{aligned}$$

Define $U_t(x_t) = U(x_t, 0, \cdots, 0)$ and $U_s(x_s) = U(0, x_s, 0, \cdots, 0)$. Then, we have

$$U(x_t, x_s, 0, \cdots, 0) = U_t(x_t) + U_s(x_s). \tag{4}$$

If a stream has three outcomes $x_t, x_s, x_r \not\sim 0$, denote it by $(x_t, x_s, x_r, 0, \cdots, 0)$. By X_ℓ^*-Dominance, $(x_t, x_s, 0, \cdots, 0) \in X_\ell^*$. From the above argument, we have (4). By X_ℓ^*-Separability,

$$\frac{1}{2} \circ c_{(0,0,x_r,0,\cdots,0)} + \frac{1}{2} \circ c_{(x_t,x_s,0,\cdots,0)} \sim \frac{1}{2} \circ c_{(x_t,x_s,x_r,0,\cdots,0)} + \frac{1}{2} \circ 0.$$

Since u is mixture linear,

$$\begin{aligned} & u(c_{(0,0,x_r,0,\cdots,0)}) + u(c_{(x_t,x_s,0,\cdots,0)}) = u(c_{(x_t,x_s,x_r,0,\cdots,0)}) + u(0) \\ \iff & U(0, 0, x_r, 0 \cdots, 0) + U(x_t, x_s, 0, \cdots, 0) = U(x_t, x_s, x_r, 0, \cdots, 0). \end{aligned}$$

Define $U_r(x_r) = U(0, 0, x_r, 0, \cdots, 0)$. Then, we have

$$\begin{aligned} U(x_t, x_s, x_r, 0, \cdots, 0) &= U_r(x_r) + U(x_t, x_s, 0, \cdots, 0) \\ &= U_t(x_t) + U_s(x_s) + U_r(x_r). \end{aligned}$$

By repeating the same argument finitely many times, we have

$$U(x)=\sum_{t\geq 0} U_t(x_t),$$

where $U_t(x_t)$ is defined as $U_t(x_t)=U(0,\cdots,0,x_t,0,\cdots,0)$. By definition, $U_t(0)=0$. By X^*_ℓ-Dominance, for any $x\in X^*_\ell$, if $x_t\not\sim 0$, $(x_t)^t\in X^*_\ell$, that is, $(x_t)^t\in\Delta_t$. Hence, U_t is defined on Δ_t.

Since U is continuous, U_t is also continuous. Take any $p\in\Delta$ and any sequence $x^n=(0,x^n_1,\cdots,x^n_T)\in X^*_\ell$, where $x^n_t\to 0$ for all $t\geq 1$. By Time-0 Irrelevance, $p\{0\}x^n=(p,x^n_1,\cdots,x^n_T)\in X^*_\ell$. Since $p\{0\}x^n\to p\in\Delta_0$, by continuity, $U(p\{0\}x^n)\to u(p)$ and $U(p\{0\}x^n)=U_0(p)+\sum_{t\geq 1}U_t(x^n_t)\to U_0(p)$. Thus, $U_0(p)=u(p)$.

Finally, we show that u must be unbounded. First, we show that u is unbounded from above. By seeking a contradiction, suppose otherwise. Then, the range of u is nonempty and has an upper bound. There exists a supremum $\overline{v}$ of the range of u. Since U_t is non-constant by Time Invariance, there exists some $\tilde{p}\in\Delta$ with $U_t(\tilde{p})>0$. Take a lottery $\overline{p}\in\Delta$ such that $\overline{v}-u(\overline{p})<U_t(\tilde{p})$. Consider the stream $\overline{x}$ which pays $\overline{p}$ in period 0, $\tilde{p}$ in period t, and zero otherwise. By Time-0 Irrelevance, $\overline{x}\in X^*_\ell$. By the representation,

$$U(\overline{x})=u(\overline{p})+U_t(\tilde{p})>\overline{v}.$$

Since $\overline{v}$ is the supremum of $u(\Delta)$, the above inequality contradicts to the Present Equivalents axiom. By the symmetric argument, we can show that u is unbounded from below. ■

Lemma 5 *The function $U: X^*_\ell\cup\Delta_0\to\mathbb{R}$ defined as in Lemma 4 can be written as follows:*

$$U(x)=u(x_0)+\sum_{t\geq 1} D_{|u(x_t)|}(t)u(x_t),$$

where for all $t\geq 1$, $D_{|u(p)|}(t)\in[0,1]$ and $D_{|u(p)|}(t)$ is continuous and strictly increasing in $|u(p)|$.

Proof. Taking the additive representation from Lemma 4, by X^*_ℓ-Time-Invariance, we have that $U_t(x_t)$ can be written as an increasing transformation of $u(x_t)$. So we can write $U_t(x_t)$ as $U_t(u(x_t))$. Define D_x by $D_{u(x_t)}(t)=\frac{U_t(u(x_t))}{u(x_t)}>0$ for any $x_t\in\Delta$ with $x_t\not\sim 0$. Define $\underline{d}_t=\inf\{D_{u(p)}(t): 0\not\sim p\in\Delta_t\}$. Then

$$U(x)=u(x_0)+\sum_{t\geq 1} D_{u(x_t)}(t)u(x_t),\ \text{for all } x\in X^*_\ell\cup\Delta_0.$$

By Lemma 2, the representation implies

$$\begin{aligned}D_{u(p)}(t)u(p)&=U(p^t)=u(c_{p^t})=-u((c_{p^t})^*)=-u(c_{(p^t)^*})\\&=-U((p^t)^*)=-D_{u(p^*)}(t)u(p^*)=D_{|u(p)|}(t)u(p),\end{aligned}$$

and hence, $D_{u(p)}(t) = D_{|u(p)|}(t)$, as desired.

To see that $D_{|u(p)|}(t)$ is strictly increasing in $|u(p)|$, it suffices to show by the above observation that $D_{u(p)}(t)$ is strictly increasing in $u(p) > 0$. Note that for any positive stream $x \in X_\ell$ and its present equivalent c_x, by definition of X_ℓ, $\alpha U(c_x) > U(\alpha x)$ for all $\alpha \in (0,1)$ and thus $\alpha U(x) > U(\alpha x)$. Applying this more specifically to a dated reward p^t with $u(p) > 0$ and exploiting mixture linearity of u, we obtain $\alpha D_{u(p)}(t)u(p) > D_{u(\alpha p)}(t)u(\alpha p) = \alpha D_{\alpha u(p)}(t)u(p)$ and thus

$$D_{u(p)}(t) > D_{\alpha u(p)}(t), \text{for all } \alpha \in (0,1),$$

as desired.

Since u and U_t are continuous, so is $D_{u(p)}(t)$ in $u(p)$ on the domain of $u(p) \neq 0$. Moreover, since $|u(p)|$ is continuous, so is $D_{|u(p)|}(t)$ for all $|u(p)| \neq 0$. Since $\underline{d}_t$ is defined as $\inf\{D_{u(p)}(t) : 0 \not\sim p \in \Delta_t\}$ and $D_{|u(p)|}(t)$ is strictly increasing in $|u(p)|$, $D_{|u(p)|}(t)$ is indeed continuous for all $|u(p)| \geq 0$.

By Impatience, for all positive p and $t \geq 1$, $u(p) = U(p^0) \geq U(p^t) = D_{|u(p)|}(t)u(p)$, which implies $D_{|u(p)|}(t) \leq 1$. ■

Lemma 6 *The function $U : X^*_\ell \cup \Delta_0 \to \mathbb{R}$ appeared in Lemma 5 can be written as follows:*

$$U(x) = u(x_0) + \sum_{t \geq 1} D_x(t)u(x_t),$$

$$s.t. \qquad D_x = \arg\max_D \{\sum_{t\geq 1} \big(D(t)|u(x_t)| - \varphi_t(D(t))\big)\}$$

where for each $t \geq 1$, $\varphi_t : [0,1] \to \mathbb{R}_+ \cup \{\infty\}$ is an increasing convex function that is strictly increasing, strictly convex, and differentiable on $\{d \,|\, 0 < \varphi_t(d) < \infty\}$, and satisfies $\varphi_t(\underline{d}_t) = 0$ and $\varphi'_t(\underline{d}_t) = 0$. Moreover, $\varphi_t(d) \leq \varphi_{t+1}(d)$ for all $t < T$ and d.

Proof. By X^*_ℓ-Dominance, if $x \in X^*_\ell$, then $xt0 \in X^*_\ell$, that is, $x^*t0 \in X_\ell$ if $x_t \not\sim 0$. Thus, φ_t can be derived from the positive dated rewards at t as follows. Define

$$S_t = \{d \in [0,1] \,|\, d = D_{|u(p)|}(t) \text{ for some } p^t \in X_\ell\}.$$

By X_ℓ-Regularity, if $p^t \in X_\ell$, then $\alpha p^t \in X_\ell$ for all $\alpha \in (0,1)$. Thus, S_t is an interval. Note $\underline{d}_t = \inf S_t$. Denote $\overline{d}_t = \sup S_t$. Define $I_t = S_t \cup \{\overline{d}_t, \underline{d}_t\}$. The cost function φ_t on I_t is implicitly defined by the first order condition

$$|u(p)| = \varphi'_t(D_{|u(p)|}(t)), \tag{5}$$

along with the assumption that $\varphi_t(\underline{d}_t) = 0$. Moreover, the continuity of $D_{|u(p)|}(t)$ wrt $|u(p)|$ requires that $0 = \varphi'_t(\underline{d}_t)$. The function is by construction once differentiable and has a positive slope. Since $D_{|u(p)|}(t)$ is strictly increasing in $|u(p)|$, (5) implies that φ'_t is strictly increasing, and hence, φ_t is strictly convex.

By construction, the set $\arg\max_D\{\sum\big(D(t)|u(x_t)| - \varphi_t(D(t))\big)\}$ is nonempty and moreover, it is a singleton since $\sum\big(D(t)|u(x_t)| - \varphi_t(D(t))\big)$ is a strictly concave function of D. Thus D_x is a unique solution.

The cost function can be extended to $[0,1]$ by

$$\varphi_t(d) = \begin{cases} 0 & \text{if } d \in [0, \underline{d}_t) \\ \varphi_t(d) & \text{if } d \in I_t \\ \infty & \text{if } d \in (\overline{d}_t, 1] \end{cases}.$$

Then, φ_t is increasing and convex on $[0,1]$.

By Impatience, for all positive p and for all $t < T$, $D_{u(p)}(t)u(p) = U(p^t) \geq U(p^{t+1}) = D_{u(p)}(t+1)u(p)$. Thus, $D_{u(p)}(t)$ is weakly decreasing wrt t. This observation implies that the effective domain *eff* (φ_t) of φ_t includes that of φ_{t+1}. For any $d := D_{u(c)}(t+1)$ in the effective domain of φ_{t+1}, it follows from the FOC that

$$\varphi_t'(d) \leq \varphi_t'(D_{u(p)}(t)) = u(p) = \varphi_{t+1}'(D_{u(p)}(t+1)) = \varphi_{t+1}'(d),$$

that is, $\varphi_t'(d) \leq \varphi_{t+1}'(d)$ for all $d \in$ *eff* (φ_{t+1}). By integrating both functions we obtain $\varphi_t(d) \leq \varphi_{t+1}(d)$ for all $d \in$ *eff* (φ_{t+1}). Consequently, $\varphi_t(d) \leq \varphi_{t+1}(d)$ for all $d \in [0,1]$. ■

B.3 Extending the Representation to X_+

Recall that X_+ is the set of positive streams. We extend the representation on $X_\ell \cup \Delta_0$ to X_+. The first lemma states that X_ℓ has a boundary point on a ray.

Lemma 7 *For any stream $x \in X_+ \setminus \Delta_0$, there exists a unique $\alpha_x \in (0,1]$ such that*

$$\begin{cases} \alpha \leq \alpha_x & \Longrightarrow \quad \alpha x \in X_\ell, \\ \alpha > \alpha_x & \Longrightarrow \quad \alpha x \notin X_\ell. \end{cases}$$

Proof. Let $A = \{\alpha \in (0,1] \,|\, \alpha x \in X_\ell\}$. By part (i) of Strong X_ℓ-Regularity, $A \neq \emptyset$. Let $\alpha_x = \sup A$. We claim that A is an interval with $\inf A = 0$. Take any $\alpha \in A$ and $\beta \in (0,\alpha)$. Since $\alpha x \in X_\ell$, by part (ii) of Strong X_ℓ-Regularity, $\beta x = \frac{\beta}{\alpha}(\alpha x) \in X_\ell$, that is, $\beta \in A$ as desired. Now, by definition of α_x, if $\alpha < \alpha_x$, then $\alpha \in A$, and hence $\alpha x \in X_\ell$. If $\alpha > \alpha_x$, then $\alpha \notin A$, and hence $\alpha x \notin X_\ell$. Uniqueness of α_x is obvious. Moreover, if $x \in X_\ell$, by part (ii) of Strong X_ℓ-Regularity, $A = (0,1)$, and hence, $\alpha_x = 1$. ■

Lemma 8 *For any $x \in X_+ \setminus \Delta_0$, take $\alpha_x \in (0,1]$ which is defined as in Lemma 7. Then,*

$$\begin{cases} \alpha < \alpha_x & \Longrightarrow \quad \alpha \circ c_x \succ \alpha x, \\ \alpha \geq \alpha_x & \Longrightarrow \quad \alpha \circ c_x \sim \alpha x. \end{cases}$$

Proof. Step 1: For all $x \in X_+ \setminus \Delta_0$, $\alpha \circ c_x \succ \alpha x$ implies $\beta \circ c_x \succ \beta x$ for all $\beta \in (0,\alpha]$. By definition, a present equivalent of αx, denoted by $c_{\alpha x}$, satisfies $\alpha \circ c_x \succ \alpha x \sim c_{\alpha x}$. For any $\gamma \in (0,1)$, let $\beta = \gamma\alpha \in (0,\alpha)$. By Weak Homotheticity and Risk Preference,

$$\beta \circ c_x = \gamma\alpha \circ c_x \succ \gamma \circ c_{\alpha x} \succsim \gamma\alpha x = \beta x,$$

as desired.

Step 2: If there exist $\alpha, \beta \in (0,1)$ such that $\alpha \circ c_x \sim \alpha x$ and $\beta \circ c_{\alpha x} \sim \beta(\alpha x)$, then $\alpha\beta \circ c_x \sim \alpha\beta x$. By definition and the assumption, $\alpha \circ c_x \sim \alpha x \sim c_{\alpha x}$. By Risk Preference, $\alpha\beta \circ c_x \sim \beta \circ c_{\alpha x}$. Hence, by assumption, $\alpha\beta \circ c_x \sim \alpha\beta x$.

Step 3: There exists a unique $\widetilde{\alpha}_x \in (0,1]$ such that

$$\begin{cases} \alpha < \widetilde{\alpha}_x & \Longrightarrow \quad \alpha \circ c_x \succ \alpha x, \\ \alpha \geq \widetilde{\alpha}_x & \Longrightarrow \quad \alpha \circ c_x \sim \alpha x. \end{cases}$$

If $x \in X_\ell$, $\widetilde{\alpha}_x = 1$ satisfies this condition. Thus, assume $x \notin X_\ell$. Let $\widetilde{A} = \{\alpha \in (0,1] \,|\, \alpha \circ c_x \succ \alpha x\}$. By part (i) of Strong X_ℓ-Regularity, $\widetilde{A}$ is non-empty. Moreover, by Step 1, $\widetilde{A}$ is an interval with $\inf \widetilde{A} = 0$. Let $\widetilde{\alpha}_x$ be a supremum of $\widetilde{A}$. If $\widetilde{A} = (0,1)$, $\widetilde{\alpha}_x = 1$ and this $\widetilde{\alpha}_x$ satisfies the desired property. If $\widetilde{A}$ is a proper subset of $(0,1)$, $\widetilde{\alpha}_x < 1$. Then, there exists a sequence $\alpha^n \to \widetilde{\alpha}_x$ with $\alpha^n > \widetilde{\alpha}_x$. Since $\alpha^n \circ c_x \sim \alpha^n x$, by Continuity, $\widetilde{\alpha}_x \circ c_x \sim \widetilde{\alpha}_x x$, as desired.

Step 4: $\widetilde{\alpha}_x \leq \alpha_x$. Seeking a contradiction, suppose $\widetilde{\alpha}_x > \alpha_x$. Lemma 7 implies $\widetilde{\alpha}_x x \notin X_\ell$. By definition, there exists $\beta \in (0,1)$ such that $\beta \circ c_{\widetilde{\alpha}_x x} \sim \beta(\widetilde{\alpha}_x x)$. Since $\widetilde{\alpha}_x \circ c_x \sim \widetilde{\alpha}_x x$, by Step 2, $\widetilde{\alpha}_x \beta \circ c_x \sim \widetilde{\alpha}_x \beta x$. Since $\widetilde{\alpha}_x \beta < \widetilde{\alpha}_x$, this contradicts to Step 3.

Step 5: $\widetilde{\alpha}_x = \alpha_x$. By Step 4, seeking a contradiction, suppose $\widetilde{\alpha}_x < \alpha_x$. Take any $\alpha \in (\widetilde{\alpha}_x, \alpha_x)$. By Step 3, $\alpha \circ c_x \sim \alpha x$. Moreover, for all γ sufficiently close to one, since $\gamma\alpha \in (\widetilde{\alpha}_x, \alpha_x)$, $\gamma\alpha \circ c_x \sim \gamma\alpha x$. Now, by definition, $c_{\alpha x} \sim \alpha x$, which implies $c_{\alpha x} \sim \alpha \circ c_x$. Since $\alpha x \in X_\ell$ by Lemma 7, for all $\gamma \in (0,1)$, $\gamma \circ c_{\alpha x} \succ \gamma\alpha x$. Thus, we have

$$\gamma \circ c_{\alpha x} \succ \gamma\alpha x \sim \gamma\alpha \circ c_x$$

for all γ sufficiently close to one. By Risk Preference, $c_{\alpha x} \succ \alpha \circ c_x$, which is a contradiction. ■

Lemma 9 *For all $x, y \in X_+ \setminus \Delta_0$, take $\alpha_x, \alpha_y \in (0,1]$ which are defined as in Lemma 7. If $x_t \sim y_t$ for all $t \geq 1$, then $\alpha_x = \alpha_y$.*

Proof. By Lemma 6, the representation depends only on utility streams $(u(x_t))_{t=0}^T$. Moreover, by Time-0 Irrelevance, x_0 is independent of whether x is ℓ-magnitude sensitive. Since $u(x_t) = u(y_t)$ for all $t \geq 1$, x is ℓ-magnitude sensitive if and only if so is y. If $x, y \in X_\ell$, $\alpha_x = \alpha_y = 1$. Assume next that $x, y \notin X_\ell$. Seeking a contradiction, suppose that $\alpha_x \neq \alpha_y$. Without loss of generality, let $\alpha_x > \alpha_y$. For any $\alpha \in (\alpha_y, \alpha_x)$, by Lemma 7, αx is ℓ-magnitude sensitive and αy is not ℓ-magnitude sensitive. Since $u(\alpha x_t) = u(\alpha y_t)$ for all t, this contradicts to the above argument. Thus, $\alpha_x = \alpha_y$, as desired. ■

As shown in Lemma 7, for any $x \in X_+ \setminus \Delta_0$,

$$\alpha_x = \sup\{\alpha \in [0,1] \,|\, \alpha x \in X_\ell\}.$$

Lemma 10 *The function $U : X_+ \to \mathbb{R}_+$ appeared in Lemma 1 can be written as*

$$U(x) = u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t),$$

$$s.t. \qquad D_x = \begin{cases} \arg\max_D \{\sum_{t\geq 1} D(t)u(x_t) - \varphi_t(D(t))\} & \text{if } x \in X_\ell \cup \Delta_0, \\ D_{\alpha_x x} & \text{if } x \notin X_\ell \cup \Delta_0. \end{cases}$$

Proof. By Lemma 6, U has the desired form on $X_\ell \cup \Delta_0$. Consider the case of $x \notin X_\ell \cup \Delta_0$. Since $u(\alpha_x \circ c_x) = U(\alpha_x x)$ by Lemma 8,

$$U(x) = u(c_x) = \frac{1}{\alpha_x} U(\alpha_x x). \tag{6}$$

By the representation on X_ℓ,

$$U(\alpha_x x) = u(\alpha_x \circ x_0) + \sum_{t\geq 1} D_{\alpha_x x}(t)u(\alpha_x \circ x_t). \tag{7}$$

By combining (6) with (7),

$$\begin{aligned} U(x) = \frac{1}{\alpha_x} U(\alpha_x x) &= \frac{1}{\alpha_x}\left(u(\alpha_x \circ x_0) + \sum_{t\geq 1} D_{\alpha_x x}(t)u(\alpha_x \circ x_t)\right) \\ &= u(x_0) + \sum_{t\geq 1} D_{\alpha_x x}(t)u(x_t), \end{aligned}$$

as desired. ■

From now on, we derive a function $K : X_+ \setminus \Delta_0 \to \mathbb{R}_{++} \cup \{\infty\}$ which serves as a general capacity constraint for the General CCE representation.

First, consider the case of $X_\ell = X_+ \setminus \Delta_0$. Since $x \in X_\ell \cup \Delta_0$ for all x, Lemma 10 directly delivers the desired representation by setting $K_x = \infty$ for all $x \in X_+ \setminus \Delta_0$. The CCE representation in this case is additively separable on the whole domain.

From now on, assume $X_\ell \subsetneq X_+ \setminus \Delta_0$. Let

$$\varphi(D) := \sum_{t\geq 1} \varphi_t(D(t)).$$

Lemma 11 *There is a function $K : X_+ \setminus \Delta_0 \to \mathbb{R}_{++} \cup \{\infty\}$ such that $\succsim$ is represented by*

$$U(x) = u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t),$$

$$s.t. \qquad D_x = \arg\max_{D\in\Lambda_x} \{\sum D(t)u(x_t) - \varphi_t(D(t))\}$$

$$\Lambda_x := \{D \in [0,1]^T \,|\, \varphi(D) \leq K_x\}.$$

Moreover, (1) the function K_x satisfies $K_x = K_{\lambda x}$ for any x and λ, and (2) for all positive streams x, y, if $u(x_t) = u(y_t)$ for all $t \geq 1$, then $K_x = K_y$.

Proof. Since $U(p) = u(p)$ for all $p \in \Delta_0$, K does not play any role for consumption stream on Δ_0. Take any $x \in X_+ \setminus \Delta_0$. If $\lambda x \in X_\ell$ for all $\lambda > 0$, define $K_x = K_{\lambda x} = \infty$ for all $\lambda > 0$. Otherwise, we can find another x on the same ray with $x \notin X_\ell$. For such x, define

$$K_x := \varphi(D_{\alpha_x x}) < \infty.$$

Extend to X_ℓ by requiring $K_x = K_{\lambda x}$ for any $\lambda > 0$.

For all $x \in X_+ \setminus \Delta_0$, by Lemma 7, there exists $\alpha_x > 0$ such that $\alpha_x x \in X_\ell$. For any $\beta \in (0, \alpha_x)$, since φ is strictly increasing and $D_{u(c)}(t)$ is strictly increasing in $u(c)$, $K_x = \varphi(D_{\alpha_x x}) > \varphi(D_{\beta x}) \geq 0$. Hence, $K_x > 0$.

For any $x \in X_+ \setminus \Delta_0$, define

$$\Lambda_x := \{D \in [0,1]^T \mid \varphi(D) \leq K_x\}.$$

From Lemma 10, for any $x \in X_\ell$ we have

$$U(x) = u(x_0) + \sum_{t \geq 1} D_x(t) u(x_t),$$

$$\text{s.t.} \quad D_x = \arg\max_D \{\sum D(t)u(x_t) - \varphi_t(D(t))\}.$$

There exists $x' \notin X_\ell \cup \Delta_0$ such that $x = \alpha x'$ for some $\alpha \in (0,1)$. Since φ is strictly increasing and $D_{\alpha x'}$ is increasing in α up to $\alpha_{x'} x'$, $\varphi(D_x) \leq \varphi(D_{\alpha_{x'} x'}) = K_x$, that is, we have $D_x \in \Lambda_x$. Thus, D_x is also the unique maximizer in the constrained problem:

$$D_x = \arg\max_{D \in \Lambda_x} \{\sum D(t)u(x_t) - \varphi_t(D(t))\},$$

thereby establishing the result for $x \in X_\ell$.

Next consider $x \notin X_\ell \cup \Delta_0$, and take $\alpha_x x \in X_\ell$. By definition, note that $K_x < \infty$. By the preceding,

$$D_{\alpha_x x} = \arg\max_{D \in \Lambda_x} \{\sum D(t)u(\alpha_x x_t) - \varphi_t(D(t))\}.$$

For notational simplicity, for any x, let $u(x)$ denote $(u(x_1), \cdots, u(x_T)) \in \mathbb{R}^T_+$. We first prove that

$$D_{\alpha_x x} \in \arg\max_{D \in \Lambda_x} D \cdot u(x). \tag{8}$$

To see this, suppose by way of contradiction that there is $D \in \Lambda_x$ s.t. $D \cdot u(x) > D_{\alpha_x x} \cdot u(x)$. Since $D_{\alpha_x x}$ is on the boundary of Λ_x and $D \in \Lambda_x$, we have $\varphi(D_{\alpha_x x}) = K_x \geq \varphi(D)$. But these inequalities imply that

$$D \cdot u(\alpha_x x) - \varphi(D) > D_{\alpha_x x} \cdot u(\alpha_x x) - \varphi(D_{\alpha_x x}),$$

contradicting the optimality of $D_{\alpha_x x}$ for $\alpha_x x$, as desired.

To conclude the proof of the lemma, observe that for any $D \in \Lambda_x$ with $D \neq D_{\alpha_x x}$,

$$
\begin{aligned}
& D_{\alpha_x x} \cdot u(\alpha_x x) - \varphi(D_{\alpha_x x}) > D \cdot u(\alpha_x x) - \varphi(D) \\
\Longrightarrow\ & D_{\alpha_x x} \cdot u(\alpha_x x) - D \cdot u(\alpha_x x) > \varphi(D_{\alpha_x x}) - \varphi(D) \\
\Longrightarrow\ & \alpha_x [D_{\alpha_x x} \cdot u(x) - D \cdot u(x)] > \varphi(D_{\alpha_x x}) - \varphi(D) \\
\Longrightarrow\ & D_{\alpha_x x} \cdot u(x) - D \cdot u(x) > \varphi(D_{\alpha_x x}) - \varphi(D) \\
& (\text{since } D_{\alpha_x x} \cdot u(x) \geq D \cdot u(x), \text{ by } (8)) \\
\Longrightarrow\ & D_{\alpha_x x} \cdot u(x) - \varphi(D_{\alpha_x x}) > D \cdot u(x) - \varphi(D).
\end{aligned}
$$

Thus,

$$D_{\alpha_x x} = \arg\max_{D \in \Lambda_x} \{\sum D(t) u(x_t) - \varphi_t(D(t))\},$$

as desired.

By Lemma 9, if $u(x_t) = u(y_t)$ for all $t \geq 1$, $\alpha_x = \alpha_y$. Thus K_x is finite if and only if K_y is finite. If K_x is finite, it is obvious from the definition that K_x depends only on the utility stream $(u(x_t))_{t=1}^T$. Thus, we have $K_x = K_y$. ■

All that remains to be established is to show properties of K: For all $S \subset \{1, \cdots, T\}$, let

$$\varphi_S(D) := \sum_{t \in S} \varphi_t(D(t)).$$

Lemma 12 *(1) $K : X_+ \setminus \Delta_0 \to \mathbb{R}_{++} \cup \{\infty\}$ is continuous.*

(2) For all p and $t \geq 1$, $K_{p^t} = \varphi_t(\bar{d}_t)$.

(3) If $K_x < \infty$, $K_x \leq K_{xS0}$ for all $S \subset \{1, \cdots, T\}$.

Proof. (1) Take any $x \in X_+ \setminus \Delta_0$. First assume $K_x < \infty$. Thus, there exists some λ with $\lambda x \notin X_\ell$. By X_ℓ-Continuity, any consumption stream y in a small neighborhood of x also satisfies $\lambda y \notin X_\ell$, which implies $K_y < \infty$. In this case, by definition, $K_y = \varphi(D_{\alpha_y y})$ for all such y. Moreover, $D_{\alpha_x x}$ is a unique maximizer of

$$\max\{\sum D(t) u(\alpha_x \circ x_t) - \varphi_t(D(t))\}.$$

If α_x is continuous in x, then $D \cdot u(\alpha_x x)$ is continuous and hence the maximum theorem implies that $D_{\alpha_x x}$ is continuous. Since φ is differentiable (and hence continuous), we have the desired result.

From now on, we will claim that α_x is continuous in x.

Claim 2 *α_x is lower semi-continuous in x, that is, if $x^n \to x$, then*

$$\liminf_n \alpha_{x^n} \geq \alpha_x.$$

Proof. Seeking a contradiction, suppose

$$\alpha_x > \alpha^* := \liminf_n \alpha_{x^n}.$$

Take any $\alpha \in (\alpha^*, \alpha_x)$. There exists a subsequence α_{x^m} converging to α^*. Since $\alpha_{x^m} \to \alpha^*$, $\alpha_{x^m} < \alpha$ for all sufficiently large m. By Lemma 8, $\alpha x^m \sim \alpha \circ c_{x^m}$. By Continuity and Claim 1, $\alpha x \sim \alpha \circ c_x$. On the other hand, since $\alpha < \alpha_x$, Lemma 8 implies $\alpha \circ c_x \succ \alpha x$, which is a contradiction. ■

Claim 3 *α_x is upper semi-continuous in x, that is, if $x^n \to x$, then*

$$\limsup_n \alpha_{x^n} \leq \alpha_x.$$

Proof. Seeking a contradiction, suppose

$$\alpha_x < \alpha^* := \limsup_n \alpha_{x^n}.$$

Take any $\alpha \in (\alpha_x, \alpha^*)$. There exists a subsequence α_{x^m} converging to α^*. Since $\alpha_{x^m} \to \alpha^*$, $\alpha < \alpha_{x^m}$ for all sufficiently large m. By Lemma 7, $\alpha x^m \in X_\ell$. Since X_ℓ is closed in $X_+ \setminus \Delta_0$ by X_ℓ-Continuity, $\alpha x \in X_\ell$. On the other hand, since $\alpha_x < \alpha$, Lemma 7 implies $\alpha x \notin X_\ell$, which is a contradiction. ■

Next consider the case of $K_x = \infty$. We want to show that K_{x^n} diverges to infinity as $x^n \to x$. Without loss of generality, assume $K_{x^n} < \infty$ for all n. Seeking a contradiction, suppose that there exists some subsequence x^m such that $K_{x^m} \leq \overline{K}$ for some $\overline{K} < \infty$. There exists y^m on the boundary of X_ℓ corresponding to each x^m. By definition, $K_{x^m} = \varphi(D_{y^m})$. Since $K_x = \infty$, all y on the same ray passing through x belong to X_ℓ. By X_ℓ-Dominance, $p^t \in X_\ell$ for all p. Thus, each φ_t is unbounded above because $\varphi_t'(D_{u(p)}(t)) = u(p)$ for all $u(p)$. Therefore, together with $\overline{K} \geq K_{x^m} = \varphi(D_{y^m})$, the sequence $\{y^m\}_{m=1}^\infty$ must be bounded. We can find a consumption stream $z^m := \lambda^m y^m \notin X_\ell$ with λ^m sufficiently larger than one. In particular, $\tilde{\alpha} z^m \notin X_\ell$ for some $\tilde{\alpha} \in (0,1)$ sufficiently close to one. Moreover, since x^m and z^m are on the same ray, z^m can be taken to converge to some point $z := \lambda x$.

By Lemma 8, together with the above observations, $\tilde{\alpha} \circ c_{z^m} \sim \tilde{\alpha} z^m$. By continuity of preference and continuity of present equivalents (Claim 1), $\tilde{\alpha} \circ c_z \sim \tilde{\alpha} z$. On the other hand, $K_x = \infty$ implies that $z \in X_\ell$, and hence, $\tilde{\alpha} \circ c_z \succ \tilde{\alpha} z$, which is a contradiction. This completes the proof.

(2) Since φ_t is defined by using $p^t \in X_\ell$, by construction, we have $K_{p^t} = \varphi_t(\overline{d}_t)$.

(3) Step 1: $K_x \leq K_{p^t}$. By part (2), it must be that

$$\{D(t) \in [0,1] \mid \varphi_t(D(t)) \leq K_{p^t}\} = \mathrm{eff}(\varphi_t),$$

and in turn,

$$\{D \in [0,1]^T \mid \varphi_t(D(t)) \leq K_{p^t} \text{ for all } t\} = \mathrm{eff}(\varphi).$$

For any stream x, trivially we must have $\{D \mid \varphi(D) \le K_x\} \subset \mathrm{eff}(\varphi)$, and so

$$\{D \in [0,1]^T \mid \varphi(D) \le K_x\} \subset \{D \in [0,1]^T \mid \varphi_t(D(t)) \le K_{p^t} \text{ for all } t\}.$$

To show that $K_x \le K_{p^t}$ for all t, take any t and any D in $\{D \mid \varphi(D) \le K_x\}$ that satisfies $D(t') = 0$ for $t' \neq t$. Then the above condition implies

$$\varphi_t(D(t)) \le K_x \Longrightarrow \varphi_t(D(t)) \le K_{p^t}.$$

In particular, if $D(t)$ satisfies $\varphi_t(D(t)) = K_x$,

$$K_x = \varphi_t(D(t)) \le K_{p^t},$$

as desired.

Step 2: For all $x \in X_+ \setminus \Delta_0$, $d = (d_t)_{t=1}^T$, and $S \subset \{1, \cdots, T\}$,

$$\varphi(d) \le K_x \implies \varphi_S(d) \le K_{xS0}.$$

Take any x and d with $\varphi(d) \le K_x$. By the properties of K shown by Lemma 11, for any $\lambda > 0$, if $y := \lambda x$, then $K_x = K_{\lambda x} = K_y$. By definition of K_x, there exists $\lambda > 0$ such that $y = \lambda x$ belongs to the boundary of X_ℓ and $K_y = \varphi(D_y)$. Since X_ℓ-Dominance implies $yS0 \in X_\ell$ for all $S \subset \{1, \cdots, T\}$ such that $y_t \succ 0$ for some $t \in S$, we have $\varphi_S(D_{yS0}) \le K_{yS0}$. Moreover, the value $D_y(t)$ is also optimal for $yS0$, that is, $D_y(t) = D_{yS0}(t)$ for all $t \in S$. Now, for any d with $\varphi(d) \le K_x$, since $K_x = K_y = \varphi(D_y)$, we have $\varphi_S(d) \le \varphi_S(D_y)$. Therefore,

$$\varphi_S(d) \le \varphi_S(D_y) = \varphi_S(D_{yS0}) \le K_{yS0} = K_{(\lambda x)S0} = K_{\lambda(xS0)} = K_{xS0}.$$

Step 3: The result. Take any $d_S \in [0,1]^T$ such that $d_S(t) \ge 0$ for all $t \in S$ and $d_S(t) = 0$ otherwise. Assume also $\varphi(d_S) \le K_x$. By part (2) and Step 1, $\varphi(d_S) \le K_x \le K_{p^t} = \varphi_t(\overline{d}_t)$. Hence, there exists some d_S^* such that $\varphi(d_S^*) = K_x$. It follows from Step 2 that $K_x = \varphi(d_S^*) = \varphi_S(d_S^*) \le K_{xS0}$, as desired. ■

B.4 Extending the Representation to X

Finally, we extend the representation on $X_\ell^* \cup \Delta_0$ to all of X.

Lemma 13 *For any stream $x \in X \setminus \Delta_0$ and any its absolute value x^*, there exists a unique $\alpha_{x^*} \in (0,1]$ such that*

$$\begin{cases} \alpha \le \alpha_{x^*} \implies \alpha x \in X_\ell^*, \\ \alpha > \alpha_{x^*} \implies \alpha x \notin X_\ell^*. \end{cases}$$

Proof. Since x^* is a positive stream, there exists $\alpha_{x^*} \in (0,1]$, ensured by Lemma 7, such that $\alpha x^* \in X_\ell$ if $\alpha \le \alpha_{x^*}$ and $\alpha x^* \notin X_\ell$ if $\alpha > \alpha_{x^*}$. By definition of X_ℓ^*, $\alpha x \in X_\ell^*$ if and only if $\alpha x^* \in X_\ell$, as desired. ■

Lemma 14 *Take any $x \in X \setminus \Delta_0$ and any its absolute values x^* and x^{**}. Then, (1) $\alpha_{x^*} = \alpha_{x^{**}}$, where α_{x^*} and $\alpha_{x^{**}}$ are defined as in Lemma 13.*
(2) If a positive stream y satisfies $x^*_t \sim y_t$ for all t, $\alpha_{x^*} = \alpha_y$.

Proof. (1) By definition of absolute values, $x^*_t \sim x^{**}_t$ for all t. By Lemma 9, $\alpha_{x^*} = \alpha_{x^{**}}$.
(2) Immediate from Lemma 9. ■

For all $x \in X \setminus \Delta_0$, by Lemma 13, define

$$\alpha_x := \alpha_{x^*}.$$

By Lemma 14, α_x is well-defined.

Lemma 15 *$\alpha_x \circ c_x \sim \alpha_x x$ for all $x \in X \setminus \Delta_0$.*

Proof. If $\alpha_x = 1$, the claim holds by definition of c_x. So, assume $\alpha_x < 1$. If x is a positive stream, $x^* = x$. Thus, Lemma 8 implies $\alpha_x \circ c_x \sim \alpha_x x$, as desired. Next assume that x is a negative stream. Since x^* is a positive stream, Lemma 8 implies $\alpha_x \circ c_{x^*} \sim \alpha_x x^*$. By Symmetry, $\alpha_x \circ (c_x)^* \sim \alpha_x x^*$. By Lemma 2 (3), $(\alpha_x \circ c_x)^* \sim (\alpha_x x)^*$. Since $\alpha_x x \in X^*_\ell$, together with Lemma 2 (1),

$$\begin{aligned}
& u((\alpha_x \circ c_x)^*) = U((\alpha_x x)^*) \\
\Longrightarrow\ & -u(\alpha_x \circ c_x) = -u(\alpha_x \circ x_0) - \sum_{t\geq 1} D_{|u(\alpha_x \circ x_t)|}(t) u(\alpha_x \circ x_t) \\
\Longrightarrow\ & -u(\alpha_x \circ c_x) = -U(\alpha_x x).
\end{aligned}$$

Thus, $\alpha_x \circ c_x \sim \alpha_x x$, as desired. Finally assume that a stream x is neither positive nor negative. By Lemma 13, a sequence $\alpha^n \to \alpha_x$ with $\alpha^n > \alpha_x$ satisfies $\alpha^n x \notin X^*_\ell$. By Lemma 2 (3), $\alpha^n x^* \notin X_\ell$. Lemmas 7 and 8 imply $\alpha^n \circ c_{x^*} \sim \alpha^n x^*$. Thus, by Symmetric Homotheticity, $\alpha^n \circ c_x \sim \alpha^n x$. By Continuity, $\alpha_x \circ c_x \sim \alpha_x x$ as $n \to \infty$. ■

Lemma 16 *The function $U : X \to \mathbb{R}$ appeared in Lemma 1 can be written as*

$$U(x) = u(x_0) + \sum_{t\geq 1} D_x(t) u(x_t),$$

$$s.t. \quad D_x = \begin{cases} \arg\max_D \{\sum D(t)|u(x_t)| - \varphi_t(D(t))\} & \text{if } x \in X^*_\ell \cup \Delta_0, \\ D_{\alpha_x x} & \text{if } x \notin X^*_\ell \cup \Delta_0. \end{cases}$$

Proof. By Lemma 6, U has the desired form on $X^*_\ell \cup \Delta_0$. Consider the case of $x \notin X^*_\ell \cup \Delta_0$. By Lemma 13, $\alpha_x x \in X^*_\ell$. Together with Lemma 15, the result follows from the same argument as in Lemma 10. ■

Lemma 17 *There is a function $K : X \setminus \Delta_0 \to \mathbb{R}_{++} \cup \{\infty\}$ such that $\succsim$ is represented by*

$$U(x) = u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t),$$

$$s.t. \qquad D_x = \arg\max_{D\in\Lambda_x}\{\sum D(t)|u(x_t)| - \varphi_t(D(t))\}$$

$$\Lambda_x := \{D \in [0,1]^T \mid \varphi(D) \leq K_x\}.$$

Moreover, the function K_x satisfies $K_x = K_{\lambda x}$ for any x and λ.

Proof. By Lemma 11, there exists such a function K on $X_+ \setminus \Delta_0$. For any $x \in X \setminus \Delta_0$, since its absolute value x^* is a positive stream, define

$$K_x := K_{x^*} > 0$$

for some absolute value x^* of x. Note that for all absolute values x^* and x^{**} of x, $x^*_t \sim x^{**}_t$. By the property of K stated in Lemma 11, $K_{x^*} = K_{x^{**}}$. Thus, K_x is well-defined. Moreover, by definition, for all $\lambda > 0$, $K_{\lambda x} = K_{\lambda x^*} = K_{x^*} = K_x$.

Note that $\lambda x \in X^*_\ell$ for all $\lambda > 0$ if and only if $\lambda x^* \in X_\ell$ for all $\lambda > 0$. Thus, $K_x = K_{x^*} = \infty$. Otherwise, we can find another x on the same ray with $x \notin X^*_\ell$. For such an x,

$$K_x = K_{x^*} = \varphi(D_{\alpha_{x^*}x^*}).$$

Since $D_{\alpha_x x}$ depends only on an absolute value of $\alpha_x x$, $\varphi(D_{\alpha_x x}) = \varphi(D_{\alpha_{x^*}x^*})$. That is, $K_x = \varphi(D_{\alpha_x x})$.

For any $x \in X \setminus \Delta_0$, define

$$\Lambda_x := \{D \in [0,1]^T \mid \varphi(D) \leq K_x\}.$$

By replacing X_ℓ with X^*_ℓ, the subsequent argument is the same as in Lemma 11. ■

All that remains to be established is to show properties of K:

Lemma 18 *(1) For all x, y, if $|u(x_t)| = |u(y_t)|$ for all t, then $K_x = K_y$.*

(2) $K : X \setminus \Delta_0 \to \mathbb{R}_{++} \cup \{\infty\}$ is continuous.

(3) For all p and $t \geq 1$, $K_{p^t} = \varphi_t(\overline{d}_t)$.

(4) If $K_x < \infty$, $K_x \leq K_{xS0}$ for all $S \subset \{1, \cdots, T\}$.

Proof. (1) Consider such x and y. By assumption, $x^*_t \sim y^*_t$ for all t. By property (2) of K appeared in Lemma 11, $K_{x^*} = K_{y^*}$. Thus, by definition, $K_x = K_y$.

(2) By part (1), K_x can be written as $K_{|u(x)|}$, that is, K depends only on the absolute value of utility streams $(|u(x_t)|)_{t=1}^T$. Thus, K can be regarded as a composite function

between the transformation from x into $(|u(x_t)|)_{t=1}^T$ and the function $K : X_+ \setminus \Delta_0 \to \mathbb{R}_{++} \cup \{\infty\}$. The desired result follows from continuity of the latter shown by Lemma 12 (1).

(3) Since K depends only on the absolute value of utility stream, the desired result follows from Lemma 12 (2).

(4) Since K depends only on the absolute value of utility stream, the desired result follows from Lemma 12 (3). ■

B.5 Necessity

Given a General CCE representation, define the set of ℓ-magnitude sensitive streams $X_\ell \subset X$ by

$$X_\ell = \{x \in X_+ \mid \alpha U(x) > U(\alpha x) \text{ for all } \alpha \in (0,1)\}.$$

First of all, we show that X_ℓ is characterized by the FOC of the unconstrained optimization problem:

$$\max_{D \in \mathbb{R}_+^T} \{\sum_{t \geq 1} D(t)|u(x_t)| - \varphi_t(D(t))\}.$$

Let D_x^{un} denote an optimal discount function for the unconstrained optimization problem, which is characterized by the FOC, $|u(x_t)| = \varphi_t'(D_x^{un}(t))$ for all $t \geq 1$ with $|u(x_t)| > 0$, or equivalently,

$$D_x^{un}(t) := (\varphi_t')^{-1}(|u(x_t)|)$$

if $u(x_t) > 0$, and $D_x^{un}(t) = 0$ if $u(x_t) = 0$. Since φ_t' is strictly increasing, $D_x^{un}(t)$ is strictly increasing in $|u(x_t)|$.

Lemma 19

$$X_\ell = \{x \in X_+ \mid \varphi(D_x^{un}) \leq K_x\}.$$

Proof. To show X_ℓ belongs to the right-hand side, take any $x \in X_\ell$. By the representation,

$$u(c_x) = U(x) = u(x_0) + \sum_{t \geq 1} D_x(t) u(x_t),$$

where $D_x = \arg\max_{D \in \Lambda_x} \{\sum_{t \geq 1} D(t)u(x_t) - \varphi_t(D(t))\}$. By definition of X_ℓ, for all $\alpha \in (0,1)$, $u(\alpha \circ c_x) > U(\alpha x)$. Together with linearity of u, this implies

$$\sum_{t \geq 1} D_x(t) u(x_t) > \sum_{t \geq 1} D_{\alpha x}(t) u(x_t).$$

Since $u(x_t) \geq 0$ and $D_x \geq D_{\alpha x}$ by Proposition 1 (3), we have $D_x(t) > D_{\alpha x}(t)$ for some t. By definition of D_x, together with properties of the representation,

$$\varphi(D_{\alpha x}) < \varphi(D_x) \leq K_x = K_{\alpha x}.$$

Hence, $D_{\alpha x} = D^{un}_{\alpha x}$. As $\alpha \to 1$, we have $\varphi(D^{un}_x) \le K_x$, as desired.

Conversely, take any x from the right-hand side. For $\alpha \in (0,1)$, By property (c) of the representation,

$$\varphi(D^{un}_{\alpha x}) < \varphi(D^{un}_x) \le K_x = K_{\alpha x}.$$

Therefore,

$$D^{un}_{\alpha x} = D_{\alpha x} = \arg\max_{\Lambda_{\alpha x}} \{\sum_{t\ge 1} D(t)u(\alpha \circ x_t) - \varphi_t(D(t))\}.$$

Since $D_x = D^{un}_x > D^{un}_{\alpha x} = D_{\alpha x}$ and u is linear,

$$\begin{aligned} u(\alpha \circ c_x) &= u(\alpha \circ x_0) + \sum_{t\ge 1} D_x(t)u(\alpha \circ x_t) \\ &> u(\alpha \circ x_0) + \sum_{t\ge 1} D_{\alpha x}(t)u(\alpha \circ x_t) = U(\alpha x), \end{aligned}$$

that is, $\alpha \circ c_x \succ \alpha x$. Hence, $x \in X_\ell$. ■

For all streams x, note that x^* is an absolute value of x if and only if $u(x^*_t) = |u(x_t)|$ for all t. Given the representation, the set of ℓ-magnitude sensitive streams is defined as

$$X^*_\ell = \{x \in X \mid \alpha U(x^*) > U(\alpha x^*) \text{ for all } \alpha \in (0,1)\}.$$

By Lemma 19,

$$X^*_\ell = \{x \in X \mid \varphi(D^{un}_x) \le K_x\}. \tag{9}$$

Note that $D_x(t)$ is continuous in x. By (9), $D_x(t)$ is strictly increasing in $|u(x_t)|$ on X^*_ℓ. It is obvious to see that $\succsim$ that U represents satisfies Weak Regularity. Lemma 19 implies Time-0 Irrelevance and X_ℓ-Continuity.

Lemma 20 *$\succsim$ satisfies Weak Homotheticity.*

Proof. Take any positive stream $x \in X_+$. By Proposition 1 (3), $D_x(t) \ge D_{\alpha x}(t)$, which implies, with linearity of u, $\alpha U(x) \ge U(\alpha x)$, or $\alpha \circ c_x \succsim \alpha x$, as desired. ■

Lemma 21 *$\succsim$ satisfies Strong X_ℓ-Regularity.*

Proof. Take any positive $x \notin X_+ \setminus \Delta_0$. Assume $x \notin X_\ell$. Then by Lemma 19, the unconstrained optimal discount function D^{un}_x violates the capacity constraint, that is, $\varphi(D^{un}_x) > K_x$. Since D^{un}_x is strictly increasing in $u(x_t)$, as $\alpha \to 0$, $D^{un}_{\alpha x}(t) \to \underline{d}_t$ of the minimum discount factor. Since $\varphi_t(\underline{d}_t) = 0$, there must exist $\alpha < 1$ for which $\varphi(D^{un}_{\alpha x}) < K_x$. By property (c), $\varphi(D^{un}_{\alpha x}) < K_{\alpha x}$, implying that $\alpha x \in X_\ell$ by Lemma 19.

Next, take any $x \in X_\ell$ and $\alpha \in (0,1)$. By Lemma 19 and property (c) of the representation, $\varphi(D^{un}_{\alpha x}) < \varphi(D^{un}_x) \le K_x = K_{\alpha x}$. Again by Lemma 19, $\alpha x \in X_\ell$, as desired. ■

Lemma 22 $\succsim$ *satisfies X_ℓ^*-Separability.*

Proof. From (9), $D_x = D_x^{un}$ on X_ℓ^*. Thus, $D_x(t)$ depends only on $|u(x_t)|$. Therefore, the representation on X_ℓ^* is additively separable and satisfies X_ℓ^*-Separability. ∎

Lemma 23 $\succsim$ *satisfies X_ℓ-Time-Invariance.*

Proof. Take any positive outcomes p and $\widehat{p}$. Suppose $p^t, \widehat{p}^t \in X_\ell$. By the representation on X_ℓ, $U(p^t) = D_{u(p)}(t)u(p)$ and $U(\widehat{p}^t) = D_{u(\widehat{p})}(t)u(\widehat{p})$. Since $D_r(t)$ is increasing in r, if $u(p) \geq u(\widehat{p})$, we have

$$U(p^t) = D_{u(p)}(t)u(p) \geq D_{u(\widehat{p})}(t)u(\widehat{p}) = U(\widehat{p}^t).$$

∎

Lemma 24 $\succsim$ *satisfies X_ℓ-Dominance.*

Proof. Take any $x \in X_\ell$ and consider an optimal D_x. By Lemma 19, $\varphi(D_x^{un}) \leq K_x$. Take any $S \subset \{1, \cdots, T\}$ with $x_t \succ 0$ for some $t \in S$. Note that $D_x^{un}(t)$ is also optimal for $xS0$, that is, $D_x^{un}(t) = D_{xS0}^{un}(t)$ for all $t \in S$. By property (c)(iv) of the General CCE representation,

$$\varphi_S(D_{xS0}^{un}) = \varphi_S(D_x^{un}) \leq \varphi(D_x^{un}) \leq K_x \leq K_{xS0}.$$

Thus, again by By Lemma 19, $xS0 \in X_\ell$, as desired. ∎

Lemma 25 $\succsim$ *satisfies Symmetry*

Proof. Take any negative stream x and its absolute value x^*. Note that the absolute value of the utility stream of x^* is written as $(|u(x_t)|)_{t=0}^T$, which is denoted by $|u(x)|$. By representation,

$$U(x^*) = |u(x_0)| + \sum_{t \geq 1} D_{|u(x)|}|u(x_t)|, \text{ and}$$
$$U(x) = u(x_0) + \sum_{t \geq 1} D_{|u(x)|}u(x_t).$$

Therefore,

$$u(c_{x^*}) = U(x^*) = -U(x) = -u(c_x) = u((c_x)^*),$$

as desired. ∎

Lemma 26 $\succsim$ *satisfies Symmetric Homotheticity.*

Proof. For all $x \notin \Delta_0$ and $\alpha \in (0,1)$, assume that $\alpha \circ c_{x^*} \sim \alpha x^*$. The representation implies that $\alpha U(x^*) = U(\alpha x^*)$, which in turn implies

$$\sum_{t \geq 1} (D_{x^*}(t) - D_{\alpha x^*}(t)) u(x_t^*) = 0.$$

Since $u(x_t^*) \geq 0$ and $D_{x^*}(t) \geq D_{\alpha x^*}(t)$ by Proposition 1 (3), we have $D_{x^*}(t) = D_{\alpha x^*}(t)$ for all t with $u(x_t^*) > 0$.

By representation, an optimal discount function satisfies $D_x = D_{x^*} = D_{|u(x)|}$. Thus, $D_x(t) = D_{\alpha x}(t)$ for all t with $u(x_t) \neq 0$. We have

$$\sum_{t \geq 1} (D_x(t) - D_{\alpha x}(t)) u(x_t) = 0,$$

which implies $\alpha U(x) = U(\alpha x)$, as desired. ■

B.6 Uniqueness

For any dated reward $x = p^t$ with $u(p) > 0$, the discount function (which requires $D_x(t) > 0$ and $D_x(\tau) = 0$ for $\tau \neq t$) is determined by preference: if $\gamma \in [0,1]$ is such that $\gamma \circ p \sim x$, then $D_x(t) = \gamma$. Thus the discount functions for dated rewards are uniquely pinned down by preference. Moreover, the set $\{D_{p^t}(t) \in [0,1] : p \succsim 0\}$ defines the effective domain of the cost function φ_t in any representation. We make use of these observations below.

Take two General CCE representations for the preference. Since u^1 and u^2 are linear and represent the same preference over lotteries, there exists $\alpha > 0$ such that $u^2 = \alpha u^1$. (Note that we impose a normalization $u_1(0) = u_2(0) = 0$.) Take a positive dated reward $x = p^t$. By the FOC, together with the above observation,

$$(\varphi_t^2)'(D_x(t)) = |u^2(p)| = \alpha |u^1(p)| = \alpha (\varphi_t^1)'(D_x(t)),$$

which implies $\varphi_t^2 = \alpha \varphi_t^1$. By property (c)(iii) of the representation, this holds on the effective domain.

Note that the unconstrained optimal discount function is identical between the two representations. Indeed, from the above observation, $|u^2(x_t)| = (\varphi_t^2)'(D_x^{un,2}(t))$ if and only if $\alpha |u^1(x_t)| = \alpha (\varphi_t^1)'(D_x^{un,2}(t))$, which is equivalent to $|u^1(x_t)| = (\varphi_t^1)'(D_x^{un,2}(t))$. Thus, we have $D_x^{un,1}(t) = D_x^{un,2}$.

By Lemma 19,

$$\{x \in X \mid \varphi^1(D_x^{un}) \leq K_x^1\} = X_\ell = \{x \in X \mid \varphi^2(D_x^{un}) \leq K_x^2\}.$$

Since $\varphi_t^2 = \alpha \varphi_t^1$,

$$\{x \in X \mid \varphi^2(D_x^{un}) \leq K_x^2\} = \{x \in X \mid \varphi^1(D_x^{un}) \leq \frac{K_x^2}{\alpha}\}.$$

Therefore, we must have $K_x^2 = \alpha K_x^1$.

C Appendix: Proof of Proposition 4

Take any $x \notin X_\ell$. Define

$$\Lambda_x = \{D \in [0,1]^T \,:\, \varphi(D) \leq K_x\}.$$

Since the empathy constraint must be binding for large streams, $\varphi(D_x) = K_x$ holds. Thus,

$$\begin{aligned} D_x &= \arg\max_{D\in\Lambda_x}\{\sum_{t\geq 1} D(t)u(x_t) - \sum_{t\geq 1}\varphi_t(D(t))\} \\ &= \arg\max_{D\in\Lambda_x}\{\sum_{t\geq 1} D(t)u(x_t) - K_x\} = \arg\max_{D\in\Lambda_x}\{\sum_{t\geq 1} D(t)u(x_t)\}. \end{aligned}$$

Therefore,

$$\begin{aligned} U(x) &= u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t) \\ &= u(x_0) + \max_{D\in\Lambda_x}\sum_{t\geq 1} D(t)u(x_t) = \max_{D\in\Lambda_x} D\cdot u(x) \end{aligned}$$

with identifying Λ_x by $\{D \in [0,1]^{T+1} \,:\, \sum_{t\geq 0}\varphi_t(D(t)) \leq K_x\}$ with $\varphi_0(d) = 0$ for all $d \in [0,1]$.

Together with this observation, the assumption implies that

$$\begin{aligned} U(\alpha x + (1-\alpha)y) &= \max_{D\in\Lambda_{\alpha x+(1-\alpha)y}} D\cdot u(\alpha x + (1-\alpha)y) \\ &\leq \alpha \max_{D\in\Lambda_{\alpha x+(1-\alpha)y}} D\cdot u(x) + (1-\alpha)\max_{D\in\Lambda_{\alpha x+(1-\alpha)y}} D\cdot u(y) \\ &\leq \alpha \max_{D\in\Lambda_x} D\cdot u(x) + (1-\alpha)\max_{D\in\Lambda_y} D\cdot u(y) \\ &= \alpha U(x) + (1-\alpha)U(y). \end{aligned}$$

D Appendix: Proof of Theorem 5

First, we show the sufficiency. Noor and Takeoka [3] show that the axioms on $\succsim$ implies Strong X_ℓ-Regularity, Time-0 Irrelevance, X_ℓ-Time-Invariance, X_ℓ-Dominance, and X_ℓ-Continuity. Thus, as shown in Appendix B.2, $\succsim$ admits the desired representation U on $X^*_\ell \cup \Delta_0$. Noor and Takeoka [3] show that U on $X_\ell \cup \Delta_0$ can be extended to X_+ as desired. In particular, φ_t admits a CRRA form. As shown in Appendix B.4, U is extended to X by using (φ, K) constructed for the positive representation. The function K is extended to $X \setminus \Delta_0$ by $K_x = K_{x^*}$.

Turn to the necessity. As shown in Noor and Takeoka [3], the axioms on X_+ is implied by the CCE representation. The CCE representation is a special case of the General CCE representation, which implies X_ℓ^*-Separability, Symmetry and Symmetric Homotheticity, as shown in Appendix B.5.

A non-trivial part is to show Monotonicity. For each components $(u, \{\varphi_t\}, K)$, the reduced form of a CCE representation is obtained as

$$U(x) = u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t),$$

where $\gamma(t) := (ma_t)^{-\frac{1}{m-1}}$, and

$$D_x(t) = \gamma(t)|u(x_t)|^{\frac{1}{m-1}} \tag{10}$$

if $U(x^*) - |u(x_0)| \leq mK$, and

$$D_x(t) = \frac{(mK)^{\frac{1}{m}}\gamma(t)|u(x_t)|^{\frac{1}{m-1}}}{\left\{\sum_{\tau=1}^{T}\gamma(\tau)|u(x_\tau)|^{\frac{m}{m-1}}\right\}^{\frac{1}{m}}} \tag{11}$$

if $U(x^*) - |u(x_0)| > mK$.

Lemma 27 $\succsim$ *satisfies Monotonicity.*

Proof. Take any x, y such that $u(x_t) \geq u(y_t)$ for all $t \geq 0$. Since $U(x)$ is additively separable between x_0 and everything else, it is enough to show Monotonicity for streams x, y with $u(x_0) = u(y_0) = 0$. From now on, we consider such streams only.

Step 1:

$$\sum \gamma(t)|u(x_t)|^{\frac{1}{m-1}}u(x_t) \geq \sum \gamma(t)|u(y_t)|^{\frac{1}{m-1}}u(y_t).$$

For each t, there are three cases: (1) $u(x_t) \geq u(y_t) \geq 0$, (2) $u(x_t) \geq 0 \geq u(y_t)$, and (3) $0 \geq u(x_t) \geq u(y_t)$. In any case, $|u(x_t)|^{\frac{1}{m-1}}u(x_t) \geq |u(y_t)|^{\frac{1}{m-1}}u(y_t)$. By adding up these inequalities across t, we have the desired result.

By Step 1, if x and y are small, we have the desired result. From now on, suppose that either x or y is large.

Step 2: If

$$\sum \gamma(t)|u(x_t)|^{\frac{1}{m-1}}u(x_t) \geq 0 \geq \sum \gamma(t)|u(y_t)|^{\frac{1}{m-1}}u(y_t), \tag{12}$$

then $U(x) \geq U(y)$. Suppose that x is large. Then, $U(x)$ is obtained by multiplying the expression in (12) by a positive multiplier, that is,

$$U(x) = \left(\frac{mK}{\sum \gamma(t)|u(x_t)|^{\frac{m}{m-1}}}\right)^{\frac{1}{m}} \sum \gamma(t)|u(x_t)|^{\frac{1}{m-1}}u(x_t).$$

Since this operation does not change the sign, we have the desired result. The same argument is applicable also when y is large.

Now assume that

$$\sum \gamma(t)|u(x_t)|^{\frac{1}{m-1}}u(x_t) \geq \sum \gamma(t)|u(y_t)|^{\frac{1}{m-1}}u(y_t) \geq 0. \tag{13}$$

Step 3: Let $S = \{t \geq 1 \,|\, u(y_t) \leq 0\}$. For any $z \in X$ satisfying $0 \geq u(z_t) \geq u(y_t)$ for $t \in S$ and $u(z_t) = u(y_t)$ elsewhere, $U(z) \geq U(y)$.

By assumption, for all $t \in S$, $0 \geq u(z_t) \geq u(y_t)$, which implies $|u(z_t)| \leq |u(y_t)|$. Thus, $\sum \gamma(t)|u(z_t)|^{\frac{m}{m-1}} \leq \sum \gamma(t)|u(y_t)|^{\frac{m}{m-1}}$. If y is small, z must be small as well. Hence, assume that y is large.

Suppose that z is large. By representation,

$$U(z) = \left(\frac{mK}{\sum \gamma(t)|u(z_t)|^{\frac{m}{m-1}}}\right)^{\frac{1}{m}} \sum \gamma(t)|u(z_t)|^{\frac{1}{m-1}}u(z_t).$$

If a consumption stream changes from z to y, the numerator is decreasing because (13) holds for z and y, while the denominator is increasing by assumption. Thus, $U(z) \geq U(y)$.

Next suppose that z is small. Since y is large, by Proposition **??**, $mK \leq \sum \gamma(t)|u(y_t)|^{\frac{m}{m-1}}$. Since (13) holds for z and y,

$$\begin{aligned} U(z) &= \sum \gamma(t)|u(z_t)|^{\frac{1}{m-1}}u(z_t) \\ &\geq \left(\frac{mK}{\sum \gamma(t)|u(y_t)|^{\frac{m}{m-1}}}\right)^{\frac{1}{m}} \sum \gamma(t)|u(y_t)|^{\frac{1}{m-1}}u(y_t) = U(y), \end{aligned}$$

as desired.

Now turn to the comparison between x and y. Define $S_N = \{t \geq 1 \,|\, u(x_t) < 0\}$ and $S_P = \{t \geq 1 \,|\, u(y_t) > 0\}$. Let y^d be the stream such that $y^d_t = x_t$ on $t \in S_N$, $y^d_t = y_t$ on $t \in S_P$, and $y^d_t = 0$ otherwise. Note that $u(y^d_t) \geq u(y_t)$ for all t. Moreover, $0 \geq u(y^d_t) \geq u(y_t)$ for $t \notin S_P$ and $u(y^d_t) = u(y_t)$ elsewhere. By Step 3, $U(y^d) \geq U(y)$.

It is enough to show that $U(x) \geq U(y^d)$. Note that $u(x_t) \geq u(y^d_t) \geq 0$ for all $t \notin S_N$ and $u(x_t) = u(y^d_t)$ elsewhere. From the representation,

$$U(x) = \sum_{t \in S_N} D_x(t)u(x_t) + \sum_{t \notin S_N} D_x(t)u(x_t)$$

and

$$U(y^d) = \sum_{t \in S_N} D_{y^d}(t)u(y^d_t) + \sum_{t \notin S_N} D_{y^d}(t)u(y^d_t).$$

Note that for all $t \in S_N$, $u(x_t) = u(y^d_t) < 0$. We show $U(x) \geq U(y^d)$ by the following two steps.

Step 4: For all $t \in S_N$, $D_x(t) \leq D_{y^d}(t)$.

Since $\sum \gamma(t)|u(x_t)|^{\frac{m}{m-1}} \geq \sum \gamma(t)|u(y^d_t)|^{\frac{m}{m-1}}$, we have either (a) x and y^d are large, or (b) x is large and y^d is small. If (a) holds, for all $t \in S_N$,

$$\begin{aligned}
D_{y^d}(t) &= \left(\frac{mK}{\sum_{\tau=1}^T \gamma(\tau)|u(y^d_\tau)|^{\frac{m}{m-1}}} \right)^{\frac{1}{m}} \gamma(t)|u(y^d_t)|^{\frac{1}{m-1}} \\
&\geq \left(\frac{mK}{\sum_{\tau=1}^T \gamma(\tau)|u(x_\tau)|^{\frac{m}{m-1}}} \right)^{\frac{1}{m}} \gamma(t)|u(x_t)|^{\frac{1}{m-1}} = D_x(t).
\end{aligned}$$

If (b) holds, $D_{y^d}(t) = \gamma(t)|u(y^d_t)|^{\frac{1}{m-1}}$. Since x is large, the multiplier for $D_x(t)$ is smaller than one. Thus we have the desired result.

Step 5: $\sum_{t \notin S_N} D_x(t)u(x_t) \geq \sum_{t \notin S_N} D_{y^d}(t)u(y^d_t)$.

Since $\sum \gamma(t)|u(x_t)|^{\frac{m}{m-1}} \geq \sum \gamma(t)|u(y^d_t)|^{\frac{m}{m-1}}$, we have either (a) x and y^d are large, or (b) x is large and y^d is small. If we define a function $f : \mathbb{R}_{++} \to \mathbb{R}_{++}$ by

$$f(\theta) = \left(\frac{a}{b+\theta} \right)^{\frac{1}{m}} \theta$$

for some $a, b > 0$, it is easy to verify that $f'(\theta) > 0$, that is, f is a strictly increasing function because $m > 1$. Given this observation, if (a) holds,

$$\begin{aligned}
&\sum_{t \notin S_N} D_x(t)u(x_t) \\
&= \left(\frac{mK}{\sum_{\tau \in S_N} |u(x_\tau)|^{\frac{m}{m-1}} + \sum_{\tau \notin S_N} u(x_\tau)^{\frac{m}{m-1}}} \right)^{\frac{1}{m}} \sum_{t \notin S_N} \gamma(t)u(x_t)^{\frac{m}{m-1}} \\
&\geq \left(\frac{mK}{\sum_{\tau \in S_N} |u(y^d_\tau)|^{\frac{m}{m-1}} + \sum_{\tau \notin S_N} u(y^d_\tau)^{\frac{m}{m-1}}} \right)^{\frac{1}{m}} \sum_{t \notin S_N} \gamma(t)u(y^d_t)^{\frac{m}{m-1}} \\
&= \sum_{t \notin S_N} D_{y^d}(t)u(y^d_t),
\end{aligned}$$

as desired.

Suppose that (b) holds. Define $x(\alpha) = \alpha x + (1-\alpha)y^d$ for all $\alpha \in (0, 1)$. Note that $u(x_t) = u(x_t(\alpha)) = u(y^d_t)$ for all $t \in S_N$ and $u(x_t) \geq u(x_t(\alpha)) \geq u(y^d_t) \geq 0$ otherwise. By continuity of the representation, there exists $\alpha^* \in (0, 1)$ such that $x(\alpha)$ is large if $\alpha > \alpha^*$

and $x(\alpha)$ is small if $\alpha \leq \alpha^*$. By the same argument as above,

$$\begin{aligned}
&\sum_{t \notin S_N} D_x(t) u(x_t) \\
&\geq \left(\frac{mK}{\sum_{\tau \in S_N} |u(x_\tau(\alpha^*))|^{\frac{m}{m-1}} + \sum_{\tau \notin S_N} u(x_\tau(\alpha^*))^{\frac{m}{m-1}}} \right)^{\frac{1}{m}} \sum_{t \notin S_N} \gamma(t) u(x_t(\alpha^*))^{\frac{m}{m-1}} \\
&= \left(\frac{mK}{mK} \right)^{\frac{1}{m}} \sum_{t \notin S_N} \gamma(t) u(x_t(\alpha^*))^{\frac{m}{m-1}} = \sum_{t \notin S_N} \gamma(t) u(x_t(\alpha^*))^{\frac{m}{m-1}} \\
&\geq \sum_{t \notin S_N} \gamma(t) u(y_t^d)^{\frac{m}{m-1}} = \sum_{t \notin S_N} D_{y^d}(t) u(y_t^d).
\end{aligned}$$

Consider the case

$$0 \geq \sum \gamma(t) |u(x_t)|^{\frac{1}{m-1}} u(x_t) \geq \sum \gamma(t) |u(y_t)|^{\frac{1}{m-1}} u(y_t). \tag{14}$$

Define $\bar{x} \in X$ by $u(\bar{x}_t) = -u(x_t)$ for all t. That is, $\bar{x}$ is the stream that has an opposite sign of utilities $u(x_t)$. Similarly, $\bar{y}$ is defined. Note that $|u(\bar{x}_t)| = |u(x_t)|$ and $|u(\bar{y}_t)| = |u(y_t)|$ for all t. Moreover, $u(\bar{y}_t) \geq u(\bar{x}_t)$ for all t because $u(y_t) \leq u(x_t)$ by assumption. From (14),

$$\sum \gamma(t) |u(\bar{y}_t)|^{\frac{1}{m-1}} u(\bar{y}_t) \geq \sum \gamma(t) |u(\bar{x}_t)|^{\frac{1}{m-1}} u(\bar{x}_t) \geq 0.$$

By the same argument as above,

$$\begin{aligned}
& U(\bar{y}) \geq U(\bar{x}) \\
\iff & \sum_{t \geq 1} D_{|u(\bar{y})|}(t) u(\bar{y}_t) \geq \sum_{t \geq 1} D_{|u(\bar{x})|}(t) u(\bar{x}_t) \\
\iff & \sum_{t \geq 1} D_{|u(x)|}(t) u(x_t) \geq \sum_{t \geq 1} D_{|u(y)|}(t) u(y_t) \\
\iff & U(x) \geq U(y).
\end{aligned}$$

■

References

[1] Maccheroni, F., M. Marinacci and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences", *Econometrica* 74, pp. 1447–1498.

[2] Noor, J., and N. Takeoka (2020): "Optimal Discounting", working paper.

[3] Noor, J., and N. Takeoka (2020): "Constrained Optimal Discounting", working paper.