Econometrica Supplementary Material

SUPPLEMENT TO "OPTIMAL DISCOUNTING"

(*Econometrica*, Vol. 90, No. 2, March 2022, 585–623)

JAWWAD NOOR
Department of Economics, Boston University

NORIO TAKEOKA
Department of Economics, Hitotsubashi University

This Supplementary Appendix for Noor and Takeoka (2022) provides (i) a characterization for the General Discounted Utility (GDU) model both on the deterministic streams and on the streams of lotteries, (ii) a weaker definition of present equivalents in the off-diagonal approach, (iii) omitted proofs for results in Noor and Takeoka (2022), and (iv) a new axiomatization result about a variant of homogeneous CE representation, called the smooth homogeneous CE model.

S1. PRELIMINARIES

IN THIS SUPPLEMENTARY APPENDIX for Noor and Takeoka (2022) (henceforth NT), we provide the omitted proofs for results in NT (NT, Sections 4, 5, and 6) as well as a characterization of the GDU model (NT, Section 2) and a characterization of the smooth homogeneous CE representation (NT, Section 7). We also show how the definition of present equivalents in the off-diagonal approach can be relaxed to allow for positive future consumption (NT, Section 3).

Recall that there are $T+1<\infty$ periods, starting with period 0. The space of outcomes is assumed to be $C=\mathbb{R}_+$. Let Δ denote the set of simple lotteries over C, with generic elements $p, q, \ldots$. Let Z denote either C or Δ. The set of consumption streams is defined as $X=Z^{T+1}$. A typical element in X is denoted by $x=(x_0, x_1, \ldots, x_T)$. The primitive of our model is a preference $\succsim$ over X.

Let $Z_0 \subset X$ denote the set of streams $x=(z, 0, \ldots, 0)$ that offer consumption $z \in Z$ immediately and 0 in every subsequent period. Abusing notation, we often use z to denote both a consumption $z \in Z$ and a stream $(z, 0, \ldots, 0) \in Z_0$. Thus, 0 also denotes the stream $(0, \ldots, 0)$. Denote by z^t the stream that pays consumption $z \in Z$ at time t and 0 in all other periods. Such a stream is called a *dated reward*.

Jawwad Noor: jnoor@bu.edu
Norio Takeoka: norio.takeoka@r.hit-u.ac.jp

The authors are very grateful three anonymous referees, and to Andrew Ellis, Larry Epstein, Shinsuke Ikeda, Bart Lipman, Jay Lu, Samuel McClure, Marco Mariotti, Anna Rubinchik, Bruno Strulovici, and Kan Takeuchi for very useful discussions, and to seminar participants at Boston University, Yokohama National University, HEC Paris, Queen Mary University of London, Caltech, UC San Diego, Otaru University of Commerce, Brown University, University of Almeria, Hitotsubashi University, Waseda University, Academia Sinica and the audience at RUD (University of Paris, Dauphine), Bounded Rationality in Choice (Queen Mary), SAET (IMPA, Rio de Janeiro), SAET (Faro, Portugal) and AMES (Doshisha University, Kyoto) for helpful comments. Noor gratefully acknowledges financial support from the NSF (Grant SES-1627906) and thanks the Yale economics department for its hospitality where part of this research was completed. Takeoka gratefully acknowledges financial support from JSPS KAKENHI (Grants JP18K01503 and JP19KK0308). Last but certainly not least, we thank Kevin Cooke and Nick Saponara for expert research assistance. The usual disclaimer applies.

 https://doi.org/10.3982/ECTA16050

S2. GDU REPRESENTATIONS

S2.1. *Basic Axioms*

First, we consider axioms that are commonly imposed on both $X = C^{T+1}$ and $X = \Delta^{T+1}$. We collect several standard conditions under a single heading.

AXIOM S1—Regularity: (a) (*Order*). *$\succsim$ is complete and transitive.*
(b) (*Continuity*). *For all $x \in X$, $\{y \in X \mid y \succsim x\}$ and $\{y \in X \mid x \succsim y\}$ are closed.*
(c) (*Impatience*). *For any $z \in Z$ and $t < t'$, $z^t \succsim z^{t'}$.*
(d) (*C-Monotonicity*). *For all $c, c' \in C$, $c \geq c' \Longleftrightarrow c \succsim c'$.*
(e) (*Monotonicity*). *For any $x, y \in X$,*

$$(x_t, 0, \dots, 0) \succsim (y_t, 0, \dots, 0) \quad \textit{for all } t \Longrightarrow x \succsim y.$$

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

Order and Continuity are standard. Impatience states that consumption is better when received sooner than later. C-Monotonicity states that more consumption is better than less. While C-Monotonicity applies only to immediate consumption, Monotonicity is a property on arbitrary streams: it requires that pointwise preferred streams are preferred.

S2.2. *GDU Representations Over Deterministic Streams*

In this subsection, we consider the case of $Z = C$, that is, preferences over $X = C^{T+1}$.

For notational convenience, for all streams $x, y \in X$ and all $S \subset \{0, 1, \dots, T\}$, let xSy denote the stream that pays according to x on S and according to y otherwise.

AXIOM S2—Strong Separability: *For all $S \subset \{0, 1, \dots, T\}$ and for all $x, x', y, y' \in X$,*

$$xSy \succsim x'Sy \quad \Longrightarrow \quad xSy' \succsim x'Sy'.$$

As an intermediate result and benchmark, we show the following.

THEOREM S1: *Assume $T \geq 3$. A preference $\succsim$ over $X = C^{T+1}$ satisfies Regularity and Strong Separability if and only if it admits a GDU representation.*

Moreover, if two GDU representations (u^i, D^i), $i = 1, 2$ represent the same preference $\succsim$, then there exists a scalar $\lambda > 0$ s.t. (i) $u^2 = \lambda u^1$, *and* (ii) *for all $c \in C$ and $t > 0$,*

$$D^1_{u^1(c)}(t) = D^2_{u^2(c)}(t).$$

PROOF: Take any $c, c' \in C$ with $c > c'$. By C-Monotonicity and Monotonicity, $c^t \succ (c')^t$ for all $t \geq 0$. Moreover, since $\succsim$ satisfies Regularity and Strong Separability, Debreu (1960, Theorem 3) ensures that $\succsim$ over X admits an additively separable utility representation

$$U(x) = \sum_{t \geq 0} U_t(x_t), \tag{S1}$$

where $U_t : C \to \mathbb{R}$ is continuous for any $t \geq 0$. Since $c^t \succ (c')^t$ for all $t \geq 0$ whenever $c > c'$, U_t is strictly increasing. Define $u(c) = U_0(c)$. Then U_t is written as $V_t(u)$ for some strictly increasing continuous function V_t. That is, $\succsim$ over X admits a utility representation of the form

$$U(x) = u(x_0) + \sum_{t\geq 1} V_t\big(u(x_t)\big). \tag{S2}$$

By normalization, we can assume $u(0) = 0$ and $V_t(0) = 0$.

Define D_x by $D_{u(x_t)}(t) = \frac{V_t(u(x_t))}{u(x_t)} > 0$ for any $x_t > 0$. Then

$$U(x) = u(x_0) + \sum_{t\geq 1} D_{u(x_t)}(t)u(x_t), \quad \text{for all } x \in X,$$

with convention that $D_{u(x_t)}(t)u(x_t) = 0$ whenever $u(x_t) = 0$.

Since u and V_t are continuous, so is $D_{u(c)}(t)$ on the domain of $c > 0$. By Impatience, for all positive c and $t \geq 1$, $u(c) = U(c^0) \geq U(c^t) = D_{u(c)}(t)u(c)$, which implies $D_{u(c)}(t) \leq 1$. Moreover, by Impatience, for all $t < T$, $D_{u(c)}(t)u(c) = U(c^t) \geq U(c^{t+1}) = D_{u(c)}(t+1)u(c)$. Thus, $D_{u(c)}(t)$ is weakly decreasing wrt t.

Finally, we establish uniqueness.

LEMMA S1: *If GDU* (u^i, D^i) *for* $i = 1, 2$ *both represent the same preference* $\succsim$, *then there exists a scalar* $\lambda > 0$ *s.t. for all* $c \in C$ *and* $t > 0$,

$$u^2(c) = \lambda u^1(c) \quad \textit{and} \quad D^1_{u^1(c)}(t) = D^2_{u^2(c)}(t).$$

PROOF: By the uniqueness property of the additive separable utility function (Debreu (1960)), we obtain $\lambda > 0$ and $\gamma \in \mathbb{R}$ s.t. $u^2 = \lambda u^1 + \gamma$. Due to the normalization $u(0) = 0$ in the representation, we have $\gamma = 0$. Thus, $u^2 = \lambda u^1$.

Next, take any $c > 0$. Since $D_{u^i(c)}(t) \in (0, 1]$, $u^i(c) \geq D_{u^i(c)}(t)u^i(c)$. By continuity of u^i, there exists $\overline{c}^i \in C$ such that $u^i(\overline{c}^i) = D_{u^i(c)}(t)u^i(c)$. From the representation, this equation means $U^i(\overline{c}^i) = U^i(c^t)$, or $\overline{c}^i \sim c^t$, $i = 1, 2$. That is, $\overline{c}^i$ is a present equivalent of c^t. Since a present equivalent is unique by C-Monotonicity, $\overline{c}^1 = \overline{c}^2 = \overline{c}$. Since $u^2 = \lambda u^1$, we therefore see that

$$D^1_{u^1(c)}(t) = \frac{u^1(\overline{c})}{u^1(c)} = \frac{u^2(\overline{c})}{u^2(c)} = D^2_{u^2(c)}(t),$$

as desired. The converse is readily established. *Q.E.D.*

S2.3. *GDU Representations Over Streams of Lotteries*

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 mixture of any pair of streams $x, y \in X$ is given by

$$\alpha \circ x + (1-\alpha) \circ y := \big(\alpha \circ x_0 + (1-\alpha) \circ y_0, \dots, \alpha \circ x_T + (1-\alpha) \circ y_T\big).$$

For any stream x, we refer to $c_x \in C$ as its *present equivalent* if it satisfies $c_x \sim x$. Present equivalents will be used instrumentally below.

AXIOM S3—Present Equivalents: *For any stream* x, *there exist* $c \in C$ *such that* $c \succsim x$.

Present Equivalents states that for any stream, there are immediate consumption levels that are better 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, $x \succ y$).

AXIOM S4—Risk Preference: *For any* $p, p', p'' \in \Delta$ *and* $\alpha \in (0, 1]$,

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

Risk Preference imposes vNM Independence only on immediate consumption.

For notational convenience, for all streams $x, y \in X$ and all t, let xty denote the stream that pays according to x at t and according to y otherwise.

AXIOM S5—Separability: *For all* $x \in X$ *and all* t,

$$\frac{1}{2} \circ c_{xt0} + \frac{1}{2} \circ c_{0tx} \sim \frac{1}{2} \circ c_x + \frac{1}{2} \circ c_0.$$

The axiom can be justified by a thought experiment appealing to standard axioms. Suppose there are only 3 periods (periods 0,1,2). Imagine that, in a hypothetical period -1, the agent has a vNM preference $\succsim^*$ over the set $\Delta(X)$ of lotteries over streams (of lotteries). Denote the lottery mixture of two streams $x, y \in X$ by $\beta \diamond x + (1-\beta) \diamond y$ for any $\beta \in [0, 1]$, and consider the following condition:

$$\frac{1}{2} \diamond (0, c_1, 0) + \frac{1}{2} \diamond (c_0, 0, c_2) \sim^* \frac{1}{2} \diamond (c_0, c_1, c_2) + \frac{1}{2} \diamond (0, 0, 0).$$

which involves a 50-50 chance at $(0, c_1, 0)$ and $(c_0, 0, c_2)$ versus a 50-50 chance at (c_0, c_1, c_2) and $(0, 0, 0)$. The indifference expresses a notion of separability: the agent only cares that in each period t, she can end up with either c_t or 0 with equal probabilities, and in particular, does not care about the correlation in consumption realizations across periods. If c_x is the present equivalent of stream x, then by vNM Independence on $\succsim^*$, the following indifference involving lotteries over present equivalents is implied:

$$\frac{1}{2} \diamond c_{(0,c_1,0)} + \frac{1}{2} \diamond c_{(c_0,0,c_2)} \sim^* \frac{1}{2} \diamond c_{(c_0,c_1,c_2)} + \frac{1}{2} \diamond c_{(0,0,0)}.$$

Separability obtains by assuming *indifference to the timing of resolution of risk*. That is, since the timing of resolution of risk is inconsequential for eventual consumption, consequentialism yields

$$\frac{1}{2} \diamond c_{(0,c',0)} + \frac{1}{2} \diamond c_{(c,0,c'')} \sim^* \frac{1}{2} \circ c_{(0,c',0)} + \frac{1}{2} \circ c_{(c,0,c'')},$$

$$\text{and} \quad \frac{1}{2} \diamond c_{(c,c',c'')} + \frac{1}{2} \diamond c_{(0,0,0)} \sim^* \frac{1}{2} \circ c_{(c,c',c'')} + \frac{1}{2} \circ c_{(0,0,0)}.$$

Note that the left-hand side of $\sim^*$ is a lottery over X, which resolves in period -1. On the other hand, the right-hand side of $\sim^*$ is an element of X. The risk resolves in period 0. Finally, by transitivity of $\succsim^*$, we obtain $\frac{1}{2} \circ c_{(0,c_1,0)} + \frac{1}{2} \circ c_{(c_0,0,c_2)} \sim^* \frac{1}{2} \circ c_{(c_0,c_1,c_2)} + \frac{1}{2} \circ c_{(0,0,0)}$, and presuming dynamic consistency between periods -1 and 0, we obtain the corresponding indifference in terms of $\succsim$. This illustrates the justification for Separability.

Note that Separability is a statement about the mixture of present equivalents of streams, rather than about the mixture of streams. These two kinds of mixtures are equivalent if the Independence axiom holds, but as is evident from the behavioral definition of MDI (Weak Homotheticity in the main paper), the Independence axiom is at odds with MDI and, therefore, we have to formulate Separability in a novel manner.

Separability is a potentially objectionable condition for our story: if an agent establishes empathy for self t, then it is conceivable that she may costlessly empathize with adjacent selves $t-1$ and $t+1$. This would violate Separability. However, we could alternatively view the duration of a period as sufficiently long that such intertemporal complementarities disappear.

S2.3.1. *Representation Result*

THEOREM S2: *A preference $\succsim$ over $X = \Delta^{T+1}$ satisfies Regularity, Present Equivalents, Risk Preference, and Separability if and only if it admits an unbounded GDU representation.*

Moreover, if two GDU representations (u^i, D^i), $i = 1, 2$ represent the same preference $\succsim$, then there exists a scalar $\lambda > 0$ s.t. (i) $u^2 = \lambda u^1$, and (ii) for all $p \in \Delta$ and $t > 0$,

$$D^1_{u^1(p)}(t) = D^2_{u^2(p)}(t).$$

PROOF: The necessity of the axioms is straightforward to establish. We establish its sufficiency in the following lemmas.

LEMMA S2: *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, unbounded above, and homogeneous (i.e., $u(\alpha \circ p) = \alpha u(p)$ for all $\alpha \geq 0$). 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 $p \in \Delta$.*

PROOF: By 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. For any $\alpha \in [0, 1]$, by mixture linearity of u, together with identifying $\alpha \circ p$ with $\alpha \circ p + (1 - \alpha) \circ 0$,

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

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

Homogeneity implies that $u(\Delta) = \mathbb{R}_+$, which in turn implies $u(C) = \mathbb{R}_+$.

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 $x^n \to \overline{x}$. There exists a corresponding present equivalent $p_{x^n} \sim x^n$. We want to show that $U(x^n) = u(p_{x^n})$ converges to $U(\overline{x}) = u(p_{\overline{x}})$. Fix $p^* \in \Delta$ with $p^* \succ 0$ arbitrarily. Since u is continuous and homogeneous, there exists a unique $\lambda(x^n) \geq 0$ such that $u(p_{x^n}) = \lambda(x^n)u(p^*) = u(\lambda(x^n) \circ p^*)$. For $\overline{\lambda} > \lambda(\overline{x})$, $\overline{x}$ belongs to the set $W = \{x \in X | \overline{\lambda} \circ p^* \succ x \succsim 0\}$. By Continuity, we can assume $x^n \in W$ for all n without loss of generality.

Since $U(x^n) = \lambda(x^n)u(p^*)$ and $U(\overline{x}) = \lambda(\overline{x})u(p^*)$, it is enough to show that $\lambda(x^n) \to \lambda(\overline{x})$. Seeking a contradiction, suppose otherwise. Then there exists a neighborhood of $\lambda(\overline{x})$, denoted by $B(\lambda(\overline{x}))$, such that $\lambda(x^m) \notin B(\lambda(\overline{x}))$ for infinitely many m. Let $\{x^m\}$ denote the corresponding subsequence of $\{x^n\}$. Since $x^n \to \overline{x}$, $\{x^m\}$ also converges to $\overline{x}$.

Since $\{\lambda(x^m)\}$ is a sequence in $[0, \overline{\lambda}]$, there exists a convergent subsequence $\{\lambda(x^\ell)\}$ with a limit $\tilde{\lambda} \neq \lambda(\overline{x})$. On the other hand, since $x^\ell \to \overline{x}$ and $x^\ell \sim \lambda(x^\ell) \circ p^*$, Continuity implies $\overline{x} \sim \tilde{\lambda} \circ p^*$. Since $\lambda(\overline{x})$ is unique, $\lambda(\overline{x}) = \tilde{\lambda}$, which is a contradiction. *Q.E.D.*

LEMMA S3: *U can be written in an additively separable utility form, that is, $U : X \to \mathbb{R}_+$ s.t. for all $x \in X$,*

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

where u is given as in Lemma S2 and $U_t : \Delta \to \mathbb{R}_+$ is continuous with $U_t(0) = 0$ for each t.

PROOF: Take any $x = (x_0, x_1, \ldots, x_T) \in X$ s.t. $x \succ 0$. By Monotonicity, there exists some $t > 0$ with $x_t \succ 0$. We start with the case where there are two $x_t, x_s \succ 0$. By notational convenience, denote such a stream by $(x_t, x_s, 0, \ldots, 0)$. By Separability,

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

Since u is mixture linear,

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

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

$$U(x_t, x_s, 0, \ldots, 0) = U_t(x_t) + U_s(x_s). \tag{S3}$$

In particular, if $t = 0$, $U_t(x_t) = u(x_t)$.

If a stream has three outcomes $x_t, x_s, x_r \succ 0$, denote it by $(x_t, x_s, x_r, 0, \ldots, 0)$. From the above argument, we have (S3). By Separability,

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

Since u is mixture linear,

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

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

$$U(x_t, x_s, x_r, 0, \ldots, 0) = U_r(x_r) + U(x_t, x_s, 0, \ldots, 0) = U_t(x_t) + U_s(x_s) + U_r(x_r).$$

By repeating the same argument finitely many times, we have

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

where $U_t(x_t)$ is defined as $U_t(x_t) = U(0, \ldots, 0, x_t, 0, \ldots, 0)$. By definition, $U_t(0) = 0$. Since U is continuous, U_t is also continuous. *Q.E.D.*

LEMMA S4: *The function* $U : X \to \mathbb{R}_+$ *defined as in Lemma S3 can be written as*

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

where for all $t > 0$ *and for all* $u(p) > 0$, $D_{u(p)}(t) \in (0, 1]$, *and* $D_{u(p)}(t)$ *is continuous in* $u(p) > 0$ *and is weakly decreasing in* t.

PROOF: Taking the additive representation from Lemma S3, by Monotonicity, 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 $u(x_t) > 0$. Then

$$U(x) = u(x_0) + \sum_{t \geq 1} D_{u(x_t)}(t) u(x_t), \quad \text{for all } x \in X,$$

with convention that $D_{u(x_t)}(t) u(x_t) = 0$ whenever $u(x_t) = 0$.

Since u and U_t are continuous, so is $D_{u(c)}(t)$ in $u(c)$ on the domain of $u(c) > 0$. By Impatience, for all positive c and $t \geq 1$, $u(c) = U(c^0) \geq U(c^t) = D_{u(c)}(t) u(c)$, which implies $D_{u(c)}(t) \leq 1$. Moreover, by Impatience, for all $t < T$, $D_{u(c)}(t) u(c) = U(c^t) \geq U(c^{t+1}) = D_{u(c)}(t+1) u(c)$. Thus, $D_{u(c)}(t)$ is weakly decreasing wrt t. *Q.E.D.*

Finally, we establish uniqueness.

LEMMA S5: *If GDU* (u^i, D^i) *for* $i = 1, 2$ *both represent the same preference* $\succsim$, *then there exists a scalar* $\lambda > 0$ *s.t. for all* $p \in \Delta$ *and* $t > 0$,

$$u^2(p) = \lambda u^1(p) \quad \textit{and} \quad D^1_{u^1(p)}(t) = D^2_{u^2(p)}(t).$$

PROOF: By considering the restriction $\succsim|_{\Delta_0}$ and applying the mixture space theorem, we obtain $\lambda > 0$ and γ s.t. $u^2 = \lambda u^1 + \gamma$. Due to the normalization $u(0) = 0$ in the representation, we have $\gamma = 0$. Thus, $u^2 = \lambda u^1$.

Next, observe that there exists a present equivalent c_{p^t} for each dated reward p^t, the representation implies that $u^i(c_{p^t}) = D^i_{u^i(p)}(t) u^i(p)$ for any $p \in \Delta$ and $t > 0$. Since $u^2 = \lambda u^1$, we therefore see that

$$D^1_{u^1(p)}(t) = \frac{u^1(c_{p^t})}{u^1(p)} = \frac{u^2(c_{p^t})}{u^2(p)} = D^2_{u^2(p)}(t),$$

as desired. The converse is readily established. *Q.E.D.*

S3. OFF-DIAGONAL APPROACH: WEAKER DEFINITION OF PRESENT EQUIVALENTS

Recall that for any stream $x \in X$, its *present equivalent* is defined by

$$c_x \sim x,$$

where c_x pays 0 consumption in all future periods $t > 0$. The requirement of 0 future consumption can be relaxed if for every $t > 0$, the agent's impatience is magnitude-*independent* for all consumption below some threshold $c^*_t > 0$ (in terms of the representation, this condition is equivalent $\underline{d}_t > 0$ and $\lim_{\delta \searrow \underline{d}_t} \varphi'_t(\delta) > 0$). In this case, present equivalents can be defined with future consumption fixed at $c_t^* > 0$ in each t, and Weak Homotheticity can be adapted accordingly to show that, in the GDU representation, it yields MDI. This is because the key requirement for our behavioral definition of MDI is that scaling down the present equivalent (along with the future stream $c_1^*, \dots, c_T^*$) does not change the optimal discount function applied to that stream. We demonstrate all this next.

Assume that there exists a stream x^* such that for all $t \geq 1$ and for all $x_t \precsim x_t^*$ and $x_0 \succsim x_0^*$ with

$$\left(x_0, x_1^*, \dots, x_{t-1}^*, x_t, x_{t+1}^* \cdots, x_T^*\right) \sim x^*,$$

we have

$$\alpha \circ \left(x_0, x_1^*, \dots, x_{t-1}^*, x_t, x_{t+1}^* \cdots, x_T^*\right) \sim \alpha \circ x^*$$

for all $\alpha \in (0, 1]$. Let us confirm that this stream defines a threshold for "magnitude-independent impatience" in the representation: If $\succsim$ admits a CE representation, this indifference means

$$\begin{aligned} u(\alpha \circ x_0) &+ D_{u(\alpha \circ x_t)}(t) u(\alpha \circ x_t) + \sum_{\tau \neq 0, t} D_{u(\alpha \circ x_\tau^*)}(\tau) u\left(\alpha \circ x_\tau^*\right) \\ &= u\left(\alpha \circ x_0^*\right) + D_{u(\alpha \circ x_t^*)}(t) u\left(\alpha \circ x_t^*\right) + \sum_{\tau \neq 0, t} D_{u(\alpha \circ x_\tau^*)}(\tau) u\left(\alpha \circ x_\tau^*\right), \end{aligned}$$

which is equivalent to

$$\begin{aligned} & u(\alpha \circ x_0) - u\left(\alpha \circ x_0^*\right) = D_{u(\alpha \circ x_t^*)}(t) u\left(\alpha \circ x_t^*\right) - D_{u(\alpha \circ x_t)}(t) u(\alpha \circ x_t) \\ \iff \quad & \beta_\alpha u(x_0) - \beta_\alpha u\left(x_0^*\right) = \beta_\alpha D_{u(\alpha \circ x_t^*)}(t) u\left(x_t^*\right) - \beta_\alpha D_{u(\alpha \circ x_t)}(t) u(x_t) \\ \iff \quad & u(x_0) - u\left(x_0^*\right) = D_{u(\alpha \circ x_t^*)}(t) u\left(x_t^*\right) - D_{u(\alpha \circ x_t)}(t) u(x_t). \end{aligned}$$

Since this equality holds for $\alpha = 1$ in particular, we have

$$D_{u(x_t^*)}(t) u\left(x_t^*\right) - D_{u(x_t)}(t) u(x_t) = D_{u(\alpha \circ x_t^*)}(t) u\left(x_t^*\right) - D_{u(\alpha \circ x_t)}(t) u(x_t).$$

By rearrangement,

$$\left(D_{u(x_t^*)}(t) - D_{u(\alpha \circ x_t^*)}(t)\right) u\left(x_t^*\right) = \left(D_{u(x_t)}(t) - D_{u(\alpha \circ x_t)}(t)\right) u(x_t) \tag{S4}$$

holds for all $\alpha \in (0, 1)$ and all $u(x_t) \leq u(x_t^*)$. We want to show that $D_r(t)$ is constant for all $r \leq u(x_t^*)$. Seeking a contradiction, suppose there exists some $r \leq u(x_t^*)$ such that $D_r(t) \neq D_{u(x_t^*)}(t)$. Take some $\alpha \in [0, 1]$ satisfying $r = u(\alpha \circ x_t^*)$. Then the left-hand side of (S4) is not equal to zero. On the other hand, as $u(x_t) \to 0$, the right-hand side of (S4) vanishes to zero because $D_{u(x_t)}(t) - D_{u(\alpha \circ x_t)}(t)$ is bounded. This is a contradiction. Note that since $D_r(t)$ is constant for all $r \leq u(x_t^*)$, it is equal to the minimal discount factor $\underline{d}_t$.

Now we show that we can replace zero future consumption in the definition of present equivalents with the consumption stream x^* derived above. For any stream x such that $x_t \succsim x_t^*$ for all $t \geq 1$, define the present equivalent $c_x \in C$ by $(c_x, x_1^*, \dots, x_T^*) \sim x$. Weak Homotheticity requires that for all $\alpha \in (0, 1)$,

$$\left(\alpha \circ c_x, \alpha \circ x_1^*, \dots, \alpha \circ x_T^*\right) \succsim (\alpha \circ x_0, \alpha \circ x_1, \dots, \alpha \circ x_T).$$

In particular, if we take any stream x that gives $p \succsim x_t^*$ at time $t \geq 1$ and gives x_τ^* at the other τ, and write its certainty equivalent as c_{p^t}, then

$$\left(\alpha \circ c_{p^t}, \alpha \circ x_1^*, \ldots, \alpha \circ x_t^*, \ldots, \alpha \circ x_T^*\right) \succsim \left(\alpha \circ x_0^*, \alpha \circ x_1^*, \ldots, \alpha \circ p, \ldots, \alpha \circ x_T^*\right).$$

This implies the desired MDI property: By the CE representation, and given the homogeneity of u, the left-hand side is equivalent to

$$\begin{aligned} u(\alpha \circ c_{p^t}) + \sum D_{u(\alpha \circ x_t^*)}(t) u\left(\alpha \circ x_t^*\right) &= u(\alpha \circ c_{p^t}) + \sum \underline{d}_t u\left(\alpha \circ x_t^*\right) \\ &= \beta_\alpha \Big(u(c_{p^t}) + \sum \underline{d}_t u\left(x_t^*\right)\Big) = \beta_\alpha U(x) \\ &= \beta_\alpha \Bigg(D_{u(p)}(t) u(p) + u\left(x_0^*\right) + \sum_{\tau \neq t, 0} \underline{d}_\tau u\left(x_\tau^*\right)\Bigg), \end{aligned}$$

while the right-hand side is equivalent to

$$\begin{aligned} & D_{u(\alpha \circ p)}(t) u(\alpha \circ p) + u\left(\alpha \circ x_0^*\right) + \sum_{\tau \neq t, 0} D_{u(\alpha \circ x_\tau^*)}(\tau) u\left(\alpha \circ x_\tau^*\right) \\ &= D_{u(\alpha \circ p)}(t) u(\alpha \circ p) + u\left(\alpha \circ x_0^*\right) + \sum_{\tau \neq t, 0} \underline{d}_\tau u\left(\alpha \circ x_\tau^*\right) \\ &= \beta_\alpha \Bigg(D_{u(\alpha \circ p)}(t) u(p) + u\left(x_0^*\right) + \sum_{\tau \neq t, 0} \underline{d}_\tau u\left(x_\tau^*\right)\Bigg). \end{aligned}$$

Thus, we have $D_{u(p)}(t) \geq D_{\beta_\alpha u(p)}(t)$, which implies that $D_r(t)$ is weakly increasing in $r \geq u(x_t^*)$, as desired.

S4. QUASI-STATIONARITY

In the context of the homogeneous CE model, NT provide a characterization of the following condition.

AXIOM S6—Quasi-Stationarity: *For any streams x, y such that $x_0 = y_0 = x_T = y_T = 0$ and any c,*

$$x \succsim y \iff cx \succsim cy.$$

PROPOSITION S1—NT (2022, Proposition 1): *Suppose that $T \geq 3$. A homogeneous CE representation $(u, m, \{\overline{d}_t, a_t\})$ satisfies Quasi-Stationarity iff there exist $c^* > 0$, $0 < \delta \leq 1$, and $0 < \beta \leq 1/(\delta u(c^*)^{\frac{1}{m-1}})$ such that*

$$a_t = \frac{1}{m\beta^{m-1}\left(\delta^{m-1}\right)^t}, \quad \overline{d}_t = \beta\delta^t u(c^*)^{\frac{1}{m-1}} \quad \textit{for each } t,$$

and the optimal discount function takes the form:

$$D_c(t) = \begin{cases} \beta\delta^t u(c)^{\frac{1}{m-1}} & \text{if } c \leq c^*, \\ \beta\delta^t u(c^*)^{\frac{1}{m-1}} & \text{if } c > c^*. \end{cases}$$

PROOF: Denote by $\succsim^1$ the preference over one-period-delayed streams: $x \succsim^1 y \Longleftrightarrow 0x \succsim 0y$. From the reduced form of the homogeneous CE model as given in NT (2022, Theorem 2), we see that $\succsim^1$ is represented by

$$U^1(x) = \sum_{t\geq 1} \kappa_t u(x_t)^{\frac{m}{m-1}} \quad \text{for all } x_t \leq u^{-1}\big(ma_t \overline{d}_t^{m-1}\big).$$

Let $\delta = \frac{\kappa_2}{\kappa_1}$. By definition of a homogeneous CE model, $\kappa_1 \geq \kappa_2 > 0$ and so we have $0 < \delta \leq 1$. There exist sufficiently small but strictly positive c, c' such that $U^1(c, 0, \dots, 0) = U^1(0, c', 0, \dots, 0)$, which is equivalent to $\kappa_1 u(c)^{\frac{m}{m-1}} = \kappa_2 u(c')^{\frac{m}{m-1}}$. By Quasi-Stationarity, $\kappa_2 u(c)^{\frac{m}{m-1}} = \kappa_3 u(c')^{\frac{m}{m-1}}$ and so $\delta = \frac{\kappa_2}{\kappa_1} = \frac{\kappa_3}{\kappa_2}$. By repeating this argument, we have $\frac{\kappa_{t+1}}{\kappa_t} = \delta$ for all $t \geq 1$, and so

$$\kappa_t = \delta^{t-1}\kappa_1 = \beta\delta^t,$$

where $\beta = \kappa_1/\delta > 0$. Since $\kappa_t = (\frac{1}{ma_t})^{\frac{1}{m-1}}$, we obtain

$$a_t = \frac{1}{m\kappa_t^{m-1}} = \frac{1}{m\big(\beta\delta^t\big)^{m-1}} = \frac{1}{m\beta^{m-1}\big(\delta^{m-1}\big)^t}.$$

The optimal discount function takes the form

$$D_c(t) = \left(\frac{u(c)}{ma_t}\right)^{\frac{1}{m-1}} = \kappa_t u(c)^{\frac{1}{m-1}} = \beta\delta^t u(c)^{\frac{1}{m-1}}$$

for all $u(c) \leq (\frac{\overline{d}_t}{\beta\delta^t})^{m-1}$.

Next, take sufficiently large c, c' satisfying $U^1((c, 0, \dots, 0)) = U^1((0, c', 0, \dots, 0))$, which then yields $\overline{d}_1 u(c) = \overline{d}_2 u(c')$. By Quasi-Stationarity, $\overline{d}_2 u(c) = \overline{d}_3 u(c')$. Let $\overline{\delta} = \frac{\overline{d}_2}{\overline{d}_1}$. By repeating the same argument as above, we have $\frac{\overline{d}_{t+1}}{\overline{d}_t} = \overline{\delta}$ for all $t \geq 1$, and so

$$\overline{d}_t = \overline{\delta}^{t-1}\overline{d}_1.$$

Moreover, let c_t^* be the threshold consumption at t satisfying $u(c_t^*) = (\frac{\overline{d}_t}{\beta\delta^t})^{m-1}$. Since $\overline{d}_t, \beta, \delta > 0$, it must be that $c_t^* > 0$. Then $\overline{d}_t = \beta\delta^t u(c_t^*)^{\frac{1}{m-1}} > 0$, and $\overline{\delta} = \frac{\overline{d}_{t+1}}{\overline{d}_t} = \delta(\frac{u(c_{t+1}^*)}{u(c_t^*)})^{\frac{1}{m-1}} > 0$. Equivalently, $\frac{u(c_{t+1}^*)}{u(c_t^*)} = (\frac{\overline{\delta}}{\delta})^{m-1} := \overline{\alpha} > 0$. We have therefore established that

$$D_c(t) = \begin{cases} \beta\delta^t u(c)^{\frac{1}{m-1}} & \text{if } u(c) \leq u\big(c_t^*\big), \\ \beta\delta^t u\big(c_t^*\big)^{\frac{1}{m-1}} & \text{otherwise.} \end{cases}$$

The value of $D_c(t)$ depends on c_t^*. The proof is complete once we can show that $c_{t+1}^* = c_t^*$ for all t. It suffices to establish that $\overline{\alpha} = \frac{u(c_{t+1}^*)}{u(c_t^*)} = 1$. Seeking a contradiction, first suppose $\overline{\alpha} > 1$. Since $u(c_{t+1}^*) = \overline{\alpha} u(c_t^*)$ for all t, it must be that $c_{t+1}^* > c_t^*$ for all t. From the preceding, we know that $c_2^* > 0$.

Consider c_2^* and take ε that satisfies

$$U\big(0, c_2^* + \varepsilon, 0, \dots, 0\big) = U\big(0, c_2^*, c_2^*, 0, \dots, 0\big).$$

By Monotonicity it must be that $\varepsilon > 0$. Since $c_1^* < c_2^* < c_2^* + \varepsilon$, we can compute that

$$\begin{aligned}
&U\big(0, c_2^* + \varepsilon, 0, \dots, 0\big) = U\big(0, c_2^*, c_2^*, 0, \dots, 0\big)\\
&\iff \quad D_{c_2^*+\varepsilon}(1)u\big(c_2^* + \varepsilon\big) = D_{c_2^*}(1)u\big(c_2^*\big) + D_{c_2^*}(2)u\big(c_2^*\big)\\
&\iff \quad \beta\delta u\big(c_1^*\big)^{\frac{1}{m-1}}u\big(c_2^* + \varepsilon\big) = \beta\delta u\big(c_1^*\big)^{\frac{1}{m-1}}u\big(c_2^*\big) + \beta\delta^2 u\big(c_2^*\big)^{\frac{1}{m-1}}u\big(c_2^*\big)\\
&\iff \quad \frac{u\big(c_2^* + \varepsilon\big)}{u\big(c_2^*\big)} = 1 + \delta\left(\frac{u\big(c_2^*\big)}{u\big(c_1^*\big)}\right)^{\frac{1}{m-1}},
\end{aligned}$$

which yields our first expression for $\frac{u(c_2^*+\varepsilon)}{u(c_2^*)}$. Given Quasi-Stationarity, we reason further that

$$\begin{aligned}
&U\big(0, c_2^* + \varepsilon, 0, \dots, 0\big) = U\big(0, c_2^*, c_2^*, 0, \dots, 0\big)\\
&\iff \quad U\big(0, 0, c_2^* + \varepsilon, 0, \dots, 0\big) = U\big(0, 0, c_2^*, c_2^*, 0, \dots, 0\big)\\
&\iff \quad D_{c_2^*+\varepsilon}(2)u\big(c_2^* + \varepsilon\big) = D_{c_2^*}(2)u\big(c_2^*\big) + D_{c_3^*}(3)u\big(c_2^*\big)\\
&\iff \quad \beta\delta^2 u\big(c_2^*\big)^{\frac{1}{m-1}}u\big(c_2^* + \varepsilon\big) = \beta\delta^2 u\big(c_2^*\big)^{\frac{1}{m-1}}u\big(c_2^*\big) + \beta\delta^3 u\big(c_2^*\big)^{\frac{1}{m-1}}u\big(c_2^*\big)\\
&\iff \quad \frac{u\big(c_2^* + \varepsilon\big)}{u\big(c_2^*\big)} = 1 + \delta.
\end{aligned}$$

Therefore, we obtain the second expression for $\frac{u(c_2^*+\varepsilon)}{u(c_2^*)}$. Putting both together, we see that

$$1 + \delta\left(\frac{u\big(c_2^*\big)}{u\big(c_1^*\big)}\right)^{\frac{1}{m-1}} = 1 + \delta.$$

Since all the terms are strictly positive and $m > 1$, we conclude that $\frac{u(c_2^*)}{u(c_1^*)} = 1$. But then $\overline{\alpha} = \frac{u(c_2^*)}{u(c_1^*)} = 1$, while we had supposed that $\overline{\alpha} > 1$, a contradiction.

Next, suppose by way of contradiction that $\overline{\alpha} < 1$. Since it is the case that $u(c_{t+1}^*) = \overline{\alpha}u(c_t^*)$ for all t, it must be that $c_{t+1}^* < c_t^*$ for all t. Consider c_1^*, c_2^*, c_3^* and take $\varepsilon > 0$ that satisfies

$$U\big(0, c_1^* + \varepsilon, 0, \dots, 0\big) = U\big(0, c_1^*, c_3^*, 0, \dots, 0\big).$$

Since by hypothesis, $c_3^* < c_2^* < c_1^*$, we can compute that

$$\begin{aligned}
&U\big(0, c_1^* + \varepsilon, 0, \dots, 0\big) = U\big(0, c_1^*, c_3^*, 0, \dots, 0\big)\\
&\iff \quad D_{c_1^*+\varepsilon}(1)u\big(c_1^* + \varepsilon\big) = D_{c_1^*}(1)u\big(c_1^*\big) + D_{c_3^*}(2)u\big(c_3^*\big)\\
&\iff \quad \beta\delta u\big(c_1^*\big)^{\frac{1}{m-1}}u\big(c_1^* + \varepsilon\big) = \beta\delta u\big(c_1^*\big)^{\frac{1}{m-1}}u\big(c_1^*\big) + \beta\delta^2 u\big(c_3^*\big)^{\frac{1}{m-1}}u\big(c_3^*\big)\\
&\iff \quad u\big(c_1^* + \varepsilon\big) = u\big(c_1^*\big) + \delta\left(\frac{u\big(c_3^*\big)}{u\big(c_1^*\big)}\right)^{\frac{1}{m-1}}u\big(c_3^*\big).
\end{aligned}$$

By Quasi-Stationarity, we should also have

$$\begin{aligned} &U\big(0,0,c_1^*+\varepsilon,0,\dots,0\big)=U\big(0,0,c_1^*,c_3^*,0,\dots,0\big)\\ \iff\quad & D_{c_1^*+\varepsilon}(2)u\big(c_1^*+\varepsilon\big)=D_{c_1^*}(2)u\big(c_1^*\big)+D_{c_3^*}(3)u\big(c_3^*\big)\\ \iff\quad & \beta\delta^2 u\big(c_2^*\big)^{\frac{1}{m-1}}u\big(c_1^*+\varepsilon\big)=\beta\delta^2 u\big(c_2^*\big)^{\frac{1}{m-1}}u\big(c_1^*\big)+\beta\delta^3 u\big(c_3^*\big)^{\frac{1}{m-1}}u\big(c_3^*\big)\\ \iff\quad & u\big(c_1^*+\varepsilon\big)=u\big(c_1^*\big)+\delta\left(\frac{u(c_3^*)}{u(c_2^*)}\right)^{\frac{1}{m-1}}u\big(c_3^*\big). \end{aligned}$$

From the preceding, we obtain two expressions for $u(c_1^*+\varepsilon)-u(c_1^*)$. Putting them together we see that $\delta(\frac{u(c_3^*)}{u(c_1^*)})^{\frac{1}{m-1}}u(c_3^*)=\delta(\frac{u(c_3^*)}{u(c_2^*)})^{\frac{1}{m-1}}u(c_3^*)$, which is equivalent to

$$u\big(c_1^*\big)=u\big(c_2^*\big).$$

This implies $\overline{\alpha}=\frac{u(c_2^*)}{u(c_1^*)}=1$ whereas we had assumed $\overline{\alpha}<1$, a contradiction.

Since $D_{c^*}(1)=\beta\delta u(c^*)^{\frac{1}{m-1}}\le 1$, which yields $\beta\le 1/(\delta u(c^*)^{\frac{1}{m-1}})$. This completes the proof of sufficiency.

We now establish necessity. Define a time-invariant function f by $f(c)=u(c)^{\frac{m}{m-1}}$ for any $c\le c^*$ and $f(c)=u(c^*)^{\frac{1}{m-1}}u(c)$ otherwise. Then the representation is written as $U(x)=u(x_0)+\sum_{t\ge1}\beta\delta^t f(x_t)$.

Now take any x, y satisfying the presumption of Quasi-Stationarity. For any c,

$$\begin{aligned} U(x)\ge U(y)\quad &\iff\quad \sum_{t\ge1}\beta\delta^t f(x_t)\ge\sum_{t\ge1}\beta\delta^t f(y_t)\\ &\iff\quad \sum_{t\ge1}\beta\delta^{t+1} f(x_t)\ge\sum_{t\ge1}\beta\delta^{t+1} f(y_t)\\ &\iff\quad u(c)+\sum_{t\ge1}\beta\delta^{t+1} f(x_t)\ge u(c)+\sum_{t\ge1}\beta\delta^{t+1} f(y_t)\\ &\iff\quad U(cx)\ge U(cy), \end{aligned}$$

as desired. *Q.E.D.*

S5. CONSUMPTION SMOOTHING

In the context of the homogeneous CE model, NT provide a characterization of the following condition.

DEFINITION S1—Consumption Smoothing: *A preference* $\succsim$ *exhibits consumption smoothing if for any* $\alpha\in[0,1]$ *and for all deterministic streams* $x, y\in C^{T+1}$ *and* $\alpha x+(1-\alpha)y\in C^{T+1}$,

$$x\sim y\quad\Longrightarrow\quad \alpha x+(1-\alpha)y\succsim x.$$

NT state that

PROPOSITION S2—NT (2022, Proposition 3): *Assume that* $\succsim$ *admits a homogeneous CE representation. If* $u(c)^{\frac{m}{m-1}}$ *is concave in* $c \in \mathbb{R}_+$, *then* $\succsim$ *exhibits consumption smoothing. Conversely, if* $\succsim$ *exhibits consumption smoothing, then at least* T *of functions* $u(x_0)$, $D_{u(x_1)}u(x_1), \ldots, D_{u(x_T)}u(x_T)$ *are concave. Moreover,* $u : \mathbb{R}_+ \to \mathbb{R}_+$ *is concave if there are* $t, s \geq 1$ *such that* $\overline{x}_t \neq \overline{x}_s$.

PROOF: We first show the "if" part. First of all, note that u is concave because u is an increasing concave transformation of a concave function $u(c)^{\frac{m}{m-1}}$. Note also that $D_{u(x_t)}(t)u(x_t) = \kappa_t u(x_t)^{\frac{m}{m-1}}$ if $x_t \leq u^{-1}(\overline{r}_t)$ and $D_{u(x_t)}(t)u(x_t) = \overline{d}_t u(x_t)$ if $x_t > u^{-1}(\overline{r}_t)$. Since $D_{u(x_t)}(t)u(x_t)$ is the pointwise minimum of two concave functions on $\mathbb{R}_+$, that is, $\kappa_t u(x_t)^{\frac{m}{m-1}}$ and $\overline{d}_t u(x_t)$, it is concave. Thus, $U(x) = u(x_0) + \sum_{t>0} D_{u(x_t)}(t)u(x_t)$ is concave in deterministic streams x, which in turn implies that $\succsim$ has preference for consumption smoothing.

Next, we show the "only if" part. We first claim that if $\succsim$ has preference for consumption smoothing, then $U(x)$ is quasi-concave in deterministic consumption streams. It suffices to show that $x \succsim y$ implies $\alpha x + (1-\alpha)y \succsim y$. If $x \sim y$, it follows directly from preference for consumption smoothing. Now assume $x \succ y$. Seeking a contradiction, suppose $y \succ \tilde{\alpha}x + (1-\tilde{\alpha})y$ for some $\tilde{\alpha} \in (0, 1)$. Let $A = \{\alpha \in [0, 1] \,|\, y \succsim \alpha x + (1-\alpha)y\}$. Since $\tilde{\alpha} \in A$, A is nonempty. By Continuity, A is closed, that is, compact. Thus, there exists a maximum of A, denoted by $\overline{\alpha}$.

Let $x_{\overline{\alpha}} = \overline{\alpha}x + (1-\overline{\alpha})y$. We show that $x_{\overline{\alpha}} \sim y$. By definition, $\overline{\alpha} \in A$. If $\overline{\alpha} = 1$, $x_{\overline{\alpha}} = x \succ y$, which is a contradiction. Thus, $\overline{\alpha} < 1$. By seeking a contradiction, suppose $x_{\overline{\alpha}} \nsim y$. That is, $y \succ x_{\overline{\alpha}}$. Since $A^\circ = \{\alpha \in [0, 1] \,|\, y \succ \alpha x + (1-\alpha)y\}$ is open and $\overline{\alpha} \in A^\circ$, we can find some $\hat{\alpha} > \overline{\alpha}$ with $\hat{\alpha} \in A^\circ$. But, this contradicts the maximality of $\overline{\alpha}$ in A.

Now, since $x_{\overline{\alpha}} \sim y$, by preference for consumption smoothing, $\lambda x_{\overline{\alpha}} + (1-\lambda)y \succsim y$ for all $\lambda \in (0, 1)$. In particular, let $\lambda = \frac{\tilde{\alpha}}{\overline{\alpha}} \in (0, 1)$. On the other hand, by assumption,

$$\frac{\tilde{\alpha}}{\overline{\alpha}}\big(\overline{\alpha}x + (1-\overline{\alpha})y\big) + \left(1 - \frac{\tilde{\alpha}}{\overline{\alpha}}\right)y = \tilde{\alpha}x + (1-\tilde{\alpha})y \prec y,$$

which is a contradiction. Therefore, $x \succ y$ implies $\alpha x + (1-\alpha)y \succsim y$ for all $\alpha \in (0, 1)$, as desired.

Yaari (1977) shows that if an additively separable function $F(x_1, \ldots, x_S) = \sum_s f_s(x_s)$ is quasi-concave, then at least $S-1$ of functions $f_1, \ldots, f_S$ are concave. Therefore, at least T of $u(x_0), D_{u(x_1)}u(x_1), \ldots, D_{u(x_T)}u(x_T)$ are concave. If u is included in the group of T functions, we are done. Hence, assume that $D_{u(x_1)}u(x_1), \cdots, D_{u(x_T)}u(x_T)$ are concave. Now assume in addition that $\overline{x}_t > \overline{x}_s$ for some t, s. Then u is concave on $[0, \overline{r}_t]$ and $[\overline{r}_s, \infty)$, which implies that u is concave on $\mathbb{R}_+$ by Lemma 2.2 of Li and Yeh (2010). *Q.E.D.*

S6. PROCRASTINATION

PROPOSITION S3—NT (2022, Proposition 4): *Consider a sophisticated DU agent. Suppose self* 2 *would not complete any task when there is only one to be done. Then, there exists a unique subgame perfect equilibrium and, in it, neither self* 0 *nor self* 2 *completes any tasks.*

PROOF: Denote by $V_t(n|m)$ the utility of self $t = 0, 2$ of completing n tasks in period t out of m uncompleted tasks in that period. As is standard, we proceed using backward induction. Suppose that self 2 would not exert the effort to complete one task when one remains to be completed:

$$V_2(1|1) = u(b-e) + D(1)u(b+R) < u(b) + D(1)u(b) = V_2(0|1),$$

which is equivalent to

$$D(1)\big[u(b+R) - u(b)\big] < u(b) - u(b-e). \tag{S5}$$

Since the expressions remain the same even if there were two tasks to be completed, we see that $V_2(1|1) < V_2(0|1)$ implies that $V_2(1|2) < V_2(0|2)$, that is, self 2 would do zero tasks rather than one task if 2 tasks remained to be done.

Given weak concavity of u, $u(b+2R) - u(b) \leq 2[u(b+R) - u(b)]$ and $u(b) - u(b-2e) \geq 2[u(b) - u(b-e)]$, and so by (S5),

$$D(1)\big[u(b+2R) - u(b)\big] < u(b) - u(b-2e). \tag{S6}$$

It follows that

$$V_2(2|2) = u(b-2e) + D(1)u(b+2R) < u(b) + D(1)u(b) = V_2(0|2),$$

that is, self 2 will not do two tasks together.

We have therefore shown that self 2 will never complete any task, regardless of how many tasks have been completed by self 0. We show next that self 0 will not complete any task either. Indeed, conditional on self 2 never completing any task, self 0 would not do any of the tasks either because her choice considerations are identical to those of self 2: like self 2, self 0 must decide whether to incur effort costs today for a return tomorrow. We conclude that no self will do any task. *Q.E.D.*

PROPOSITION S4—NT (2022, Proposition 5): *Consider a sophisticated CE agent. If self 2 would not complete any task when there is only one to be done, then there exists a unique subgame perfect equilibrium, and it permits only the following three possibilities:*

(i) *Neither of self 0 nor self 2 completes any tasks.*
(ii) *Self 0 completes no task and self 2 completes 2 tasks.*
(iii) *Self 0 completes 2 tasks.*

PROOF: Consider the homogeneous CE model where the optimal discount factor is given by $D_r(t) = \kappa_t r^{\frac{1}{m-1}}$ if $r \leq \overline{r}_t$ and $D_r(t) = \overline{d}_t$ otherwise, where $\kappa_t = (ma_t)^{-\frac{1}{m-1}}$ and $\overline{r}_t = ma_t\overline{d}_t^{m-1}$. Let $U_t(n|m)$ denote the utility of self $t = 0, 2$ of completing n tasks in period t given that there are m uncompleted tasks in that period.

The hypothesis states that self 2 would not exert effort when there is 1 task to complete:

$$\begin{aligned} &U_2(1|1) < U_2(0|1) \\ &\iff \quad u(b-e) + D_{u(b+R)}(1)u(b+R) < u(b) + D_{u(b)}(1)u(b) \\ &\iff \quad D_{u(b+R)}(1)u(b+R) - D_{u(b)}(1)u(b) < u(b) - u(b-e). \end{aligned} \tag{S7}$$

Next, consider the subgame where self 2 faces two tasks. It is easy to see that the expressions are no different if there were 2 tasks to complete. Therefore,

$$U_2(1|1) < U_2(0|1) \quad \Longrightarrow \quad U_2(1|2) < U_2(0|2). \tag{S8}$$

Since, by (S8), completing 0 tasks is preferred to completing 1 task, we see that self 2 completes both tasks if and only if

$$\begin{aligned} & U_2(2|2) \geq U_2(0|2) \\ \iff \quad & u(b-2e) + D_{u(b+2R)}(1)u(b+2R) \geq u(b) + D_{u(b)}(1)u(b) \\ \iff \quad & D_{u(b+2R)}(1)u(b+2R) - D_{u(b)}(1)u(b) \geq u(b) - u(b-2e). \end{aligned} \tag{S9}$$

A novel feature of the CE model is that self 2 may complete two tasks due to magnitude-dependent impatience, even when she is reluctant to complete one task. To see that inequalities (S7) and (S9) can both hold, suppose that $u(b+R)$ is below $\bar{r}_1$, and $u(b+2R)$ is above $\bar{r}_1$. Then, (S7) and (S9) are reduced to

$$\kappa_1\big[u(b+R)^{\frac{m}{m-1}} - u(b)^{\frac{m}{m-1}}\big] < u(b) - u(b-e), \tag{S10}$$

$$\overline{d}_1 u(b+2R) - \kappa_1 u(b)^{\frac{m}{m-1}} \geq u(b) - u(b-2e), \tag{S11}$$

respectively. Moreover, suppose m is sufficiently close to one. Then, the curvature of the convex transformation $z^{\frac{m}{m-1}}$ is so strong that payoffs $u(b)^{\frac{m}{m-1}}$ and $u(b+R)^{\frac{m}{m-1}}$ become negligible compared with a payoff $\overline{d}_1 u(b+2R)$. Hence, there exist parameter values for which both (S10) and (S11) can hold simultaneously. For example, when $a_1 = \frac{1}{m}$ and $\overline{d}_1 = 1$, we have $\bar{r}_1 = \kappa_1 = 1$. Moreover, assume $u(b) - u(b-2e) \leq 1$. Since $u(b+2R) > 1 > u(b+R)$, $u(b)^{\frac{m}{m-1}} \to 0$ and $u(b+R)^{\frac{m}{m-1}} \to 0$ as $m \to 1$, and hence, both (S10) and (S11) hold.

Consider two cases for self 2's behavior and derive the corresponding self 0 behavior.

Case (i): $U_2(2|2) < U_2(0|2)$

That is, self 2 would not complete two tasks. By hypothesis she would not complete 1 task either when facing one task to be completed. Given self 2's optimal actions on the subgames, self 0's considerations are identical with those of self 2 facing with two tasks. By hypothesis and case (i), self 0 would not complete any tasks either. This establishes the first possibility in the statement of the proposition.

Case (ii): $U_2(2|2) \geq U_2(0|2)$

That is, self 2 would complete both tasks. First, rule out the possibility that self 0 will complete 1 task. Recall that for self 2, completing one task is dominated by completing none, which is in turn dominated by completing two tasks by case (ii). Since self 2 does not complete any task after self 0 completes one task, self 0's comparison between completing one and two tasks is identical with self 2's comparison between these two actions. Thus, for self 0, completing one task is dominated by completing two tasks.

Now we compare self 0's utilities from completing two tasks or none. If self 0 completes both tasks, then her utility is given by

$$U_0(2|2) = u(b-2e) + D_{u(b+2R)}(1)u(b+2R) + D_{u(b)}(2)u(b) + D_{u(b)}(3)u(b),$$

and if she completes none, then given that self 2 will complete both, her utility is

$$U_0(0|2) = u(b) + D_{u(b)}(1)u(b) + D_{u(b-2e)}(2)u(b-2e) + D_{u(b+2R)}(3)u(b+2R).$$

Therefore, self 0 completes both tasks iff

$$\begin{aligned} & U_0(2|2) \geq U_0(0|2) \\ \iff & [D_{u(b+2R)}(1)u(b+2R) - D_{u(b)}(1)u(b)] + [D_{u(b)}(2)u(b) - D_{u(b-2e)}(2)u(b-2e)] \\ & \geq u(b) - u(b-2e) + [D_{u(b+2R)}(3)u(b+2R) - D_{u(b)}(3)u(b)]. \end{aligned} \tag{S12}$$

Another novel feature of the CE model is that self 0 may exploit self 2's incentive to complete both tasks by leaving them to self 2. For example, suppose that $u(b+R)$ is below $\bar{r}_1$, and $u(b+2R)$ is above $\bar{r}_1$. Moreover, assume $\overline{d}_t = 1$ for all t, and a_1 and a_3 are sufficiently close. Then, $\bar{r}_1$ and $\bar{r}_3$ are close, and κ_1 and κ_3 are also close to each other. Then, (S12) is reduced to

$$\begin{aligned} & \big[\overline{d}_1 u(b+2R) - \kappa_1 u(b)^{\frac{m}{m-1}}\big] + \kappa_2\big[u(b)^{\frac{m}{m-1}} - u(b-2e)^{\frac{m}{m-1}}\big] \\ & \geq [u(b) - u(b-2e)] + \big[\overline{d}_3 u(b+2R) - \kappa_3 u(b)^{\frac{m}{m-1}}\big]. \end{aligned} \tag{S13}$$

By assumption, the first bracket of the left-hand side and the second bracket of the right-hand side are almost the same. Moreover, if m is close to one, the curvature of the convex transformation $z^{\frac{m}{m-1}}$ is so strong that small payoffs $u(b)^{\frac{m}{m-1}}$ and $u(b-2e)^{\frac{m}{m-1}}$ become negligible, and the inequality (S13) is almost dominated by the magnitude of $u(b) - u(b-2e) > 0$. Thus, $U_0(0|2) > U_0(2|2)$ may hold.

Conclude that, depending on parameters, self 0 either completes both tasks by herself or leaves both to self 2 to complete, who then completes them. *Q.E.D.*

S7. SMOOTH HOMOGENEOUS CE REPRESENTATION

Consider preference $\succsim$ on $X = C^{T+1}$. In NT, the homogeneous CE representation is characterized on this domain by exploiting the magnitude sensitivity toward the scaling operation. In this section, we provide an alternative way to capture the magnitude sensitivity in terms of the MRS and derive the corresponding representation result.

S7.1. *Axiomatization*

DEFINITION S2—Magnitude-Sensitivity on the Diagonal: *A stream $x \in X$ on the diagonal is magnitude sensitive wrt period $t > 0$ if*

$$MRS_{\alpha x}(t) < MRS_x(t) \quad \textit{for all } \alpha \in (0,1).$$

The set of all magnitude sensitive streams wrt period t is denoted by $X_m(t) \subset X$.

Magnitude sensitivity behaviorally identifies streams on the diagonal for which the discount function is strictly increasing in consumption. We place structure on magnitude sensitivity.

AXIOM S7—X_m-Regularity: *For any $t > 0$ and any stream $x \in X$ on the diagonal,*
(i) *if $x \notin X_m(t)$, then $\alpha x \in X_m(t)$ for some $\alpha \in (0,1]$, and*
(ii) *if $x \in X_m(t)$, then $\alpha x \in X_m(t)$ for all $\alpha \in (0,1)$.*

Thus, if a stream on the diagonal is not magnitude sensitive then scaling it down makes it so, and keeps it so as it is scaled down further. The content of the axiom is that, along the diagonal, $MRS_x(t)$ is strictly increasing between the origin and some stream, and is constant beyond that stream.

The next condition requires the MRS to define a homogeneous function along the diagonal and parallel to the x_0-axis. Recall our notation: if x is a stream, then αx is a stream that yields αx_t for any t, and $\alpha^0 \cdot x$ is a stream that yields αx_0 in period 0 and x_t in any future period.

AXIOM S8—MRS-Homogeneity: *For any* $t, s > 0$, *any* $\alpha \in (0, 1]$, *and any* $\beta > 0$, (i) *for any* $x \in X_m(t)$ *and* $y \in X_m(s)$ *on the diagonal*,

$$MRS_{\alpha x}(t) = \beta MRS_x(t) \quad \Longrightarrow \quad MRS_{\alpha y}(s) = \beta MRS_y(s),$$

and (ii) *for any* $x, y \in X$,

$$MRS_{\alpha^0 \cdot x}(t) = \beta MRS_x(t) \quad \Longrightarrow \quad MRS_{\alpha^0 \cdot y}(s) = \beta MRS_y(s).$$

MRS-Homogeneity (i) implies that, for the values of r for which $D_r(t)$ is strictly increasing, $D_{u(x_t)}(t)u(x_t)$ is homogeneous in x_t. In particular, there exists a function that maps each $\alpha \in (0, 1]$ to some $\beta(\alpha) \in (0, 1]$ such that

$$D_{u(\alpha x_t)}(t) = \frac{\beta(\alpha) u(x_t)}{u(\alpha x_t)} \times D_{u(x_t)}(t).$$

In order to obtain homogeneity of D, we require the same of u. This is achieved by MRS-Homogeneity (ii). We obtain the following.

THEOREM S3: *Suppose* $\succsim$ *over* X *admits a smooth GDU representation. Then* $\succsim$ *satisfies Increasing MRS*, X_m*-Regularity, and MRS-Homogeneity if and only if* $\succsim$ *admits a CE representation with a utility index* u *of the power form such that* $\{\varphi_t\}$ *is given by*

$$\varphi_t(d) = \begin{cases} a_t d^m & \text{if } d \in [0, \tilde{d}_t], \\ -A_t \ln(\overline{d}_t - d) + C_t & \text{if } d \in (\tilde{d}_t, \overline{d}_t), \\ \infty & \text{if } d \in [\overline{d}_t, 1], \end{cases} \tag{S14}$$

where $m > 1$, $a_t > 0$ *increasing in* t,

$$\tilde{d}_t := \frac{m-1}{m}\overline{d}_t, \qquad A_t := a_t\left(\frac{m-1}{m}\right)^{m-1}\overline{d}_t^m > 0, \quad \text{and}$$

$$C_t := A_t\left(\frac{m-1}{m} + \ln\frac{\overline{d}_t}{m}\right) \in \mathbb{R}.$$

Since φ_t is smooth, a CE model with the cost function (S14) is called the *smooth homogeneous CE model*. Moreover, since φ_t is strictly increasing and strictly convex, it is easy to obtain an optimal discount factor from the FOC.

COROLLARY S1: *In the smooth homogeneous CE model, the optimal discount function takes the form*

$$D_r(t) = \begin{cases} \kappa_t r^{\frac{1}{m-1}} & \text{if } r \leq \tilde{r}_t, \\ \overline{d}_t - \dfrac{A_t}{r} & \text{if } r > \tilde{r}_t, \end{cases}$$

where $\kappa_t = (m a_t)^{-\frac{1}{m-1}} > 0$ *and* $\tilde{r}_t = m a_t (\tilde{d}_t)^{m-1} > 0$.

This optimal discount function takes a power form on the subdomain of $[0, \widetilde{r}_t]$. Since the right and left derivatives of $D_r(t)$ coincide at $\widetilde{r}_t$, the above $D_r(t)$ is differentiable throughout. Moreover, $D_r(t)$ is strictly increasing on the whole domain and converges to the upper bound $\overline{d}_t$ as $r \to \infty$.

Finally, the smooth homogeneous CE model admits the following uniqueness property.

THEOREM S4: *If there are two smooth homogeneous CE representations* $(u^i, m^i, \{\overline{d}_t^i, a_t^i\})$, $i = 1, 2$ *of the same preference* $\succsim$, *then there exists* $\lambda > 0$ *such that* (*i*) $u^2 = \lambda u^1$, (*ii*) $m^2 = m^1$, $\overline{d}_t^2 = \overline{d}_t^1$, *and* $a_t^2 = \lambda a_t^1$ *for each* t.

S7.2. *Comparison With the Homogeneous CE Model*

As shown by NT (Theorem 2), the optimal discount function in the homogeneous CE model is given by

$$D_r(t) = \begin{cases} \kappa_t r^{\frac{1}{m-1}} & \text{if } r \leq \overline{r}_t, \\ \overline{d}_t & \text{if } r > \overline{r}_t, \end{cases} \tag{S15}$$

where $\kappa_t = (ma_t)^{-\frac{1}{m-1}} > 0$, and $\overline{r}_t = ma_t \overline{d}_t^{m-1} > 0$. This optimal discount function is not differentiable at the threshold.

It is easy to see that the homogeneous CE representation violates the Increasing MRS axiom in Section 4.1 of NT. Since

$$\partial[D_r(t)r] = \begin{cases} \dfrac{m}{m-1} \kappa_t r^{\frac{1}{m-1}} & \text{if } r < \overline{r}_t, \\ \overline{d}_t & \text{if } r > \overline{r}_t, \end{cases}$$

together with Lemma 8 of NT,

$$\begin{aligned} MRS^-_{\overline{r}_t}(t) &= \partial[D_r(t)r]|_{r \nearrow \overline{r}_t} = \frac{m}{m-1} \kappa_t \overline{r}_t^{\frac{1}{m-1}} = \frac{m}{m-1} \overline{d}_t > \overline{d}_t \\ &= \partial[D_r(t)r]|_{r \searrow \overline{r}_t} = MRS^+_{\overline{r}_t}(t), \end{aligned}$$

which violates Increasing MRS. Note that in the smooth homogeneous CE model,

$$MRS^-_{\widetilde{r}_t}(t) = \partial[D_r(t)r]|_{r \nearrow \widetilde{r}_t} = \frac{m}{m-1} \widetilde{d}_t = \overline{d}_t = \partial[D_r(t)r]|_{r \searrow \widetilde{r}_t} = MRS^+_{\widetilde{r}_t}(t),$$

which is consistent with the Increasing MRS axiom.

S7.3. *Illustrations*

We provide a numerical example of the two representations (homogeneous CE and smooth homogeneous CE) under the same parameters as given below.

Assume $m = 2$. Take any $a_t > 0$, $t \geq 1$, increasing in t, and any $0 < \overline{d}_t \leq 1$, $t \geq 1$, decreasing in t. Then, by the formula derived above,

$$\kappa_t = \frac{1}{2a_t}, \qquad \overline{r}_t = 2a_t \overline{d}_t, \qquad \widetilde{d}_t = \frac{1}{2} \overline{d}_t, \qquad \widetilde{r}_t = 2a_t \widetilde{d}_t = \frac{1}{2} \overline{r}_t,$$

$$A_t = \frac{1}{2} a_t \overline{d}_t^2, \qquad C_t = \frac{1}{2} a_t \overline{d}_t^2 \left(\frac{1}{2} + \ln \frac{\overline{d}_t}{2} \right).$$

S7.3.1. *The Case of Homogeneous CE Representation*

Under these parameter values, we first consider the case of homogeneous CE model. The cost function on $[0, \overline{d}_t]$ is given as a quadratic function $\varphi_t(d) = a_t d^2$. The corresponding optimal discount function below the threshold is given as a linear function $D_r(t) = \frac{1}{2a_t} r$. The period-$t$ utility function is obtained as

$$U_t(r) = D_r(t) r = \begin{cases} \frac{1}{2a_t} r^2 & \text{if } r \le \overline{r}_t, \\ \overline{d}_t r & \text{if } r > \overline{r}_t. \end{cases}$$

Note that this is a star-shaped function; it is a quadratic function on $[0, \overline{r}_t]$ and becomes a linear function (passing through the origin) thereafter.

The marginal rate of substitution between time 0 and time t on the diagonal, denoted $MRS_r(t)$, is obtained as

$$MRS_r(t) = \partial\big[D_r(t) r\big] = \begin{cases} \frac{1}{a_t} r & \text{if } r < \overline{r}_t, \\ \overline{d}_t & \text{if } r > \overline{r}_t. \end{cases}$$

Thus, $MRS_r(t)$ is a linear function on $[0, \overline{r}_t]$, takes a value of $2\overline{d}_t$ at $\overline{r}_t = 2a_t\overline{d}_t$, and jumps down to a constant value of $\overline{d}_t$ afterwards.

The solid lines of Figure 1 stand for (a) cost function $\varphi_t(d)$, (b) optimal discount function $D_r(t)$, (c) period-t utility function $U_t(r)$, and (d) $MRS_r(t)$ when $a_t = \frac{1}{2}$ and $\overline{d}_t = 1$.

S7.3.2. *The Case of Smooth Homogeneous CE Representation*

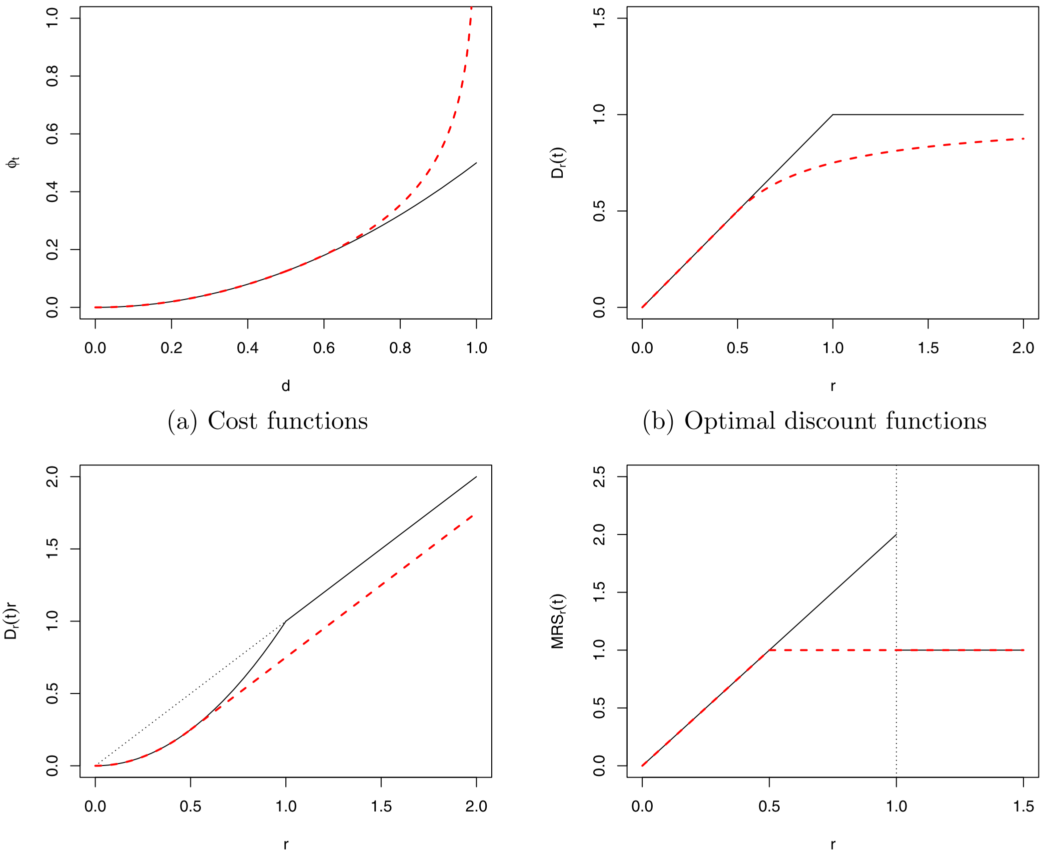
Under the same parameter values, we next consider the case of smooth homogeneous CE model. The cost function is given by

$$\varphi_t(d) = \begin{cases} a_t d^2 & \text{if } d \in [0, \widetilde{d}_t], \\ -\frac{1}{2} a_t \overline{d}_t^2 \log(\overline{d}_t - d) + C_t & \text{if } d \in (\widetilde{d}_t, \overline{d}_t), \\ \infty & \text{if } d \in [\overline{d}_t, 1]. \end{cases}$$

Note that this cost function is smooth on the effective domain because $\varphi_t'(\widetilde{d}_t) = a_t \overline{d}_t$.

The corresponding optimal discount function and period-t utility function are obtained as

$$D_r(t) = \begin{cases} \frac{1}{2a_t} r & \text{if } r \le \widetilde{r}_t, \\ \overline{d}_t - \frac{a_t \overline{d}_t^2}{2r} & \text{if } r > \widetilde{r}_t, \end{cases} \qquad U_t(r) = D_r(t) r = \begin{cases} \frac{1}{2a_t} r^2 & \text{if } r \le \widetilde{r}_t, \\ \overline{d}_t r - \frac{a_t \overline{d}_t^2}{2} & \text{if } r > \widetilde{r}_t. \end{cases}$$



(a) Cost functions (b) Optimal discount functions

(c) Time-t utility functions (d) MRS on the diagonal

FIGURE 1.—$m = 2$, $a_t = \frac{1}{2}$, $\overline{d}_t = 1$: Solid lines stand for graphs of the homogeneous CE model, while dashed lines stand for graphs of the smooth homogeneous CE model.

Note that this is a smooth function such that the derivative at $\widetilde{r}_t$ is given as $\overline{d}_t$. Moreover,

$$MRS_r(t) = \partial\big[D_r(t)r\big] = \begin{cases} \dfrac{1}{a_t}r & \text{if } r < \widetilde{r}_t, \\ \overline{d}_t & \text{if } r > \widetilde{r}_t. \end{cases}$$

Thus, $MRS_r(t)$ is a linear function on $[0, \widetilde{r}_t]$, and takes a constant value of $\overline{d}_t$ afterwards. Therefore, this model is consistent with Increasing MRS.

The dashed lines of Figure 1 stand for (a) cost function $\varphi_t(d)$, (b) optimal discount function $D_r(t)$, (c) period-t utility function $U_t(r)$, and (d) $MRS_r(t)$ when $a_t = \frac{1}{2}$ and $\overline{d}_t = 1$.

S7.4. *Intermediate Characterization*

To characterize the smooth homogeneous CE representation, we use the following proposition as an intermediate lemma.

PROPOSITION S5: *Suppose $\succsim$ over X admits a smooth GDU representation. Then $\succsim$ satisfies Increasing MRS and X_m-Regularity if and only if $\succsim$ admits a CE representation where the optimal discount function $D_r(t)$ has the property that $D_r(t)r$ is a convex function of r, and $[D_r(t)r]'$ is strictly increasing on $[0, \widetilde{r}_t]$ and is constant otherwise.*

PROOF: First, we verify necessity. Since $MRS_r(t) = [D_r(t)r]'$ on the diagonal, $MRS_r(t)$ is strictly increasing up to $\widetilde{r}_t$ and becomes constant beyond the threshold. Thus, Increasing MRS holds. It is clear that a stream x is magnitude sensitive wrt period $t > 0$ iff $u(x_t) \leq \widetilde{r}_t$. Thus, the necessity of X_m-Regularity follows.

We turn to the sufficiency. Denote $R_m(t) = \{r \mid r = u(x) \text{ for some } x \in X_m(t)\}$, where the constant stream x is identified with a single period consumption x.

LEMMA S6: *$R_m(t)$ is an nonempty interval with* $\inf R_m(t) = 0$.

PROOF: By X_m-Regularity (i), for any $x \in X$ on the diagonal, there is α such that $\alpha x \in X_m(t)$. Therefore $X_m(t)$ and $R_m(t)$ are nonempty. We show that $R_m(t)$ is an interval: Take any $r \in R_m(t)$. There exists $x \in X_m(t)$ with $r = u(x)$. By X_m-Regularity (ii), $\alpha x \in X_m(t)$ for all $\alpha \in (0, 1)$. Thus, $u(\alpha x) \in R_m(t)$ for all $\alpha \in (0, 1)$, which means that $R_m(t)$ is an interval with $\inf R_m(t) = 0$. *Q.E.D.*

Denote $\widetilde{r}_t = \sup R_m(t)$. Together with Lemma 8 of NT, X_m-Regularity implies that for all $\widetilde{r}_t \geq u(x_t) > u(y_t) \geq 0$,

$$[D_r(t)r]'|_{r=u(x_t)} > [D_r(t)r]'|_{r=u(y_t)}. \tag{S16}$$

LEMMA S7: *For any $u(x_t) \geq u(y_t) > \widetilde{r}_t$,*

$$[D_r(t)r]'|_{r=u(x_t)} = [D_r(t)r]'|_{r=u(y_t)}.$$

PROOF: Take any $r = u(c) > \widetilde{r}_t$. Define $A(c, t) = \{\alpha \in (0, 1] \mid MRS_{\alpha c}(t) = MRS_c(t)\}$. We have the following observations about $A(c, t)$: (i) since $c \notin X_m(t)$, $A(c, t) \neq \emptyset$, and (ii) by Increasing MRS, $A(c, t)$ is convex, and hence it is an interval with $\sup A(c, t) = 1$.

Let $\underline{\alpha} = \inf A(c, t)$. Note that from (S16), $MRS_{\alpha c}(t) < MRS_c(t)$ for any $u(\alpha c) < \widetilde{r}_t$. In other words, for any $\alpha < \frac{u^{-1}(\widetilde{r}_t)}{c} = \frac{u^{-1}(\widetilde{r}_t)}{u^{-1}(r)} := \alpha^*$, $\alpha \notin A(c, t)$. Hence, $\alpha^* \leq \underline{\alpha}$.

We will claim that $\alpha^* = \underline{\alpha}$. By seeking a contradiction, suppose $\alpha^* < \underline{\alpha}$. Since $MRS_r(t)$ is continuous, $MRS_{\underline{\alpha}c}(t) = MRS_c(t)$. On the other hand, $\alpha^* < \underline{\alpha}$ implies $\widetilde{r}_t < u(\underline{\alpha}c)$. Since $\underline{\alpha}c \notin X_m(t)$, there exists $\beta \in (0, 1)$ such that $MRS_{\beta\underline{\alpha}c}(t) = MRS_{\underline{\alpha}c}(t)$, which contradicts to the fact that $\underline{\alpha}$ is an infimum of $A(c, t)$.

Now take any $s = u(c') \in (\widetilde{r}_t, r)$. By the above claim, $u(c') > \widetilde{r}_t = u(\alpha^* c) = u(\underline{\alpha}c)$. There exists some $\hat{\alpha} \in (\underline{\alpha}, 1)$ such that $u(c') = u(\hat{\alpha}c)$. Since $A(c, t) = [\underline{\alpha}, 1]$, we have $\hat{\alpha} \in A(c, t)$, and hence, $MRS_{\hat{\alpha}c}(t) = MRS_c(t)$. By Lemma 8 of NT, we have the desired result. *Q.E.D.*

Define $f_t(r) := D_r(t)r$. Together with Lemma 8 of NT, Condition (S16) and Lemma S7 imply that $f_t'(r)$ is strictly increasing on $[0, \widetilde{r}_t]$ and is constant on $(\widetilde{r}_t, \infty)$, respectively. Thus, f_t is strictly convex on $[0, \widetilde{r}_t]$ and is affine otherwise. By the same argument as in Lemma 10 of NT, $D_r(t)$ is strictly increasing on $[0, \widetilde{r}_t]$. *Q.E.D.*

S7.5. *Proof of Theorem S3*

First, we verify necessity. By Corollary S1, $D_r(t) = \kappa_t r^{\frac{1}{m-1}}$ if $r \leq \widetilde{r}_t$, and $D_r(t) = \overline{d}_t - \frac{A_t}{r}$ otherwise. Hence, $D_r(t)r = \kappa_t r^{\frac{m}{m-1}}$ if $r \leq \widetilde{r}_t$, and $D_r(t)r = \overline{d}_t r - A_t$ otherwise. Therefore, $[D_r(t)r]' = \frac{m}{m-1}\kappa_t r^{\frac{1}{m-1}}$ if $r \leq \widetilde{r}_t$, and $[D_r(t)r]' = \overline{d}_t$ otherwise. Since $MRS_r(t) = [D_r(t)r]'$ on the diagonal, $MRS_r(t)$ is strictly increasing up to $\widetilde{r}_t$ and becomes constant beyond the threshold. Thus, Increasing MRS holds. It is clear that a stream x is magnitude sensitive wrt period $t > 0$ iff $u(x_t) \leq \widetilde{r}_t$. Thus, the necessity of X_m-Regularity follows. Next, since $MRS_x(t) = \frac{u'(x_t)}{u'(x_0)} \frac{\partial D_r(t)r}{\partial r}|_{r=u(x_t)}$ in the GDU model (by Lemma 8 of NT) and since $D_r(t)r$ and u are homogeneous on $X_m(t)$ and C, respectively, it follows by Euler's theorem that $\frac{u'(x_t)}{u'(x_0)} \frac{\partial D_r(t)r}{\partial r}$ is homogeneous as well on $X_m(t)$, and in turn that $MRS_x(t)$ is homogeneous along the diagonal and the x_0-axis. This confirms MRS-Homogeneity.

We turn to the sufficiency. By Proposition S5, $[D_r(t)r]' = MRS_r(t)$ is strictly increasing on $[0, \widetilde{r}_t]$ and is constant otherwise.

LEMMA S8: *$D_r(t)$ is an affine function on $(\widetilde{r}_t, \infty)$.*

PROOF: Since $[D_r(t)r]'$ is constant on this domain, $D_r(t)r$ is an affine function and is written as $D_r(t)r = \overline{d}_t r - A_t$ for some $\overline{d}_t$ and A_t. Hence, for all $r > \widetilde{r}_t$, $D_r(t) = \overline{d}_t - \frac{A_t}{r}$. To derive $\overline{d}_t$ and A_t more explicitly, denote $f_t(r) = D_r(t)r$ for notational convenience. Since $f_t(r)$ is differentiable at $r = \widetilde{r}_t$, we have $\overline{d}_t = f_t'(\widetilde{r}_t)$. Since $U_t(c)$ is continuous, $f_t(\widetilde{r}_t) = \lim_{r\searrow \widetilde{r}_t} f_t(r) = \lim_{r\searrow \widetilde{r}_t}(\overline{d}_t r - A_t) = \overline{d}_t \widetilde{r}_t - A_t$. That is, $A_t = \overline{d}_t \widetilde{r}_t - f_t(\widetilde{r}_t) = f_t'(\widetilde{r}_t)\widetilde{r}_t - f_t(\widetilde{r}_t)$.

Take any consumption c with $r = u(c) > \widetilde{r}_t$. By Impatience, $r = U(c) \geq U(c^t) = \overline{d}_t r - A_t$. Since $1 \geq \overline{d}_t - \frac{A_t}{r}$ for all $r > \widetilde{r}_t$, we have $1 \geq \overline{d}_t$ as $r \to \infty$. *Q.E.D.*

LEMMA S9: *$D_r(t)$ and $u(c)$ are homogeneous on $[0, \widetilde{r}_t]$ and on $\mathbb{R}_+$, respectively.*

PROOF: By Lemma 8 of NT, on the diagonal, $MRS_x(t) = \frac{\partial D_r(t)r}{\partial r}|_{r=u(x_t)}$, and so $\int_0^1 MRS_{\alpha x}(t)\,\mathrm{d}\alpha = D_{u(x_t)}(t)u(x_t)$. By MRS-Homogeneity (i),

$$D_{u(x_t)}(t)u(x_t) = \int_0^1 MRS_{\alpha x}(t)\,\mathrm{d}\alpha = \int_0^1 \beta(\alpha) MRS_x(t)\,\mathrm{d}\alpha$$

$$= MRS_x(t) \times \int_0^1 \beta(\alpha)\,\mathrm{d}\alpha = k \times MRS_x(t), \tag{S17}$$

where $k := \int_0^1 \beta(\alpha)\,\mathrm{d}\alpha$. It follows that $D_{u(x_t)}(t)u(x_t)$ is homogeneous because

$$D_{u(\alpha x_t)}(t)u(\alpha x_t) = k \times MRS_{\alpha x}(t) = k\beta(\alpha) \times MRS_x(t). \tag{S18}$$

(Homogeneity of $D_{u(x_t)}(t)u(x_t)$ could alternatively be obtained using Euler's theorem, since it is the integral of a homogeneous function). As a consequence, we see that

$$D_{u(\alpha x_t)}(t) = \frac{\beta(\alpha)u(x_t)}{u(\alpha x_t)} \times D_{u(x_t)}(t)$$

since $D_{u(\alpha x_t)}(t)u(\alpha x_t) = \beta(\alpha) \times D_{u(x_t)}(t)u(x_t)$ by (S17) and (S18). The proof is complete once we establish that u is homogeneous since then $\frac{u(x_t)}{u(\alpha x_t)}$ is independent of x_t.

Recall that Lemma 8 of NT yields

$$MRS_x(t) = \frac{1}{u'(x_0)}\left[\frac{\partial D_{u(x_t)}(t)}{\partial r}u'(x_t)u(x_t) + D_{u(x_t)}(t)u'(x_t)\right],$$

where the term in the square bracket does not depend on x_0. Since MRS is homogeneous in x_0 by MRS-Homogeneity (ii), it follows that $u'(x_0)$ is a homogeneous function. By Euler's theorem, $u(x_0)$ must be homogeneous, as desired. *Q.E.D.*

Since $D_r(t)$ is homogeneous on $[0, \widetilde{r}_t]$, by Theorem 2 of NT, there exist $\theta > 0$ and $\kappa_t > 0$ such that

$$D_r(t) = \kappa_t r^\theta,$$

which is a strictly increasing function from $[0, \widetilde{r}_t]$ onto $[0, \tilde{d}_t]$, where $\widetilde{d}_t = \kappa_t \widetilde{r}_t^\theta$.

Therefore, we fully characterize the discount function as follows:

$$D_r(t) = \begin{cases} \kappa_t r^\theta & \text{if } 0 \le r \le \widetilde{r}_t, \\ \overline{d}_t - \dfrac{A_t}{r} & \text{if } r > \widetilde{r}_t. \end{cases}$$

Note that since $f_t(r) = D_r(t)r = \kappa_t r^{\theta+1}$ on $r \le \widetilde{r}_t$, $\overline{d}_t = f_t'(\widetilde{r}_t) = (\theta+1)\kappa_t \widetilde{r}_t^\theta$ and $A_t = f_t'(\widetilde{r}_t)\widetilde{r}_t - f_t(\widetilde{r}_t) = \theta\kappa_t \widetilde{r}_t^{\theta+1} > 0$.

Since the above $D_r(t)$ exhibits MDI, $\succsim$ is a CE representation. We show that φ_t of this model has an explicit form as below.

LEMMA S10: *There exist $a_t > 0$, $m > 1$ such that $\varphi_t(d) = a_t d^m$ for all $d \le \widetilde{d}_t$, $\varphi_t(d) = -A_t \ln(\overline{d}_t - d) + C_t$ for all $d \in (\widetilde{d}_t, \overline{d}_t)$, and $\varphi_t(d) = \infty$ for all $d \ge \overline{d}_t$. Moreover, $\widetilde{d}_t = \frac{m-1}{m}\overline{d}_t$ and $a_{t+1} \ge a_t$.*

PROOF: Recall $\widetilde{d}_t = \kappa_t \widetilde{r}_t^\theta$. Take any $u(c) = r \le \widetilde{r}_t$. By the same argument as in Lemma 7 of NT, there exist $m = \frac{1+\theta}{\theta} > 1$ and $a_t = 1/(m\kappa_t^{\frac{1}{\theta}}) > 0$ such that $\varphi_t(d) = a_t d^m$ for all $d \le \widetilde{d}_t$. Note also that $\overline{d}_t = (\theta+1)\kappa_t \widetilde{r}_t^\theta = \frac{m}{m-1}\widetilde{d}_t$, or $\widetilde{d}_t = \frac{m-1}{m}\overline{d}_t$.

Next, take any $u(c) = r > \widetilde{r}_t$. From the FOC of the cognitive optimization problem, it must be that $u(c) = \varphi_t'(\overline{d}_t - \frac{A_t}{u(c)})$, which implies that by setting $d = \overline{d}_t - \frac{A_t}{u(c)}$, we have $d \in (\widetilde{d}_t, \overline{d}_t)$ and

$$\varphi_t'(d) = \frac{A_t}{\overline{d}_t - d} > 0.$$

Since $\theta = \frac{1}{m-1}$ and $\widetilde{r}_t = a_t m(\widetilde{d}_t)^{\frac{1}{\theta}}$,

$$A_t = \theta\kappa_t \widetilde{r}_t^{\theta+1} = \frac{\widetilde{d}_t}{m-1}\widetilde{r}_t = \frac{a_t m}{m-1}(\widetilde{d}_t)^m = a_t\left(\frac{m-1}{m}\right)^{m-1}\overline{d}_t^m.$$

We have

$$\varphi_t(d) = -A_t \ln(\overline{d}_t - d) + C_t,$$

where $C_t \in \mathbb{R}$ is a constant. More explicitly, continuity requires that

$$C_t = \varphi_t(\widetilde{d}_t) + A_t \ln(\overline{d}_t - \widetilde{d}_t) = a_t(\widetilde{d}_t)^m + \frac{a_t m}{m-1}(\widetilde{d}_t)^m \ln\frac{\overline{d}_t}{m} = A_t\left(\frac{m-1}{m} + \ln\frac{\overline{d}_t}{m}\right).$$

Moreover, since $\overline{d}_t - \widetilde{d}_t = \theta\kappa_t \widetilde{r}_t^{\theta}$,

$$\lim_{d \searrow \widetilde{d}_t} \varphi_t'(d) = \frac{A_t}{\overline{d}_t - \widetilde{d}_t} = \frac{\theta\kappa_t \widetilde{r}_t^{\theta+1}}{\theta\kappa_t \widetilde{r}_t^{\theta}} = \widetilde{r}_t = a_t m(\widetilde{d}_t)^{m-1} = \lim_{d \nearrow \widetilde{d}_t} \varphi_t'(d),$$

which implies that φ_t is differentiable at $\widetilde{d}_t$.

Note that φ_t is strictly increasing and strictly convex on $d \in (\widetilde{d}_t, \overline{d}_t)$ and diverges to infinity as $d \to \overline{d}_t$. Thus, we can set $\varphi_t(d) = \infty$ for all $d \in [\overline{d}_t, 1]$. *Q.E.D.*

S7.6. *Proof of Theorem S4*

Since $u(C)$ is unbounded above by homogeneity of u, φ_t is automatically maximal. From Theorem 4 of NT, we have already shown that there exists $\lambda > 0$ such that $u_2 = \lambda u_1$ and $\varphi_t^2 = \lambda \varphi_t^1$. In particular, $\overline{d}_t^1 = \overline{d}_t^2 = \overline{d}_t$. Thus, $a_t^2 d^{m^2} = \lambda a_t^1 d^{m^1}$ for all $d \leq \min[\frac{m_1-1}{m_1}\overline{d}_t, \frac{m_2-1}{m_2}\overline{d}_t,]$. Note that $d^{m^1-m^2}$ is constant and equal to $\frac{a_t^2}{\lambda a_t^1}$ for all such d, which happens only when $m^1 = m^2$. Consequently, $a_t^2 = \lambda a_t^1$, as desired.

REFERENCES

DEBREU, G. (1960): "Topological Methods in Cardinal Utility Theory," in *Mathematical Methods in the Social Sciences*, ed. by K. Arrow, S. Karlin, and P. Suppes. Stanford University Press. [2,3]

LI, Y.-C., AND C.-C. YEH (2010): "Some Characterization of Convex Functions," *Computers and Mathematics with Applications*, 59, 327–337. [13]

NOOR, J., AND N. TAKEOKA (2022): "Optimal Discounting," *Econometrica*, 90 (2), 585–623. [1,9,10,13,14]

YAARI, M. (1977): "A Note on Separability and Quasiconcavity," *Econometrica*, 45, 1183–1186. [13]

Co-editor Itzhak Gilboa handled this manuscript.

Manuscript received 26 January, 2018; final version accepted 26 June, 2021; available online 29 June, 2021.