Supplementary Appendix for
“Optimal Discounting” *

Jawwad Noor Norio Takeoka

May 12, 2020

Abstract

This supplementary appendix for Noor and Takeoka [3] provides (i) a characterization for the smooth General Discounted Utility model (ii) a characterization of the homogeneous CE model using the MRS approach and (iii) the omitted proofs for results in Noor and Takeoka [3].

1 Introduction

In this supplementary appendix for Noor and Takeoka [3] (henceforth NT), we provide the omitted proofs for results in NT as well as two new results: a characterization of the smooth GDU model (Section 3) and a characterization of the homogeneous CE model using the MRS approach (Section 5).

Recall that there are $T + 1 < \infty$ periods starting with period 0, the space $C$ of outcomes is assumed to be $C = \mathbb{R}_+$, and $\Delta$ denotes the set of simple lotteries over $C$, with generic elements $p, q, \ldots$. The space of consumption streams is $X = \Delta^{T+1}$. It is endowed with the product topology and $x = (x_0, x_1, \ldots, x_T)$ is a generic stream. The primitive of our model is a preference $\succsim$ over $X$.

Let $\Delta_0 \subset X$ denote the set of streams $x = (p, 0, \ldots, 0)$ 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 $(p, 0, \ldots, 0) \in \Delta_0$. Thus, 0 also denotes the stream $(0, \ldots, 0)$. An element of $\Delta_0$ that is a mixture between two consumption alternatives $p, q \in \Delta_0$ 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:

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

*Noor (the corresponding author) is at 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.
2 GDU Representations

NT state that

Theorem 1 (NT [3, Theorem 2]) A preference \( \succsim \) over \( X \) satisfies Regularity and Separability if and only if it admits a GDU representation.

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

\[
D_{u^1(p)}(t) = D_{u^2(p)}(t).
\]

We provide the omitted proof here.

The necessity of Regularity and Separability is straightforward to establish. We establish its sufficiency in the following lemmas.

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, homogeneous (that is, \( u(\alpha \circ p) = \alpha u(p) \) 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. If \( \alpha \in [0,1] \), by mixture linearity of \( u \), together with identifying \( \alpha \circ p \) with \( \alpha \circ p + (1-\alpha) \circ 0 \),

\[
u(\alpha \circ p) = u(\alpha \circ p + (1-\alpha) \circ 0) = \alpha u(p) + (1-\alpha)u(0) = \alpha u(p).
\]

If \( \alpha > 1 \), 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.

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 \( p^* \to 0 \) 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 \( \overline{x} > \lambda(\overline{x}) \), \( \overline{x} \) belongs to the set \( W = \{ x \in X | \overline{x} \circ p^* > x \geq 0 \} \). 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 \([0,\overline{x}]\), there exists a convergent subsequence \( \{ \lambda(x^m)_i \} \) with a limit \( \lambda \neq \lambda(\overline{x}) \). On the other hand, since \( x^i \to \overline{x} \) and \( x^i \sim \lambda(x^i) \circ p^* \), Continuity implies \( \overline{x} \sim \lambda \circ p^* \). Since \( \lambda(\overline{x}) \) is unique, \( \lambda(\overline{x}) = \lambda \), which is a contradiction. ■
Lemma 2 U can be written as an additively separable utility form, i.e. \( U : \mathbb{X} \to \mathbb{R}^+ \) s.t. for all \( x \in \mathbb{X} \),
\[
U(x) = u(x_0) + \sum_{t \geq 1} U_t(x_t),
\]
where \( u \) is given as in Lemma 1 and \( U_t : \Delta \to \mathbb{R}^+ \) is continuous with \( U_t(0) = 0 \) for each \( t \). Moreover, \( u \) is unbounded from above.

Proof. Take any \( x = (x_0, x_1, \cdots, x_T) \in \mathbb{X} \) s.t. \( x > 0 \). By Monotonicity, there exists some \( t > 0 \) with \( x_t > 0 \). We start with the case where there are two \( x_t, x_s > 0 \). By notational convenience, denote such a stream by \((x_t, x_s, 0, \cdots, 0)\). By Separability,
\[
\frac{1}{2} \circ c_{(0,0,0,\cdots,0)} + \frac{1}{2} \circ c_{(x_t,0,\cdots,0)} \sim \frac{1}{2} \circ c_{(x_t,0,\cdots,0)} + \frac{1}{2} \circ 0.
\]
Since \( u \) is mixture linear,
\[
 u(c_{(0,0,0,\cdots,0)}) + u(c_{(x_t,0,\cdots,0)}) = u(c_{(x_t,0,\cdots,0)}) + u(0)
\]
\[\Longleftrightarrow U(0, x_0, 0, \cdots, 0) + U(x_t, 0, \cdots, 0) = U(x_t, 0, x_s, 0, \cdots, 0). \]
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
\[
U(x_t, 0, x_s, 0, \cdots, 0) = U_t(x_t) + U_s(x_s). \tag{1}
\]
In particular, if \( t = 0 \), \( U_t(x_t) = u(x_t) \).
If a stream has three outcomes \( x_t, x_s, x_r > 0 \), denote it by \((x_t, x_s, x_r, 0, \cdots, 0)\). From the above argument, we have (1). By Separability,
\[
\frac{1}{2} \circ c_{(0,0,0,\cdots,0)} + \frac{1}{2} \circ c_{(x_t,x_t,0,\cdots,0)} \sim \frac{1}{2} \circ c_{(x_t,x_t,0,\cdots,0)} + \frac{1}{2} \circ 0.
\]
Since \( u \) is mixture linear,
\[
 u(c_{(0,0,0,\cdots,0)}) + u(c_{(x_t,x_t,0,\cdots,0)}) = u(c_{(x_t,x_t,0,\cdots,0)}) + u(0)
\]
\[\Longleftrightarrow U(0, 0, x_r, 0, \cdots, 0) + U(x_t, x_s, 0, \cdots, 0) = U(x_t, x_s, x_r, 0, \cdots, 0). \]
Define \( U_r(x_r) = U(0, 0, x_r, 0, \cdots, 0) \). Then, we have
\[
U(x_t, x_s, x_r, 0, \cdots, 0) = U_r(x_r) + U_t(x_t) + U_s(x_s) + U_r(x_r).
\]
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 \( U_t(x_t) = U(0, \cdots, 0, x_t, 0, \cdots, 0) \). By definition, \( U_t(0) = 0 \). Since \( U \) is continuous, \( U_t \) is also continuous.
Finally, we show that \( u \) must be 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 Monotonicity, there exists some \( \tilde{c} \in C \) with \( U_t(\tilde{c}) > 0 \). Take a consumption \( \overline{v} \in C \) such that \( \overline{v} \geq \tilde{c} \) and \( \overline{v} - u(\overline{v}) < \sum_{t \geq 1} U_t(\tilde{c}) \). Consider the stream \( \pi = (\overline{v}, \ldots, \overline{v}) \). By the representation,

\[
U(\pi) = u(\overline{v}) + \sum_{t \geq 1} U_t(\overline{v}) \geq u(\overline{v}) + \sum_{t \geq 1} U_t(\tilde{c}) > \overline{v}.
\]

Since \( \overline{v} \) is the supremum of \( u(C) \), the above inequality contradicts to the Present Equivalents axiom. ■

**Lemma 3** The function \( U : X \rightarrow \mathbb{R}_+ \) defined as in Lemma 2 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 2, 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 \( D_t \) 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

\[
U(x) = u(x_0) + \sum_{t \geq 1} D_u(x_t)(t)u(x_t), \text{ for all } x \in X,
\]

with convention that \( D_u(x_0)(t)u(x_0) = 0 \) whenever \( u(x_0) = 0 \).

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

Finally we establish uniqueness.

**Lemma 4** If \( GDU(u^i, D^i) \) for \( i = 1, 2 \) both represent the same preference \( \succeq \), then there exists a scalar \( \beta > 0 \) s.t. for all \( p \in \Delta \) and \( t > 0 \),

\[
u^1(p) = \beta u^2(p) \text{ and } D^1_{u^1(p)}(t) = D^2_{u^2(p)}(t).
\]

**Proof.** By considering the restriction \( \succeq \big|_{\Delta_0} \) and applying the Mixture Space Theorem, we obtain \( \beta > 0 \) and \( \gamma \) s.t. \( u^1 = \beta u^2 + \gamma \). Due to the normalization \( u(0) = 0 \) in the representation, we have \( \gamma = 0 \). Thus \( u^1 = \beta u^2 \).

Next, observe that there exists a present equivalent \( c_{p^t} \) for each dated \( p^t \), 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 \( u^1 = \beta u^2 \), 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. ■

4
3 Smooth GDU Representations

NT introduce the notion of a General Discounted Utility Representation and assume differentiability of the discount function $D_r(t)$ in payoffs $r$ in order to identify the magnitude effects in terms of the change in marginal rate of intertemporal substitution. This section provides an axiomatic foundation to guarantee the existence of such smooth representations. Recall the definition of General Discounted Utility Representations and their smooth counterparts:

**Definition 1 (General Discounted Utility Representation)** A General Discounted Utility (GDU) representation is a tuple $(u; D_r)$ where $u : \Delta \to \mathbb{R}_+$ is continuous and mixture linear with $u(0) = 0$ and $D_r(t) \in [0, 1]$ is a magnitude-dependent discounting function which is weakly decreasing in $t$ and continuous for all $r > 0$, such that $\succsim$ is represented by the function $U : X \to \mathbb{R}_+$ defined by

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

A GDU representation $(u, D)$ is smooth if $u(c)$ and $D_r(t)$ are differentiable in $c$ and $r$ respectively.

NT show that a preference $\succsim$ over $X$ satisfies Regularity and Separability if and only if $\succsim$ admits a GDU representation.

Taking into account the marginal rate of substitution of preference, we require two kinds of differentiability conditions on preference. The first axiom requires smoothness of the present equivalent for deterministic streams.

**Axiom 1 (Smooth Present Equivalent)** A present equivalent function $c : C^{T+1} \to C$ is differentiable.

Since $c_x$ is uniquely determined for all streams $x$, a present equivalent function $c : X \to C$ is well-defined. The function $c : C^{T+1} \to C$ is its restriction on the deterministic streams. Since $C = \mathbb{R}_+$ and $C^{T+1} = \mathbb{R}_+^{T+1}$, the differentiability of $c$ should be understood as standard differentiability on $\mathbb{R}_+^{T+1}$.

The next axiom requires smoothness of VNM function on $C$. Here we adopt a version of Nakamura [2].

**Axiom 2 (Smooth Curvature)** $\succsim|_{\Delta_0}$ satisfies (1) risk aversion, that is, for all $c, c' \in C$ and $\alpha \in (0, 1)$, $\alpha c + (1 - \alpha)c' \succsim \alpha \circ c + (1 - \alpha) \circ c'$, and (2) Axiom A2 (Preference for Small Positive Risk Taking) of Nakamura [2].

We have already assumed $C$-Monotonicity, that is, $c > c'$ implies $c > c'$. Nakamura [2] shows that a strictly increasing and concave VNM function is differentiable if and only if it satisfies the Preference for Small Positive Risk Taking axiom.

Then we can show that:
Proposition 1 \( \succeq \) satisfies Regularity, Separability, Smooth Present Equivalent, and Smooth Curvature if and only if \( \succeq \) admits a GDU representation such that \( u(c) \) and \( D_r(t) \) are differentiable in \( c \) and \( r \) respectively.

Proof. The proof of sufficiency builds on the proof of Theorem 1 above. As in Lemma 1 above, by Regularity, \( \succeq |_{\partial \Delta} \) satisfies the vNM axioms. There exists a continuous mixture linear function \( u : \Delta \to \mathbb{R}_+ \) which represents \( \succeq |_{\partial \Delta} \) and which can be chosen so that \( u(0) = 0 \). Moreover, if \( \succeq \) satisfies Smooth Curvature, by Nakamura [2], \( u : C \to \mathbb{R}_+ \) can be taken to be differentiable.

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 \( \succeq \). Moreover, for all \( p \in \Delta \), \( U(p) = u(p) \). In particular, we have \( U(0) = u(0) = 0 \). Take any deterministic streams \( x \in C^{T+1} \). If \( \succeq \) satisfies Smooth Present Equivalent and Smooth Curvature, then both \( c_x \) and \( u : C \to \mathbb{R}_+ \) are differentiable, and hence \( U(x) = u(c_x) \) is also differentiable. That is, \( U(x) \) is differentiable on \( C^{T+1} \).

Since \( \succeq \) satisfies Separability on the whole domain, by applying Lemma 2 above, 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. Since \( U_t \) is differentiable on \( C^{T+1} \), so is \( U_t \) on \( C \). By the same argument as in Lemma 2 above, we can show that \( u \) is unbounded from above, that is, \( u(C) = \mathbb{R}_+ \).

Given the above additive representation, by the Monotonicity property contained in the Regularity axiom, 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

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

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

Note that since \( u \) and \( U_t \) are continuous, so is \( D_{u(c)}(t) \) in \( u(c) \) on the domain of \( u(c) > 0 \). Since \( u \) and \( U_t \) are differentiable on \( C \),

\[
D_{u(c)}(t) = \frac{U_t(u(c))}{u(c)}
\]

is differentiable on \( C \). Since \( u(C) = \mathbb{R}_+ \), \( D_r(t) \) is also differentiable in \( r \).

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 \( t \), as desired.

For necessity, consider the representation with differentiable \( u(c) \) and \( D_r(t) \). It is enough to verify Smooth Curvature and Smooth Present Equivalents. The necessity of the former
is proved by Nakamura [2]. By representation, $U(x)$ is differentiable on $C^{T+1}$. Thus, on $C^{T+1}$, the present equivalent function $c_x = u^{-1}(U(x))$ is differentiable. ■

4 Homogeneous CE Representation: Necessity

NT state that

Theorem 2 (NT [3, Theorem 7]) A regular CE preference $\succsim$ over $X$ satisfies $X_\ell$-Homogeneity if and only if $\succsim$ admits a homogeneous CE representation.

We present here the omitted proof of necessity:

Proof. We check whether $X_\ell$-Homogeneity holds. By NT [3, Proposition 2], the reduced form of a homogeneous CE representation is given as

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

where

$$D_{x_t}(t) = \begin{cases} \left( \frac{u(x_t)}{ma_t} \right)^{\frac{1}{m-1}} & \text{if } u(x_t) \leq \frac{m}{d_t} m a_t, \\ \frac{1}{d_t} & \text{if } u(x_t) > \frac{m}{d_t} m a_t. \end{cases}$$

Therefore,

$$X_\ell = \{ x \in X : \text{there exists some } t \geq 1 \text{ such that } 0 < u(x_t) \leq \frac{m}{d_t} m a_t \}.$$ 

To show $X_\ell$-Homogeneity, take any $x \in X_\ell$ with $u(x_0) = 0$. Then, $\beta \circ c_x \sim \alpha x$ if and only if

$$\beta \sum_{t \geq 1} D_{x_t}(t)u(x_t) = \sum_{t \geq 1} D_{\alpha \circ x_t}(t)u(\alpha \circ x_t)$$

$$\iff \beta \sum_{t \geq 1} \left( \frac{u(x_t)}{ma_t} \right)^{\frac{1}{m-1}} u(x_t) = \sum_{t \geq 1} \left( \frac{u(\alpha \circ x_t)}{ma_t} \right)^{\frac{1}{m-1}} u(\alpha \circ x_t)$$

$$\iff \beta \sum_{t \geq 1} \left( \frac{u(x_t)}{ma_t} \right)^{\frac{1}{m-1}} u(x_t) = \alpha^{\frac{m}{m-1}} \sum_{t \geq 1} \left( \frac{u(x_t)}{ma_t} \right)^{\frac{1}{m-1}} u(x_t).$$

Thus, $\beta = \alpha^{\frac{m}{m-1}}$. Since $\beta$ is independent of $x$, we have established $X_\ell$-Homogeneity. ■

5 Smooth Homogeneous CE Representation: MRS Approach

While NT axiomatize the homogeneous CE model using the lottery approach, here we provide the counterpart using the MRS approach.
First we present the analog of ℓ-Magnitude Sensitivity defined in NT. We define it for deterministic streams (since MRS is defined only for such streams) and then extend the definition to streams of lotteries.

**Definition 2** A deterministic stream \( x \in C^{T+1} \) is Magnitude Sensitive if there exists \( t > 0 \) such that for all \( \varepsilon \in (0, x_t) \),

\[
MRS_{f(x)}(t) > MRS_{f(x-\varepsilon)}(t).
\]

An arbitrary stream \( x \in X \) is Magnitude Sensitive if there exists a Magnitude Sensitive deterministic stream \( (c_t)_{t=0}^{T} \in C^{T+1} \) such that \( c_t \sim x_t \) for all \( t \).

The set of all Magnitude Sensitive streams is denoted by \( X_m \subset X \).

Since Magnitude Sensitivity is defined using \( t > 0 \), Magnitude Sensitivity of a stream \( x \) does not rely on the immediate consumption \( x_0 \) offered. Moreover, a stream that offers only 0 consumption in the future cannot be Magnitude Sensitive. When the stream \( x \) is a dated reward at \( t > 0 \), that is, \( x = c_t \), then \( x \) is Magnitude Sensitive iff for this \( t \) and for all \( \varepsilon \in (0, c) \), the agent exhibits

\[
MRS_{f(x)}(t) > MRS_{f(x-\varepsilon)}(t).
\]

Consider the counterparts of \( X_\ell \)-Regularity and \( X_\ell \)-Homogeneity axioms in NT that replace \( X_\ell \) with \( X_m \):

**Axiom 3 (\( X_m \)-Regularity)** For any deterministic \( c \in C \) and \( t > 0 \),

(i) if \( (c_t)^t \notin X_m \) then \( (c'^t) \in X_m \) for some \( 0 < c' < c \).

(ii) if \( (c_t)^t \in X_m \) then \( (c'^t) \in X_m \) for any \( 0 < c' < c \).

**Axiom 4 (\( X_m \)-Homogeneity)** For any \( x, y \in X_m \) s.t. \( x_0 \sim y_0 \sim 0 \) and any \( \alpha, \beta \in (0, 1) \),

\[
\beta \circ c_x \sim \alpha x \implies \beta \circ c_y \sim \alpha y.
\]

### 5.1 Representation Theorem

Define a smooth homogeneous CE representation as a CE representation where \( u \) is differentiable and the cost functions \( \{\varphi_t\} \) are defined as follows: there exist parameters \( m > 1 \), \( a_t > 0 \) increasing in \( t \geq 1 \), \( 0 < \bar{d}_t \leq \frac{m-1}{m} \) decreasing in \( t \geq 1 \), such that \( \{\varphi_t\} \) is given by

\[
\varphi_t(d) = \begin{cases} 
    a_t \cdot d^m & \text{if } d \in [0, \bar{d}_t], \\
    a_t \log(\bar{D}_t - d) + C_t & \text{if } d \in (\bar{d}_t, \bar{D}_t), \\
    \infty & \text{if } d \in [\bar{D}_t, 1],
\end{cases}
\]

where, in order to simplify exposition, we define the scalars

- \( \bar{D}_t := \frac{m}{m-1} \bar{d}_t \in (\bar{d}_t, 1), \)
- \( A_t := -\frac{a_t m}{m-1} \bar{d}_t^m < 0, \) and
- \( C_t := a_t \bar{d}_t^m (1 + \frac{m}{m-1} \log \frac{\bar{d}_t}{\bar{d}_t}) \in \mathbb{R}. \)

We can now state:
**Theorem 3** Suppose $\succsim$ over $X$ admits a smooth GDU representation. Then, $\succsim$ satisfies Weak Magnitude Effect, $X_m$-Regularity, and $X_m$-Homogeneity if and only if $\succsim$ admits a smooth homogeneous CE representation.

We show that by Weak Magnitude Effect and $X_m$-Regularity, we have $MRS_{f_t(x)}(t) > MRS_{f_t(x-\varepsilon)}(t)$ for any $x \in X_m$ and in particular, $D_r(t)$ is strictly increasing for all $r$ under some threshold. By $X_m$-Homogeneity on $X_m$, $D_r(t)$ admits a CRRA form for all such $r$. For any $x \notin X_m$, Weak Magnitude Effect and $X_m$-Regularity imply that $MRS_{f_t(x)}(t) = MRS_{f_t(x-\varepsilon)}(t)$ for all small $\varepsilon$. This condition is used to show that $D_r(t)$ must be a positive affine function for all sufficiently large payoffs. Thus the discount function of the GDU representation is given by

$$D_r(t) = \begin{cases} \gamma t^\theta & \text{if } r \leq \tau_t, \\ \frac{A_t}{r} & \text{if } r > \tau_t, \end{cases}$$

where $\theta > 0$, $\gamma > 0$, and $\tau_t > 0$ have particular restrictions. We show that this discount function would arise from a homogeneous CE-style model where the cost function has a CRRA form $\varphi_t(d) = a_t d^\alpha$ on some interval $[0, \overline{d}_t]$ and the form $\varphi_t(d) = A_t \log(\overline{D}_t - d) + C_t$ on an interval $(\overline{d}_t, \overline{D}_t)$. Since $D_r(t)$ never takes values on $[\overline{D}_t, 1]$, we set $\varphi_t(d) = \infty$.

The affine component of the cost function arises from the assumed smoothness of the GDU, which requires that $D_r(t)$ must be a differentiable function of $r$. With smoothness, $D_r(t)$ is strictly increasing on the whole domain and converges to the upper bound $\overline{D}_t$ as $r \to \infty$. Recall the reduced form of the “non-smooth” homogeneous CE representation in NT [3, Proposition 2], where the optimal discount function $D_r(t)$ first takes the above power form for $r \leq \tau_t$, but becomes constant at some constant $\overline{D}_t$ beyond $\tau_t$. This optimal discount function is not differentiable at $\tau_t$, and indeed corresponds to the case where $A_t = 0$. This is a consequence of homothetic property of the representation beyond the threshold, which requires a proportional attitude when payoffs are scaled down toward the origin.

### 5.2 Proof of Theorem 3

Denote $R_m(t) = \{r \mid r = u(p) \text{ for some } p \in X_m\}$. Define

$$D^m_t = \{d \in [0, 1] \mid d = D_u(p)(t) \text{ for some } p \in X_m\}.$$

By a symmetric argument as in NT [3, Lemma 3], $R_m(t)$ is an interval with $\inf R_m(t) = 0$, and $D^m_t$ is an interval.

Denote $\overline{d}_t = \sup D^m_t$ and $\tau_t = \sup R_m(t)$. Weak Magnitude Effect and $X_m$-Regularity imply that

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

for all $\tau_t \geq u(x_t) \geq u(y_t) > 0$, and

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

9
for all \( u(x_t) \geq u(y_t) > \tau_t \). If we define \( f_t(r) \equiv D_r(t)r \), the above condition means \( f_t'(r) \) is strictly increasing on \([0, \tau_t]\) and is constant on \((\tau_t, \infty)\), respectively. Thus, \( f_t \) is strictly convex on \([0, \tau_t]\) and is affine otherwise. By the same argument as in NT [3, Lemma 11], \( D_r(t) \) is strictly increasing on \([0, \tau_t]\). Moreover, by the same argument as in NT [3, Lemmas 6 and 7], there exists \( \theta > 0 \) and \( \tau_t > 0 \) such that

\[
D_r(t) = \gamma(t)^\theta,
\]

which is a strictly increasing function from \([0, \tau_t]\) onto \([0, \overline{d}_t]\), where \( \overline{d}_t = \gamma t^\theta \).

For all \( r > \tau_t \), \( [D_r(t)r]' = f_t'(r) \) is also constant, and hence written as \( f_t(r) = \overline{D}(t)r + \overline{A}_t \).

Since \( f_t(r) \) is differentiable at \( r = \tau_t \), we have \( \overline{D}(t) = f_t'(\tau_t) \). Since \( U_t(c) \) is continuous, \( f_t(\tau_t) = \lim_{r \to \tau_t} f_t(r) = \lim_{r \to \tau_t} (\overline{D}(t)r + \overline{A}_t) = \overline{D}(t)\tau_t + \overline{A}_t \). That is, \( \overline{A}_t = f_t(\tau_t) - \overline{D}(t)\tau_t = f_t(\tau_t) - f_t'(\tau_t)\tau_t < 0 \). That is, for all \( r > \tau_t \), \( D_r(t)r \) is a positive affine function as given by

\[
D_r(t)r = f_t'(\tau_t)r + f_t(\tau_t) - f_t'(\tau_t)\tau_t.
\]

Therefore, we fully characterize the discount function as follows :

\[
D_r(t) = \begin{cases} 
\gamma(t)^\theta \overline{D}(t) + \frac{\overline{A}_t}{r} & \text{if } 0 \leq r \leq \tau_t \\
\gamma(t)^{2 \theta} & \text{if } r > \tau_t 
\end{cases}
\]

Recall that when \( f_t(r) = D_r(t)r = \gamma(t)^{\theta + 1} \) on \( r \leq \tau_t \), \( \overline{D}(t) = f_t'(\tau_t) = (\theta + 1)\gamma(t)^\theta \) and \( A_t = f_t(\tau_t) - \tau_t f_t'(\tau_t) = -\theta \gamma(t)^{\theta + 1} < 0 \).

Finally, we show that \( \varphi_t \) of this model has an explicit form as below.

**Lemma 5** There exist \( a_t > 0 \), \( m > 1 \) such that \( \varphi_t(d) = a_id^m \) for all \( d \leq \overline{d}_t \), \( \varphi_t(d) = A_t \log(\overline{D}(t) - d) + C_t \) for all \( d \in (\overline{d}_t, \overline{D}(t)) \), and \( \varphi_t(d) = \infty \) for all \( d > \overline{D}(t) \). Moreover, \( a_{t+1} \geq a_t \).

**Proof.** Recall \( \overline{d}_t = \gamma(t)^\theta \). Note that \( \overline{D}(t) = f_t'(\tau_t) = (1 + \theta)\gamma(t)^\theta = (1 + \theta)\overline{d}_t > \overline{d}_t \).

Take any \( u(p) = r \leq \tau_t \). By the same argument as in NT [3, Lemma 8], there exist \( m = \frac{1 + \theta}{\theta} > 1 \) and \( a_t = \frac{1}{m\gamma^\theta} > 0 \) such that \( \varphi_t(d) = a_id^m \) for all \( d \leq \overline{d}_t \).

Next, take any \( u(p) = r > \tau_t \). From the FOC of the cognitive optimization problem it must be that \( u(p) = \varphi_t'(\overline{D}(t) + \frac{A_t}{u(p)}) \), which implies that by setting \( d = \overline{D}(t) + \frac{A_t}{u(p)} \in (\overline{d}_t, \overline{D}(t)) \),

\[
\varphi_t'(d) = \frac{-A_t}{\overline{D}(t) - d} > 0.
\]

Since \( \theta = \frac{1}{m-1} \) and \( \tau_t = a_t m \gamma^\theta \),

\[
A_t = -\theta \gamma(t)^{\theta + 1} = -\frac{d_t}{m-1} \tau_t = -\frac{a_t m}{m-1} \overline{d}_t < 0.
\]

We have

\[
\varphi_t(d) = A_t \log(\overline{D}(t) - d) + C_t.
\]
where \( C_t \in \mathbb{R} \) is a constant. More explicitly, continuity requires that
\[
C_t = \varphi_t(d_t) - A_t \log(D(t) - d_t) = a_t d_t^m + \frac{a_t m}{m-1} \log \frac{d_t}{m-1}.
\]

Note that \( \varphi_t \) is strictly increasing and strictly convex on \( d \in (d_t, D(t)) \) and diverges to infinity as \( d \to D(t) \). Note also that \( \frac{D(t)}{d_t} = (\theta + 1)\gamma t \gamma^\theta - \gamma t \gamma^\theta = \theta \gamma t \gamma^\theta \). Thus,
\[
\lim_{d \searrow d_t} \varphi_t'(d) = \frac{-A_t}{D(t) - d_t} = \frac{\theta \gamma t \gamma^\theta + 1}{\theta \gamma t \gamma^\theta} = r_t = \varphi_t'(d_t),
\]
which implies that \( \varphi_t \) is differentiable at \( d_t \).

Take any lottery \( p \) with \( r = u(p) > r_t \). By Impatience, \( r = U(p) \geq U(p') = D(t)r + A_t \). Since \( 1 \geq D(t) + \frac{A_t}{r} \) for all \( r > r_t \), we have \( 1 \geq D(t) \) as \( r \to \infty \). Hence,
\[
d_t = \frac{1}{\theta + 1} D(t) = \frac{m}{m-1} D(t) \leq \frac{m}{m-1},
\]
as desired. \( \blacksquare \)

We turn to necessity. Given the cost function of smooth homogeneous CE representation, an optimal discount function is given as (3). Since the second derivative of \( D_r(t)r \) is non-negative, MRS = \([D_r(t)r]' = (1 + \theta)\gamma t \gamma^\theta \), which is strictly increasing in \( r \), while if \( r > r_t \), MRS = \([D_r(t)r]' = D(t) \), which is constant in \( r \). Thus, \( X_m \)-Regularity holds. Since \( D_r(t) \) has a CRRA form on \( X_m \), \( X_m \)-homogeneity is shown by the same argument as in NT [3, Theorem 7] or in Section 4 of this supplementary appendix.

## 6 Quasi-Stationarity

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

**Axiom 5 (Quasi-Stationarity)** For any streams \( x, y \) such that \( x_0 = y_0 = x_T = y_T = 0 \) and any \( c \),
\[
x \succeq y \iff cx \succeq cy.
\]

NT state that

**Proposition 2 (NT [3, Proposition 4])** Suppose that \( T \geq 3 \). A homogeneous CE representation \((u, \{a_t d_t^m\}_{t \geq 1})\) 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
\[
a_t = \frac{1}{m\beta^m-1(\delta^m-1)^t}, \text{ for each } t,
\]
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: \( x \succsim^1 y \iff 0x \succsim 0y \). From the reduced form of the homogenous CE model as given in NT [3, Proposition 2], we see that \( \succsim^1 \) is represented by

\[
U^1(x) = \sum_{t \geq 1} \gamma(t)u(x_t)\frac{m}{m-t}, \quad \text{for all } x_t \leq u^{-1}(ma_t\bar{d}_t^{m-1}).
\]

Let \( \delta = \frac{\gamma^2}{\gamma(1)} \). By definition of a Homogeneous CE model, \( \gamma(1) \geq \gamma(2) > 0 \) and so we have \( 0 < \delta < 1 \). There exist sufficiently small but strictly positive \( c, c' \) such that \( U^1(c, 0, \cdots, 0) = U^1(0, c', 0, \cdots, 0) \), which is equivalent to \( \gamma(1)u(c)\frac{m}{m-t} = \gamma(2)u(c')\frac{m}{m-t} \). By Quasi-Stationarity, \( \gamma(2)u(c)\frac{m}{m-t} = \gamma(3)u(c')\frac{m}{m-t} \) and so \( \delta = \frac{\gamma(2)}{\gamma(1)} = \frac{\gamma(3)}{\gamma(2)} \). By repeating this argument, we have \( \frac{\gamma(t+1)}{\gamma(t)} = \delta \) for all \( t \geq 1 \), and so

\[
\gamma(t) = \delta^{t-1}\gamma(1) = \beta \delta^t,
\]

where \( \beta = \gamma(1)/\delta > 0 \). Since \( \gamma(t) = \left( \frac{1}{ma_t} \right) \frac{1}{m-t} \), we obtain

\[
a_t = \frac{1}{m\gamma(t)^{m-t}} = \frac{1}{m(\beta \delta^t)^{m-1}} = \frac{1}{m\beta^{m-1}(\delta^{m-1})^t}.
\]

The optimal discount function takes the form

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

for all \( u(c) \leq \left( \frac{\bar{a}_d}{\beta \delta^t} \right)^{m-1} \).

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

\[
\bar{d}_t = \delta^{t-1}\bar{d}_1.
\]

Moreover, let \( c^*_t \) be the threshold consumption at \( t \) satisfying \( u(c_t^*) = \left( \frac{\bar{a}_t}{\beta \delta^t} \right)^{m-1} \). Since \( \bar{d}_t, \beta, \delta > 0 \), it must be that \( c^*_t > 0 \). Then, \( \bar{d}_t = \beta \delta^t u(c^*_t) \frac{1}{m-t} > 0 \), and \( \bar{d} = \frac{\bar{a}_{t+1}}{\bar{a}_t} = \frac{\bar{d}_t}{\bar{d}_{t-1}} = \delta - \frac{\beta \delta^t u(c^*_t)}{\beta \delta^t u(c^*_t)} = \delta \).
By Monotonicity it must be that $c_t > 0$. Equivalently, $\frac{u(c_{t+1})}{u(c_t)} = \left(\frac{\delta}{\bar{\sigma}}\right)^m := \bar{\sigma} > 0$. We have therefore established that
\[
D_c(t) = \begin{cases} 
\beta \delta \mu(c) \frac{1}{\bar{\sigma}^{1/m}} u(c) & \text{if } u(c) \leq u(c_t), \\
\beta \delta \mu(c_t) \frac{1}{\bar{\sigma}} & \text{otherwise}.
\end{cases}
\]

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 $\bar{\sigma} = \frac{u(c_{t+1})}{u(c_t)} = 1$. Seeking a contradiction, first suppose $\bar{\sigma} > 1$. Since $u(c_{t+1}) = \bar{\sigma} u(c_t)$ 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 $\epsilon$ that satisfies
\[
U(0, c_2^* + \epsilon, 0, \ldots, 0) = U(0, c_2^*, 0, 0, \ldots, 0).
\]
By Monotonicity it must be that $\epsilon > 0$. Since $c_1^* < c_2^* < c_2^* + \epsilon$, we can compute that
\[
U(0, c_2^* + \epsilon, 0, \ldots, 0) = U(0, c_2^*, c_2^*, 0, 0, \ldots, 0)
\]
\[
\iff\quad D_{c_2^* + \epsilon}(1) u(c_2^* + \epsilon) = D_{c_2^*}(1) u(c_2^*) + D_{c_2}(2) u(c_2^*)
\]
\[
\iff\quad \beta \delta u(c_2^*) \frac{1}{\bar{\sigma}^{1/m}} u(c_2^* + \epsilon) = \beta \delta u(c_2^*) \frac{1}{\bar{\sigma}^{1/m}} u(c_2^*) + \beta \delta^2 u(c_2^*) \frac{1}{\bar{\sigma}} u(c_2^*)
\]
\[
\iff\quad \frac{u(c_2^* + \epsilon)}{u(c_2^*)} = 1 + \delta \left(\frac{u(c_2^*)}{u(c_2^*)}\right)^{1/m} = 1 + \delta.
\]
which yields our first expression for $\frac{u(c_2^* + \epsilon)}{u(c_2^*)}$. Given Quasi-Stationarity, we reason further that
\[
U(0, c_2^* + \epsilon, 0, \ldots, 0) = U(0, c_2^*, c_2^*, 0, 0, \ldots, 0)
\]
\[
\iff\quad U(0, 0, c_2^* + \epsilon, 0, \ldots, 0) = U(0, 0, c_2^*, c_2^*, 0, 0, \ldots, 0)
\]
\[
\iff\quad D_{c_2^* + \epsilon}(2) u(c_2^* + \epsilon) = D_{c_2^*}(2) u(c_2^*) + D_{c_2}(3) u(c_2^*)
\]
\[
\iff\quad \beta \delta^2 u(c_2^*) \frac{1}{\bar{\sigma}^{1/m}} u(c_2^* + \epsilon) = \beta \delta^2 u(c_2^*) \frac{1}{\bar{\sigma}^{1/m}} u(c_2^*) + \beta \delta^3 u(c_2^*) \frac{1}{\bar{\sigma}} u(c_2^*)
\]
\[
\iff\quad \frac{u(c_2^* + \epsilon)}{u(c_2^*)} = 1 + \delta.
\]
Therefore we obtain the second expression for $\frac{u(c_2^* + \epsilon)}{u(c_2^*)}$. Putting both together we see that
\[
1 + \delta \left(\frac{u(c_2^*)}{u(c_2^*)}\right)^{1/m} = 1 + \delta.
\]
Since all the terms are strictly positive and $m > 1$, we conclude that $\frac{u(c_2^*)}{u(c_2^*)} = 1$. But then $\bar{\sigma} = \frac{u(c_2^*)}{u(c_2^*)} = 1$, while we had supposed that $\bar{\sigma} > 1$, a contradiction.

Next suppose by way of contradiction that $\bar{\sigma} < 1$. Since it is the case that $u(c_{t+1}) = \bar{\sigma} u(c_t)$ for all $t$, it must be that $c_{t+1} < c_t$ for all $t$. Consider $c_1^*, c_2^*, c_3^*$ and take $\epsilon > 0$ that satisfies
\[
U(0, c_1^* + \epsilon, 0, \ldots, 0) = U(0, c_1^*, c_2^*, 0, 0, \ldots, 0),
\]
Since by hypothesis, \( c_3^* < c_2^* < c_1^* \), we can compute that

\[
U(0, c_1^* + \epsilon, 0, \cdots, 0) = U(0, c_1^*, c_3^*, 0, \cdots, 0)
\]

\[
\iff \quad D_{c_1^* + \epsilon}(1)u(c_1^* + \epsilon) = D_{c_3^*}(1)u(c_1^*) + D_{c_3^*}(2)u(c_3^*)
\]

\[
\iff \quad \beta \delta u(c_1^*) \frac{1}{\mu - 1} u(c_1^* + \epsilon) = \beta \delta u(c_1^*) \frac{1}{\mu - 1} u(c_1^*) + \beta \delta^2 u(c_3^*) \frac{1}{\mu - 1} u(c_3^*)
\]

\[
\iff \quad u(c_1^* + \epsilon) = u(c_1^*) + \delta \left( \frac{u(c_3^*)}{u(c_1^*)} \right)^{\frac{1}{\mu - 1}} u(c_3^*).
\]

By Quasi-Stationarity, we should also have

\[
U(0, 0, c_1^* + \epsilon, 0, \cdots, 0) = U(0, 0, c_1^*, c_3^*, 0, \cdots, 0)
\]

\[
\iff \quad D_{c_1^* + \epsilon}(2)u(c_1^* + \epsilon) = D_{c_1^*}(2)u(c_1^*) + D_{c_3^*}(3)u(c_3^*)
\]

\[
\iff \quad \beta \delta^2 u(c_2^*) \frac{1}{\mu - 1} u(c_1^* + \epsilon) = \beta \delta^2 u(c_2^*) \frac{1}{\mu - 1} u(c_1^*) + \beta \delta^3 u(c_3^*) \frac{1}{\mu - 1} u(c_3^*)
\]

\[
\iff \quad u(c_1^* + \epsilon) = u(c_1^*) + \delta \left( \frac{u(c_3^*)}{u(c_2^*)} \right)^{\frac{1}{\mu - 1}} u(c_3^*).
\]

From the preceding we obtain two expressions for \( u(c_1^* + p) - u(c_1^*) \). Putting them together we see that

\[
\delta \left( \frac{u(c_3^*)}{u(c_1^*)} \right)^{\frac{1}{\mu - 1}} u(c_3^*) = \delta \left( \frac{u(c_3^*)}{u(c_2^*)} \right)^{\frac{1}{\mu - 1}} u(c_3^*),
\]

which is equivalent to

\[
u(c_1^*) = u(c_2^*).
\]

This implies \( \bar{\mu} = \frac{u(c_3^*)}{u(c_1^*)} = 1 \) whereas we had assumed \( \bar{\mu} < 1 \), a contradiction. Since \( D_{c}(t) \leq 1 \) for all \( t \) and \( c \), \( D_{c}(t) \leq \beta \delta^t u(c^*) \frac{1}{\mu - 1} \leq \beta \delta u(c^*) \frac{1}{\mu - 1} \leq 1 \), which yields \( \beta \leq 1/\delta u(c^*) \frac{1}{\mu - 1} \). This completes the proof of sufficiency.

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

Now take any \( x, y \) satisfying the presumption of the Quasi Stationarity axiom. For any \( c \),

\[
U(x) \geq U(y) \iff \sum_{t \geq 1} \beta \delta^t f(x_t) \geq \sum_{t \geq 1} \beta \delta^t f(y_t)
\]

\[
\iff \sum_{t \geq 1} \beta \delta^{t+1} f(x_t) \geq \sum_{t \geq 1} \beta \delta^{t+1} f(y_t)
\]

\[
\iff u(c) + \sum_{t \geq 1} \beta \delta^{t+1} f(x_t) \geq u(c) + \sum_{t \geq 1} \beta \delta^{t+1} f(y_t)
\]

\[
\iff U(cx) \geq U(cy),
\]

as desired. \( \blacksquare \)
7 Consumption Smoothing

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

**Definition 3 (Consumption Smoothing)** A preference \( \succsim \) exhibits consumption smoothing if for any \( \alpha \in [0, 1] \) and for all deterministic streams \( x, y \in CT^{+1} \),

\[
x \sim y \implies \alpha x + (1 - \alpha)y \succsim x.
\]

NT state that

**Proposition 3 (NT [3, Proposition 6])** Assume that \( \succsim \) admits a homogeneous CE representation. If \( u(c)^\pi \) is concave in \( c \in \mathbb{R}_+ \), then \( \succsim \) exhibits consumption smoothing. Conversely, if \( \succsim \) exhibits consumption smoothing, then at least \( T \) of functions \( u(x_0), D_u(x_1), \ldots, D_u(x_T)u(x_T) \) are concave. Moreover, \( u : \mathbb{R}_+ \to \mathbb{R}_+ \) is concave if there are \( t, s \geq 1 \) such that \( \pi_t \neq \pi_s \).

**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 \( u(c)^\pi \). Note also that \( D_u(x_t)(t)u(x_t) = \gamma^t u(x_t)^\pi \) if \( x_t \leq u^{-1}(\pi_t) \) and \( D_u(x_t)(t)u(x_t) = \overline{D}_u(x_t) \) if \( x_t > u^{-1}(\pi_t) \). Since \( D_u(x_t)(t)u(x_t) \) is the point-wise minimum of two concave functions on \( \mathbb{R}_+ \), that is, \( \gamma^t u(x_t)^\pi \) and \( \overline{D}_u(x_t) \), 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 \( x \succ y \). By 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 non-empty. By Continuity, \( A \) is closed, that is, compact. Thus, there exists a maximum of \( A \), denoted by \( \overline{\alpha} \).

Let \( x_\pi = \overline{\pi} x + (1 - \overline{\pi})y \). We show that \( x_\pi \sim y \). By definition, \( \overline{\pi} \in A \). If \( \overline{\pi} = 1 \), \( x_\pi = x \succ y \), which is a contradiction. Thus, \( \overline{\pi} < 1 \). By seeking a contradiction, suppose \( x_\pi \not\succ y \). That is, \( y \succ x_\pi \). Since \( A^\circ = \{ \alpha \in [0, 1] | y \succ \alpha x + (1 - \alpha)y \} \) is open and \( \overline{\pi} \in A^\circ \), we can find some \( \alpha > \overline{\pi} \) with \( \alpha \in A^\circ \). But, this contradicts to the maximality of \( \overline{\pi} \) in \( A \).

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

\[
\frac{\tilde{\alpha}}{\overline{\alpha}}(\overline{\pi} x + (1 - \overline{\pi})y) + (1 - \frac{\tilde{\alpha}}{\overline{\alpha}})(\tilde{\alpha} x + (1 - \tilde{\alpha})y) \prec y,
\]

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 [4] shows that if an additively separable function \( F(x_1, \cdots, x_S) = \sum_s f_s(x_s) \) is quasi-concave, then at least \( S - 1 \) of functions \( f_1, \cdots, f_S \) are concave. Therefore, at least \( T \) of \( u(x_0), D_{u(x_1)}u(x_1), \cdots, D_{u(x_T)}u(x_T) \) 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 \( \bar{x}_t > \bar{x}_s \) for some \( t, s \). Then, \( u \) is concave on \([0, \bar{x}_t]\) and \([\bar{x}_s, \infty)\), which implies that \( u \) is concave on \( \mathbb{R}_+ \) by Lemma 2.2 of Li and Yeh [1].

8 Procrastination

In their application of the CE model to procrastination, NT present two propositions:

8.1 Sophisticated DU

**Proposition 4 (NT [3, Proposition 7])** Consider a sophisticated DU agent. Suppose self 2 would not complete a task when there is one to be done. On the path of a unique equilibrium, neither self 0 nor self 2 completes any tasks.

**Proof.** 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:

\[
V_2(1\mid 1) = u(b - e) + D(1)u(b + R) < u(b) + D(1)u(b) = V_2(0\mid 1),
\]

which is equivalent to

\[
D(1)[u(b + R) - u(b)] < u(b) - u(b - e).
\] (4)

Since the expressions remain the same even if there were two tasks to be completed, we see that \( V_2(1\mid 1) < V_2(0\mid 1) \) implies that

\[
V_2(1\mid 2) < V_2(0\mid 2),
\]

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 \( u(b) - u(b - 2e) \geq 2[u(b) - (b - e)] \), and so by (4),

\[
D(1)[u(b + 2R) - u(b)] < u(b) - u(b - 2e).
\] (5)

It follows that

\[
V_2(2\mid 2) = u(b - 2e) + D(1)u(b + 2R) < u(b) + D(1)u(b) = V_2(0\mid 2),
\]

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. Conclude that no self will do any task. ■
8.2 Sophisticated CE

Proposition 5 (NT [3, Proposition 8]) Consider a sophisticated CE agent. If self 2 would not complete a task when there is one to be done, then the model permits three possibilities on the path of a unique equilibrium:
(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. The hypothesis states that self 2 would not exert effort when there is 1 task to complete:
\[ U_2(1|1) < U_2(0|1) \]
\[ \iff u(b - e) + \gamma(1)u(b + R)^{\frac{m}{m-1}} < u(b) + \gamma(1)u(b)^{\frac{m}{m-1}} \]
\[ \iff \gamma(1)[u(b + R) - u(b)]^{\frac{m}{m-1}} < u(b) - u(b - e). \] (6)

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:
\[ U_2(1|1) < U_2(0|1) \implies U_2(1|2) < U_2(0|2). \] (7)

Since, by (7), completing 0 tasks is preferred to completing 1 task, we see that self 2 completes both tasks if and only if
\[ U_2(2|2) \geq U_2(0|2) \]
\[ \iff u(b - 2e) + \gamma(1)u(b + 2R)^{\frac{m}{m-1}} \geq u(b) + \gamma(1)u(b)^{\frac{m}{m-1}} \]
\[ \iff \gamma(1)[u(b + 2R) - u(b)]^{\frac{m}{m-1}} \geq u(b) - u(b - 2e). \] (8)

A novel feature of the CE model is that self 2 may complete two tasks by the magnitude effect even when she is reluctant to complete one task. To see that inequalities (6) and (8) can both hold, suppose \( u \) is linear, and observe that the inequalities become \( \gamma(1)[R]^{\frac{m}{m-1}} < e \) and \( 2^{\frac{m}{m-1}} \times \gamma(1)[R]^{\frac{m}{m-1}} \geq e \), respectively. Since \( 2^{\frac{m}{m-1}} > 1 \), there exist parameter values for which both can hold simultaneously.

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

\[ U_0(2|2) = u(b - 2e) + \gamma(1)u(b + 2R)\frac{m}{m-1} + \gamma(2)u(b)\frac{m}{m-1} + \gamma(3)u(b)\frac{m}{m-1}, \]

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

\[ U_0(0|2) = u(b) + \gamma(1)u(b)\frac{m}{m-1} + \gamma(2)u(b - 2e)\frac{m}{m-1} + \gamma(3)u(b + 2R)\frac{m}{m-1}. \]

Therefore, self 0 completes both tasks iff

\[ U_0(2|2) \geq U_0(0|2) \]
\[ \iff \gamma(1)[u(b + 2R)\frac{m}{m-1} - u(b)\frac{m}{m-1}] + \gamma(2)[u(b)\frac{m}{m-1} - u(b - 2e)\frac{m}{m-1}] \]
\[ \geq u(b) - u(b - 2e) + \gamma(3)[u(b + 2R)\frac{m}{m-1} - u(b)\frac{m}{m-1}]. \]  \hspace{1cm} (9)

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, when \( m \) 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 compared with a large payoff \( u(b + 2R)\frac{m}{m-1} \). Thus, when \( \gamma(1) \) and \( \gamma(3) \) are sufficiently close, the inequality (9) 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.

References


