# Supplementary Appendix to "Constrained Optimal Discounting"*

Jawwad Noor Norio Takeoka

February 28, 2023

**Abstract**

This supplementary appendix to Noor and Takeoka [1] provides a necessity of axioms for the CCE* representation, derivation of the reduced forms of the CCE* and the homogeneous CCE models, and proofs for the applications.

# 1 Necessity and Reduced Form of the CCE* Model

Theorem 1 in Noor and Takeoka [1] identifies a set of necessary and sufficient axioms for the CCE* representation. This section provides the necessity. First, we derive the reduced form of the CCE* representation, whereby the necessity is proved.

## 1.1 Reduced Form Representation

In this subsection, a proof of Proposition 5 (derivation of the reduced-form representation) is provided. As a preliminary, we first verify that for any $x \in X \setminus \Delta_0$, there exists a unique $\lambda_x > 0$ such that

$$\sum_{t>0} \lambda u(x_t)(\varphi_t')^{-1}(\lambda u(x_t)) = \overline{v},$$

which is associated with the definition of CCE* model. Recall the regularity* condition, $\varphi_t'(\overline{D}(t))\overline{D}(t) = \overline{v}$. Since $\varphi_t'$ is defined on $[0, 1]$, it is evident that $\overline{D}(t) \leq 1$. Denote

*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. Takeoka gratefully acknowledges financial support from JSPS KAKENHI Grant Number JP19KK0308. Part of this research was conducted while Norio Takeoka was visiting the Department of Economics, Boston University, whose hospitality is gratefully acknowledged. We thank Jiaqi Yang for expert research assistance. The usual disclaimer applies.

$\overline{r}_t = \varphi_t'(\overline{D}(t))$. Since $\varphi_t'$ is increasing, its inverse function $(\varphi_t')^{-1}$ is well-defined on $[0, \overline{r}_t]$. The function

$$V(r) = \sum_{t>0} r_t (\varphi_t')^{-1}(r_t)$$

is well-defined on $\prod_{t>0}[0, \overline{r}_t]$. In particular, at $\overline{r} = (\overline{r}_1, \cdots, \overline{r}_T)$, we have

$$\sum_{t>0} \overline{r}_t (\varphi_t')^{-1}(\overline{r}_t) > \overline{r}_t (\varphi_t')^{-1}(\overline{r}_t) = \overline{v}.$$

For any steam $x$, there exists some $\lambda > 0$ such that $\lambda u(x_t) \leq \overline{r}_t$ for all $t > 0$. Since $(\varphi_t')^{-1}(r_t)$ is well-defined and increasing on this region, there exists a unique $\lambda_x > 0$ satisfying the above equation.

**Lemma 1** *For all $x$ and $\mu > 0$, $\lambda_x = \mu \lambda_{\mu \circ x}$.*

**Proof.** By definition, $\lambda_{\mu \circ x}$ is a unique solution to

$$\sum_{t>0} \lambda u(\mu \circ x_t)(\varphi_t')^{-1}(\lambda u(\mu \circ x_t)) = \overline{v}.$$

By rearrangment, we have

$$\sum_{t>0} \lambda_{\mu\circ x} u(\mu \circ x_t)(\varphi_t')^{-1}(\lambda_{\mu\circ x} u(\mu \circ x_t)) = \sum_{t>0} \mu \lambda_{\mu\circ x} u(x_t)(\varphi_t')^{-1}(\mu \lambda_{\mu\circ x} u(x_t)) = \overline{v}.$$

Since the solution is unique also for $x$, we must have $\lambda_x = \mu \lambda_{\mu \circ x}$. ■

**Lemma 2** *For all $x$ and $\mu > 0$, $K_x = K_{\mu \circ x}$.*

**Proof.** By definition of $K_x$ and Lemma 1,

$$K_x = \sum_{t>0} \varphi_t((\varphi_t')^{-1}(\lambda_x u(x_t))) = \sum_{t>0} \varphi_t((\varphi_t')^{-1}(\mu \lambda_{\mu\circ x} u(x_t))) = \sum_{t>0} \varphi_t((\varphi_t')^{-1}(\lambda_{\mu\circ x} u(\mu \circ x_t))) = K_{\mu\circ x}.$$

■

Now we solve the cognitive optimization problem and derive the reduced form of the representation. We first solve the cognitive optimization only under the capacity constaint with ignoring the boundary consraint $D_x(t) \leq 1$. Then, we verify that the discount function derived in the first step, called a quasi-optimal discount function, also satisfies the boundary constraint, and hence, it is acutally an optimal discount function.

Take any $x$ and consider the following two cases.

Case 1: $\sum_{t>0} u(x_t)(\varphi_t')^{-1}(u(x_t)) \leq \overline{v}$. In this case, we have $\lambda_x \geq 1$. Hence,

$$\sum_{t>0} \varphi_t((\varphi_t')^{-1}(u(x_t))) \leq \sum_{t>0} \varphi_t((\varphi_t')^{-1}(\lambda_x u(x_t))) = K_x.$$

This means the unconstrained optimal discount function $D_x^{un}(t) = (\varphi_t')^{-1}(u(x_t))$ is feasible in the capacity constraint. Therefore, a quasi-optimal disount function is given by $D_x = D_x^{un}$, whereby, the corresponding representation is written as

$$U(x) = u(x_0) + \sum_{t>0} (\varphi_t')^{-1}(u(x_t))u(x_t).$$

Note that $U(x)$ is additive separable across time and a quasi-optimal discount function $D_{u(x_t)}(t)$ is strictly increasing in $u(x_t)$.

Case 2: $\sum_{t>0} u(x_t)(\varphi_t')^{-1}(u(x_t)) > \overline{v}$. In this case, we have $\lambda_x < 1$, and

$$\sum_{t>0} \varphi_t((\varphi_t')^{-1}(\lambda_x u(x_t))) = K_x.$$

This means that $D^*(t) = (\varphi_t')^{-1}(\lambda_x u(x_t))$ is feasible in the capacity constraint. Moreover, since $\lambda_x u(x_t) = \varphi_t'(D^*(t))$ for all $t > 0$, it follows that $D^*$ solves the Lagrangian with the multiplier $1/\lambda_x$. Hence, $D^*$ is a quasi-optimal discount function for $x$, that is, $D_x = D^*$. The corresponding representation is written as

$$U(x) = u(x_0) + \sum_{t>0} (\varphi_t')^{-1}(\lambda_x u(x_t))u(x_t).$$

Now compare two streams $x$ and $\mu \circ x$ that satisfy case (2). Since $\lambda_x = \mu \lambda_{\mu \circ x}$ by Lemma 1,

$$D_x(t) = (\varphi_t')^{-1}(\lambda_x u(x_t)) = (\varphi_t')^{-1}(\mu \lambda_{\mu\circ x} u(x_t)) = (\varphi_t')^{-1}(\lambda_{\mu\circ x} u(\mu \circ x_t)) = D_{\mu\circ x}(t),$$

that is, a quasi-optimal discount function is constant on the ray.

Finally, we verify that the boundary constraint $D_x(t) \leq 1$ is satisfied at the quasi-optimal discont function, and hence, it is indeed optimal. As derived above, a quasi-optimal discount function $D_x$ is given by

$$D_x(t) = \begin{cases} (\varphi_t')^{-1}(u(x_t)) & \text{if } \lambda_x \geq 1, \\ (\varphi_t')^{-1}(\lambda_x u(x_t)) & \text{if } \lambda_x < 1. \end{cases}$$

Recall that we define $\overline{r}_t = \varphi_t'(\overline{D}(t))$. Since we must have $\lambda_x u(x_t) \leq \overline{r}_t$, it follows that $(\varphi_t')^{-1}(\lambda_x u(x_t)) \leq (\varphi_t')^{-1}(\overline{r}_t) = \overline{D}(t)$. This is, $D_x(t) \leq \overline{D}(t) \leq 1$, as desired.

## 1.2 Necessity

We first establish necessity of some conditions defining Regularity. Order, C-Monotonicity and Risk Preference are obvious. Given that the cognitive objective function is strictly concave and thus yields a unique solution $D_x$, the Maximum Theorem ensures that $D_x$ is a continuous function of $x$. Therefore the representation $U(x) = u(x_0) + \sum_{t\geq 1} D_x(t)u(x_t)$ is continuous, thereby establishing Continuity. Present Equivalent is satisfied due to the assumption that $u(C) = \mathbb{R}_+$. The necessity of Monotonicity and Impatience is established below, along with the necessity of the remaining conditions.

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

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

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

Define

$$\Lambda_x = \{D \in \mathbb{R}_+^{\mathrm{T}} : \sum_{t>0} \varphi_t(D(t)) \leq K_x\}.$$

**Lemma 3** *For all $x \in X \setminus \Delta_0$,*

$$\varphi(D_x^{un}) \leq K_x \iff \sum_{t>0} u(x_t)(\varphi_t')^{-1}(u(x_t)) \leq \overline{v}.$$

**Proof.** To prove ' $\Longleftarrow$ ', suppose $x$ satisfies $\sum_{t>0} u(x_t)(\varphi_t')^{-1}(u(x_t)) \leq \overline{v}$. Then, the unique solution $\lambda_x$ to the equation

$$\sum_{t>0} \lambda u(x_t)(\varphi_t')^{-1}(\lambda u(x_t)) = \overline{v}$$

must satisfy $\lambda_x \geq 1$. Consider the discount function $D_x^{un} = (D_{u(x_t)}^{un}(t))_{t>0}$ defined by the FOC, $D_{u(x_t)}^{un}(t) = (\varphi_t')^{-1}(u(x_t))$ for all $t$. From the representation,

$$K_x = \sum_{t>0} \varphi_t\Big((\varphi_t')^{-1}(\lambda_x u(x_t))\Big) \geq \sum_{t>0} \varphi_t\Big((\varphi_t')^{-1}(u(x_t))\Big) = \sum_{t>0} \varphi_t(D_{u(x_t)}^{un}(t)) = \varphi(D_x^{un}).$$

To prove ' $\Longrightarrow$ ', suppose $\varphi(D_x^{un}) \leq K_x$, where $D_x^{un}(t) = (\varphi_t')^{-1}(u(x_t))$. By the representation,

$$\sum_{t>0} \varphi_t\Big((\varphi_t')^{-1}(\lambda_x u(x_t))\Big) = K_x \geq \varphi(D_x^{un}),$$

where $\lambda_x$ solves $\sum_{t>0} \lambda u(x_t)(\varphi_t')^{-1}(\lambda u(x_t)) = \overline{v}$. Since $\varphi(D_{\lambda_x \circ x}^{un}) \geq \varphi(D_x^{un})$, we must have $\lambda_x \geq 1$. Thus,

$$\sum_{t>0} u(x_t)(\varphi_t')^{-1}(u(x_t)) \leq \sum_{t>0} \lambda_x u(x_t)(\varphi_t')^{-1}(\lambda_x u(x_t)) = \overline{v},$$

as desired. ■

**Lemma 4** *For all $x \in X \setminus \Delta_0$,*

$$x \in X_{ms} \iff \sum_{t>0} u(x_t)(\varphi_t')^{-1}(u(x_t)) \leq \overline{v}.$$

**Proof.** To show that $X_{ms}$ is a subset of the right-hand side, take any $x \in X_{ms}$. By the representation,

$$u(c_x) = U(x) = u(x_0) + \sum_{t>0} D_x(t)u(x_t),$$

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

$$\sum_{t>0} D_x(t)u(x_t) > \sum_{t>0} D_{\alpha\circ x}(t)u(x_t).$$

Since $u(x_t) \geq 0$ and $D_x \geq D_{\alpha\circ x}$ as shown in Section 1.1, we have $D_x(t) > D_{\alpha\circ x}(t)$ for some $t$. By definition of $D_x$, together with Lemma 2,

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

Hence, $D_{\alpha\circ x} = D^{un}_{\alpha\circ x}$. As $\alpha \to 1$, we have $\varphi(D^{un}_x) \leq K_x$. By Lemma 3, $x$ belongs to the right-hand side.

Conversely, take any $x$ in the right hand side. By Lemma 3, $\varphi(D^{un}_x) \leq K_x$. Since $D^{un}_x$ is an optimal discount function, $U(x) = u(x_0) + \sum_{t>0}(\varphi'_t)^{-1}(u(x_t))u(x_t)$. Since $(\varphi'_t)^{-1}(u(x_t))$ is increasing in $u(x_t)$, it is obvious that both $(x_t, 0_{-t})$ and $(0_t, x_{-t})$ belong to the right-hand side. Thus, by the same reason as above, $U(0_t, x_{-t}) = u(x_0) + \sum_{s\neq t}(\varphi'_s)^{-1}(u(x_s))u(x_s)$ and $U(x_t, 0_{-t}) = (\varphi'_t)^{-1}(u(x_t))u(x_t)$. This implies that

$$\begin{aligned}
&U(x) = U(x_t, 0_{-t}) + U(0_t, x_{-t}) \\
\iff &u(c_x) = u(c_{(x_t,0_{-t})}) + u(c_{(0_t,x_{-t})}) \\
\iff &\frac{1}{2}u(c_x) + \frac{1}{2}u(0) = \frac{1}{2}u(c_{(x_t,0_{-t})}) + \frac{1}{2}u(c_{(0_t,x_{-t})}) \\
\iff &u(\frac{1}{2} \circ c_x + \frac{1}{2} \circ 0) = u(\frac{1}{2} \circ c_{(x_t,0_{-t})} + \frac{1}{2} \circ c_{(0_t,x_{-t})}),
\end{aligned}$$

which means $x$ is separable.

Next, we show that $x$ is magnitude sensitive. For $\alpha \in (0,1)$, by the FOC,

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

Therefore,

$$D^{un}_{\alpha\circ x} = D_{\alpha\circ x} = \arg\max_{\Lambda_{\alpha\circ x}}\{\sum_{t>0} D(t)u(\alpha \circ x_t) - \varphi_t(D(t))\}.$$

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

$$\begin{aligned}
u(\alpha \circ c_x) = \alpha u(c_x) = \alpha U(x) &= u(\alpha \circ x_0) + \sum_{t>0} D_x(t)u(\alpha \circ x_t) \\
&> u(\alpha \circ x_0) + \sum_{t>0} D_{\alpha x}(t)u(\alpha \circ x_t) = U(\alpha \circ x),
\end{aligned}$$

or, $\alpha \circ c_x \succ \alpha \circ x$. That is, $x$ is magnitude sensitive. Therefore, $x \in X_{ms}$. ■

**Lemma 5** $X_{ms} = \{x \in X \setminus \Delta_0 : U(0, x_{-0}) \leq \overline{v}\}$.

**Proof.** Take any $x \in X_{ms}$. By Lemma 4, $x$ satisfies $\sum_{t\geq 1} u(x_t)(\varphi_t')^{-1}(u(x_t)) \leq \overline{v}$, and then, the unique $\lambda_x$ satisfying $\sum_{t=1}^{T} \lambda u(x_t)(\varphi_t')^{-1}(\lambda u(x_t)) = \overline{v}$ must satisfy $\lambda_x \geq 1$. Thus, $\overline{\lambda}_x = \min[\lambda_x, 1] = 1$, and hence,

$$U(0, x_{-0}) = \sum_{t>0} (\varphi_t')^{-1}(\overline{\lambda}_x u(x_t))u(x_t) = \sum_{t>0} (\varphi_t')^{-1}(u(x_t))u(x_t) \leq \overline{v}.$$

Conversely, assume $x$ satisfies $U(0, x_{-0}) \leq \overline{v}$. Seeking a contradiction, suppose $x \notin X_{ms}$. By Lemma 4, $\sum_{t>0} u(x_t)(\varphi_t')^{-1}(u(x_t)) > \overline{v}$. Thus, the unique solution $\lambda_x$ to the equation must satifiy $\lambda_x < 1$. Notice that $\sum_{t>0} \lambda_x u(x_t)(\varphi_t')^{-1}(\lambda_x u(x_t)) = \overline{v}$ implies that the stream $\lambda_x \circ x$ satisfies the same equation when $\lambda_{\lambda_x \circ x} = 1$. By Lemma 3, this measn that $D^{un}(t) = (\varphi_t')^{-1}(u(\lambda_x \circ x_t))$ is feasible for all $t > 0$. Therefore, by the representation,

$$U(0, \lambda_x \circ x_{-0}) = \sum_{t>0} (\varphi_t')^{-1}(u(\lambda_x \circ x_t))u(\lambda_x \circ x_t) = \overline{v}.$$

On the other hand, since $\lambda_x < 1$, Monotonicity implies $U(0, \lambda_x \circ x_{-0}) < U(0, x_{-0}) \leq \overline{v}$, which contradicts the above equation. ■

**Lemma 6** $\succsim$ *satisfies Weak Homotheticity.*

**Proof.** Take any stream $x \in X$. As shown in Section 1.1, $D_{\lambda\circ x}(t)$ is increasing in $\lambda > 0$ if $U(0, \lambda \circ x_0) \leq \overline{v}$ and constant otherwise. Thus, for any $x$ and $\alpha \in (0, 1)$, $D_x(t) \geq D_{\alpha x}(t)$, which implies, with linearity of $u$, $\alpha U(x) \geq U(\alpha x)$, or $\alpha \circ c_x \succsim \alpha \circ x$, as desired. ■

**Lemma 7** $\succsim$ *satisfies Magnitude-Sensitive Separability.*

**Proof.** Take any magnitude sensitive stream $x$. If $x \in X_{ms}$, from Lemmas 3 and 4, $D_x = D_x^{un}$ on $X_{ms}$. Thus, $D_x(t)$ depends only on $u(x_t)$, and hence, $x$ is separable. Next, consider $x \notin X_{mx}$. We will claim that then $x$ is not magnitude sensitive. By Lemma 4, $\sum_{t>0} u(x_t)(\varphi_t')^{-1}(u(x_t)) > \overline{v}$. Thus, the unique solution $\lambda_x$ to the equation must satifiy $\lambda_x < 1$. Notice that $\sum_{t>0} \lambda_x u(x_t)(\varphi_t')^{-1}(\lambda_x u(x_t)) = \overline{v}$ implies that the stream $\lambda_x \circ x$ satisfies the same equation when $\lambda_{\lambda_x \circ x} = 1$. This means that $x$ and $\lambda_x \circ x$ have the same optimal discount function $D_x(t) = D_{\lambda_x \circ x}(t) = (\varphi_t')^{-1}(u(\lambda_x \circ x_t))$ for all $t > 0$. Therefore, we have

$$\begin{aligned} u(\lambda_x \circ c_x) = \lambda_x u(c_x) = \lambda_x U(x) = \lambda_x(u(x_0) + \sum_{t>0} D_x(t)u(x)) \\ = \lambda_x(u(x_0) + \sum_{t>0} D_{\lambda_x \circ x}(t)u(x)) = u(\lambda_x \circ x_0) + \sum_{t>0} D_{\lambda_x \circ x}(t)u(\lambda_x \circ x) = U(\lambda_x \circ x), \end{aligned}$$

which means $\lambda_x \circ c_x \sim \lambda_x \circ x$ for some $\lambda_x \in (0, 1)$. That is, $x$ is magnitude insensitive. ■

**Lemma 8** $\succsim$ *satisfies PBT.*

**Proof.** If $x \notin X_{ms}$, by Lemma 5, $U(0, x_{-0}) > \overline{v}$. Then since $\lim_{\lambda \to 0} U(0, \lambda x_{-0}) = 0$, it follows there exists $\lambda_x < 1$ s.t. $U(0, \lambda_x \circ x_{-0}) \leq \overline{v}$ or $\lambda_x x \in X_{ms}$ by Lemma 5, as required by PBT (i). To confirm PBT (ii), note that if $x \in X_{ms}$ then $U(0, x_{-0}) \leq \overline{v}$ and it follows that for any $y$ satisfying $U(0, y_{-0}) \leq U(0, x_{-0})$ it must be that $y \in X_{ms}$. ■

**Lemma 9** $\succsim$ *satisfies Monotonicity.*

**Proof.** Since the representation is additively separable in $x_0$ it suffices to establish Monotonicity on the set of all streams of the form $(0, x_{-0})$. We saw in Lemma 5 that $X_{ms} = \{x \in X \setminus \Delta_0 : U(0, x_{-0}) \leq \overline{v}\}$. Take two streams $(0, x_{-0}) \geq (0, y_{-0})$. If we are in the case $U(0, y_{-0}) \leq \overline{v} \leq U(0, x_{-0})$, then we are done. Consider the two remaining cases:

(1) $U(0, y_{-0}), U(0, x_{-0}) \leq \overline{v}$.

Since $U(0, y_{-0}), U(0, x_{-0}) \leq \overline{v}$ implies $(0, y_{-0}), (0, x_{-0}) \in X_{ms}$ , it follows that $U(0, y_{-0}) \leq U(0, x_{-0})$, because $U$ is additively separable on $X_{ms}$, $u$ is strictly increasing and $D_r(t)$ is strictly increasing in $r$.

(2) $\overline{v} \leq U(0, y_{-0})$

Suppose by way of contradiction that $U(0, x_{-0}) < U(0, y_{-0})$. Consider scaling down $(0, y_{-0})$ by $\lambda^* < 1$ s.t. $U(0, \lambda^* \circ y_{-0}) = \overline{v}$. We claim that in fact $U(0, \lambda^* \circ x_{-0}) < U(0, \lambda^* \circ y_{-0}) = \overline{v}$. This is obvious in the case where $(0, x_{-0}) \in X_{ms}$ since $U$ satisfies Monotonicity on $X_{ms}$ (as noted in the proof of (1) above). In the case where $(0, x_{-0}) \notin X_{ms}$, this is true because, as shown in Section 1.1, $U(0, \lambda \circ x_{-0})$ initially decreases linearly if we decrease $\lambda$ (because $D_{(0, \lambda \circ x_{-0})}$ is constant as long as $(0, \lambda \circ x_{-0}) \notin X_{ms}$) and eventually faster than linear (because $D_{(0, \lambda \circ x_{-0})}$ decreases as $\lambda$ decreases when $(0, \lambda \circ x_{-0}) \in X_{ms}$), while $U(0, \lambda \circ y_{-0})$ reduces only linearly over $\lambda \in [\lambda^*, 1]$. Thus, we have $U(0, \lambda^* \circ x_{-0}) < U(0, \lambda^* \circ y_{-0}) = \overline{v}$. But this contradicts (1) since $(0, \lambda^* \circ x_{-0}) \geq (0, \lambda^* \circ y_{-0})$ and both streams are in $X_{ms}$. ■

**Lemma 10** $\succsim$ *satisfies Impatence.*

**Proof.** Since the cognitive optimization problem requires the boundary constraint $D_x(t) \leq 1$, in particular for dated rewards $p^t$, we have $D_{u(p)}(t) \leq 1$. Hence, $u(p) \geq D_{u(p)}(t)u(p)$, which implies $p \succsim p^t$ for all $t$.

Next, take any $p \in \Delta$ and $0 < t < s$. Suppose first that $p^t, p^s \in X_{ms}$. Note that by the FOC

$$\varphi_t'(D_{u(p)}(t)) = u(p) = \varphi_s'(D_{u(p)}(s)).$$

By the property $\varphi_t' \leq \varphi_s'$ of the representation, we must have $D_{u(p)}(t) \geq D_{u(p)}(s)$ and thus $p^t \succsim p^s$ as desired.

Next suppose $p^s \notin X_{ms}$ and suppose by way of contradiction that $p^t \prec p^s$. Arguing as in case (2) in Lemma 9, there exists $\lambda^* < 1$ s.t. $U(\lambda^* \circ p^t) < U(\lambda^* \circ p^s) = \overline{v}$. But then $(\lambda^* \circ p^t), (\lambda^* \circ p^s) \in X_{ms}$ and $(\lambda^* \circ p^t) \prec (\lambda^* \circ p^s)$ contradicts what we just established in the preceding paragraph. ■

## 2 Appendix: Proof of Proposition 6

We solve the cognitive optimization problem for each $x$. Let $\varphi_t(d) = a_t d^m$ on $d \in [0,1]$. As explained in Section 1.1, the boundary constraint $D(t) \le 1$ is effectively ignored. For each $x$, an optimal discount function $\{D_x(t)\}_{t>0}$ is determined by

$$\max_{D\ge 0} \sum_{t>0} D(t)u(x_t) - \sum_{t>0} \varphi_t(D(t)), \quad \text{subject to} \sum_{t>0} \varphi_t(D(t)) \le K.$$

The FOC of the above maximization problem is obtained as the FOC of the following Lagrangian:

$$\mathcal{L} = \sum_{t>0} D(t)u(x_t) - \sum_{t>0} a_t D(t)^m + \xi(K - \sum_{t>0} a_t D(t)^m),$$

where $\xi \ge 0$ is a Lagrange multiplier for the capacity constraint. By differentiating $\mathcal{L}$ with respect to $D(t)$, we have

$$D_x(t) = \left(\frac{u(x_t)}{(1+\xi)ma_t}\right)^{\frac{1}{m-1}}, \tag{1}$$

for all $t = 1, \cdots, T$.

Suppose $x$ is small. Since the capacity constraint is not binding, we have $\xi = 0$. Thus,

$$D_x(t) = \left(\frac{u(x_t)}{ma_t}\right)^{\frac{1}{m-1}}$$

and

$$U(x) = u(x_0) + \sum_{t>0} D_x(t)u(x_t) = u(x_0) + \sum_{t>0} \gamma(t) u(x_t)^{\frac{m}{m-1}}, \tag{2}$$

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

Next, suppose $x$ is large. Then, the capacity constraint is binding. By substituting (1) into the capacity constraint,

$$\sum_{t>0} a_t \left(\frac{u(x_t)}{(1+\xi)ma_t}\right)^{\frac{m}{m-1}} = K.$$

By rearrangement,

$$\frac{1}{(1+\xi)^{\frac{1}{m-1}}} = \frac{K^{\frac{1}{m}}}{\left\{\sum_{t\ge 1} a_t \left(\frac{u(x_t)}{ma_t}\right)^{\frac{m}{m-1}}\right\}^{\frac{1}{m}}}.$$

By substituting it into (1),

$$D_x(t) = \frac{K^{\frac{1}{m}} \left(\frac{u(x_t)}{ma_t}\right)^{\frac{1}{m-1}}}{\left\{\sum_{t\ge 1} a_t \left(\frac{u(x_t)}{ma_t}\right)^{\frac{m}{m-1}}\right\}^{\frac{1}{m}}} = \frac{(mK)^{\frac{1}{m}} \gamma(t) u(x_t)^{\frac{1}{m-1}}}{\left\{\sum_{t\ge 1} \gamma(t) u(x_t)^{\frac{m}{m-1}}\right\}^{\frac{1}{m}}}.$$

Therefore,

$$\begin{aligned} U(x) &= u(x_0) + \sum_{t>0} D_x(t)u(x_t) \\ &= u(x_0) + (mK)^{\frac{1}{m}} \left\{ \sum_{t>0} \gamma(t)u(x_t)^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}}. \end{aligned} \tag{3}$$

Finally, we derive a threshold where small and large streams are distinguished. At this boundary of consumption streams,

$$\sum_{t>0} \varphi_t(D_x(t)) = \sum_{t>0} a_t \left( \frac{u(x_t)}{ma_t} \right)^{\frac{m}{m-1}} = K.$$

Equivalently,

$$\sum_{t>0} \gamma(t)u(x_t)^{\frac{m}{m-1}} = mK.$$

Therefore, at the boundary,

$$U(x) = u(x_0) + \sum_{t>0} \gamma(t)u(x_t)^{\frac{m}{m-1}} = u(x_0) + mK.$$

# 3 Proof of Proposition 7

**Lemma 11** *(Sophisticated CCE Saving Rule) Assume $0 < \sigma\frac{m}{m-1} < 1$ and suppose that both self 0 and self 1 are cognitively constrained at their respective optimal consumption path. The optimal saving rules for the sophisticated CCE model are given by*

$$s_1^* = A_1(I_1 + Rs_0) \qquad \text{where } A_1 := \frac{1}{1 + [\left(\frac{K_1}{a_1}\right)^{\frac{1}{m}} R^\sigma]^{\frac{1}{\sigma-1}}},$$

*and*

$$s_0^* = \frac{I_0 - (\left(\frac{K_0}{a_1}\right)^{\frac{1}{m}} A_0 R)^{\frac{1}{\sigma-1}} I_1}{1 + (\left(\frac{K_0}{a_1}\right)^{\frac{1}{m}} A_0 R)^{\frac{1}{\sigma-1}} R} \qquad \text{where } A_0 := R^\sigma \frac{\left( \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} + \left( \left(\frac{K_1}{a_1}\right)^{\frac{1}{m}} R \right)^{\frac{\sigma\theta}{\sigma-1}} \right)^{\frac{1}{\theta}}}{\left( 1 + \left(\frac{K_1}{a_1}\right)^{\frac{1}{m}\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}} \right)^\sigma}.$$

**Proof.** Write $0 < \sigma\theta < 1$, where $\theta = \frac{m}{m-1}$. Let $\gamma_t = (a_t m)^{-\frac{1}{m-1}}$. By Proposition 6 of Noor and Takeoka[1], the homogeneous CCE representation takes the form:

$$U(c_0, c_1, c_2) = \begin{cases} c_0^\sigma + \gamma_1 c_1^{\sigma\theta} + \gamma_2 c_2^{\sigma\theta} & \text{if } \gamma_1 c_1^{\sigma\theta} + \gamma_2 c_2^{\sigma\theta} \le mK_0, \\ \\ c_0^\sigma + (mK_0)^{\frac{1}{m}} [\gamma_1 c_1^{\sigma\theta} + \gamma_2 c_2^{\sigma\theta}]^{\frac{1}{\theta}} & \text{if } \gamma_1 c_1^{\sigma\theta} + \gamma_2 c_2^{\sigma\theta} > mK_0, \end{cases}$$

or equivalently,

$$U(c_0, c_1, c_2) = \begin{cases} c_0^\sigma + \gamma_1 c_1^{\sigma\theta} + \gamma_2 c_2^{\sigma\theta} & \text{if } \delta_1 c_1^{\sigma\theta} + \delta_2 c_2^{\sigma\theta} \leq m^\theta K_0, \\ c_0^\sigma + (K_0)^{\frac{1}{m}} [(\delta_1 c_1^{\sigma\theta} + \delta_2 c_2^{\sigma\theta}]^{\frac{1}{\theta}} & \text{if } \delta_1 c_1^{\sigma\theta} + \delta_2 c_2^{\sigma\theta} > m^\theta K_0, \end{cases}$$

where $\delta_t = (a_t^{-1})^{\frac{1}{m-1}}$. Note that $\delta_2 \leq \delta_1 \leq 1$. The utility function for self 1 is analogous. We proceed by backward induction.

Solve for the optimal consumption of self 1 assuming that she will be constrained at the solution, that is, $\delta_2 c_2^{\sigma\theta} > m^\theta K_1$. Her problem is

$$\max_{s_1} \, (I_1 + Rs_0 - s_1)^\sigma + \left(\frac{K_1}{a_1}\right)^{\frac{1}{m}} (Rs_1)^\sigma.$$

The FOC $(I_1 + Rs_0 - s_1)^{\sigma-1} = \left(\frac{K_1}{a_1}\right)^{\frac{1}{m}} R^\sigma s_1^{\sigma-1}$ yields the saving rule

$$s_1^* = A_1(I_1 + Rs_0), \text{ where } A_1 = \frac{1}{1 + [\left(\frac{K_1}{a_1}\right)^{\frac{1}{m}} R^\sigma]^{\frac{1}{\sigma-1}}}. \tag{4}$$

Now turn to self 0 and assume that she is sophisticated. We solve for her optimal consumption assuming that she is constrained at the solution, that is, $\delta_1 c_1^{\sigma\theta} + \delta_2 c_2^{\sigma\theta} > m^\theta K_0$. Her maximization problem is

$$\max_{s_0, s_1} \, (I_0 - s_0)^\sigma + K_0^{\frac{1}{m}} [\delta_1 (I_1 + Rs_0 - s_1)^{\sigma\theta} + \delta_2 (Rs_1)^{\sigma\theta}]^{\frac{1}{\theta}}.$$

Being sophisticated, she take $s_1^* = A_1(I_1 + Rs_0)$, and so the problem becomes

$$\begin{aligned}
& \max_{s_0} \, (I_0 - s_0)^\sigma + K_0^{\frac{1}{m}} [\delta_1 (I_1 + Rs_0 - s_1^*)^{\sigma\theta} + \delta_2 (Rs_1^*)^{\sigma\theta}]^{\frac{1}{\theta}} \\
\iff & \max_{s_0} \, (I_0 - s_0)^\sigma + K_0^{\frac{1}{m}} [\delta_1 (I_1 + Rs_0 - A_1(I_1 + Rs_0))^{\sigma\theta} + \delta_2 (RA_1(I_1 + Rs_0))^{\sigma\theta}]^{\frac{1}{\theta}} \\
\iff & \max_{s_0} \, (I_0 - s_0)^\sigma + K_0^{\frac{1}{m}} [\delta_1 (1 - A_1)^{\sigma\theta} (I_1 + Rs_0)^{\sigma\theta} + \delta_2 (RA_1)^{\sigma\theta} (I_1 + Rs_0)^{\sigma\theta}]^{\frac{1}{\theta}} \\
\iff & \max_{s_0} \, (I_0 - s_0)^\sigma + K_0^{\frac{1}{m}} \left[\delta_1 (1 - A_1)^{\sigma\theta} + \delta_2 (RA_1)^{\sigma\theta}\right]^{\frac{1}{\theta}} (I_1 + Rs_0)^\sigma \\
\iff & \max_{s_0, s_1} \, (I_0 - s_0)^\sigma + \left(\frac{K_0}{a_1}\right)^{\frac{1}{m}} A_2 (I_1 + Rs_0)^\sigma,
\end{aligned}$$

where $A_0 := \left[(1 - A_1)^{\sigma\theta} + \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} (RA_1)^{\sigma\theta}\right]^{\frac{1}{\theta}}$ and we use the fact that $\delta_2/\delta_1 = \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}}$. The FOC is

$$\sigma(I_0 - s_0)^{\sigma-1} = \sigma \left(\frac{K_0}{a_1}\right)^{\frac{1}{m}} A_0 R (I_1 + Rs_0)^{\sigma-1}$$

which yields the time 0 saving rule

$$s_0^* = \frac{I_0 - (\left(\frac{K_0}{a_1}\right)^{\frac{1}{m}} A_0 R)^{\frac{1}{\sigma-1}} I_1}{1 + (\left(\frac{K_0}{a_1}\right)^{\frac{1}{m}} A_0 R)^{\frac{1}{\sigma-1}} R}.$$

Finally we verify that

$$\begin{aligned} A_0 &= \left[(1-A_1)^{\sigma\theta} + \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} (RA_1)^{\sigma\theta}\right]^{\frac{1}{\theta}} \\ &= \left[(\frac{[\left(\frac{K_1}{a_1}\right)^{\frac{1}{m}} R^\sigma]^{\frac{1}{\sigma-1}}}{1+[\left(\frac{K_1}{a_1}\right)^{\frac{1}{m}} R^\sigma]^{\frac{1}{\sigma-1}}})^{\sigma\theta} + \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} (R\frac{1}{1+[\left(\frac{K_1}{a_1}\right)^{\frac{1}{m}} R^\sigma]^{\frac{1}{\sigma-1}}})^{\sigma\theta}\right]^{\frac{1}{\theta}} \\ &= R^\sigma \left(\frac{\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} + \left(\left(\frac{K_1}{a_1}\right)^{\frac{1}{m}} R\right)^{\frac{\sigma\theta}{\sigma-1}}}{\left(1+\left(\frac{K_1}{a_1}\right)^{\frac{1}{m}\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}\right)^{\sigma\theta}}\right)^{\frac{1}{\theta}} = R^\sigma \frac{\left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} + \left(\left(\frac{K_1}{a_1}\right)^{\frac{1}{m}} R\right)^{\frac{\sigma\theta}{\sigma-1}}\right)^{\frac{1}{\theta}}}{\left(1+\left(\frac{K_1}{a_1}\right)^{\frac{1}{m}\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}\right)^{\sigma}}. \end{aligned}$$

■

**Lemma 12** *Assume $0 < \sigma\frac{m}{m-1} < 1$ and suppose that both self 0 and self 1 are cognitively constrained at their respective optimal consumption path. Then the following hold for the sophisticated CCE model:*

*(i) Given any wealth $I_1 + Rs_0$, $s_1^*$ is increasing in $K_1$.*

*(ii) $s_0^*$ is increasing in $K_0$.*

*(iii) $s_0^*$ is increasing in $K_1$ if and only if $a_1/a_2 > (R^{-\sigma}(K_1/a_1)^{1-\sigma-\frac{1}{m}})^{\frac{1}{1-\sigma}}$.*

*(iv) Let $K_0 = K_1 = K$. $s_0^*$ is increasing in $K$ if $a_1/a_2 > (R^{-\sigma}(K/a_1)^{1-\sigma-\frac{1}{m}})^{\frac{1}{1-\sigma}}$.*

**Proof.** Parts (i) and (ii) follow from the expression of the saving rules.

Proof of (iii): Write self 0's saving rule as

$$s_0^* = \frac{I_0 - (\left(\frac{K_0}{a_1}\right)^{\frac{1}{m}} A_0 R)^{\frac{1}{\sigma-1}} I_1}{1 + (\left(\frac{K_0}{a_1}\right)^{\frac{1}{m}} A_0 R)^{\frac{1}{\sigma-1}} R} = \frac{(R^{\sigma+1} \left(\frac{K_0}{a_1}\right)^{\frac{1}{m}} f(K_1))^{\frac{1}{1-\sigma}} I_0 - I_1}{(R^{\sigma+1} \left(\frac{K_0}{a_1}\right)^{\frac{1}{m}} f(K_1))^{\frac{1}{1-\sigma}} + R},$$

where

$$f(K_1) := \frac{(1 + \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R^{\frac{\sigma\theta}{1-\sigma}} \left(\frac{K_1}{a_1}\right)^{\frac{\sigma\theta}{m(1-\sigma)}})^{\frac{1}{\theta}}}{(1 + R^{\frac{\sigma}{1-\sigma}} \left(\frac{K_1}{a_1}\right)^{\frac{1}{m(1-\sigma)}})^{\sigma}}.$$

For notational simplicity, denote $\zeta = \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}}$ and $k = \left(\frac{K_1}{a_1}\right)^{\frac{1}{m}}$. Then, the sign of $f'(K_1)$ is the same as the sign of the derivative of

$$h(k) := \frac{(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}}}{(1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})^{\sigma}}. \tag{5}$$

The sign of $h'(k)$ is the same as the sign of its numerator, that is,

$$\begin{aligned}
&\frac{1}{\theta}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}-1} \frac{\sigma\theta\zeta}{1-\sigma} R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}-1} (1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})^{\sigma} \\
&- \sigma(1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})^{\sigma-1} \frac{1}{1-\sigma} R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}-1} (1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}}.
\end{aligned}$$

Equivalently

$$\begin{aligned}
&\frac{\sigma}{1-\sigma}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}} (1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})^{\sigma} \\
&\times \left( (1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})^{-1} \zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}-1} - (1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})^{-1} R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}-1} \right) \\
\iff &\frac{\sigma}{1-\sigma}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}} (1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})^{\sigma} \\
&\times \frac{R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}-1} \zeta (1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}}) - R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}-1} (1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})}{(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})(1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})} \\
\iff &\frac{\sigma}{1-\sigma}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}} (1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})^{\sigma} \\
&\times \frac{\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}-1} - R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}-1}}{(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})(1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})} \\
\iff &\frac{\sigma}{1-\sigma}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}-1} (1+R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})^{\sigma-1} \left( R^{\frac{\sigma}{1-\sigma}} (\zeta R^{\frac{\sigma(\theta-1)}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}-1} - k^{\frac{1}{1-\sigma}-1} \right). \quad (6)
\end{aligned}$$

Therefore, the sign of the above expression is equivalent to the sign of

$$\zeta R^{\frac{\sigma(\theta-1)}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}-1} - k^{\frac{1}{1-\sigma}-1}.$$

This expression is positive if and only if

$$\begin{aligned}
&\zeta R^{\frac{\sigma(\theta-1)}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}-1} > k^{\frac{1}{1-\sigma}-1} \\
\iff &\zeta > R^{-\frac{\sigma(\theta-1)}{1-\sigma}} k^{\frac{1-\sigma\theta}{1-\sigma}} \\
\iff &\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} > R^{-\frac{\sigma}{1-\sigma}\frac{1}{m-1}} \left(\frac{K_1}{a_1}\right)^{\frac{(1-\sigma)m-1}{m(m-1)(1-\sigma)}} = \left( R^{-\sigma} \left(\frac{K_1}{a_1}\right)^{\frac{(1-\sigma)m-1}{m}} \right)^{\frac{1}{(m-1)(1-\sigma)}} \\
\iff &\frac{a_1}{a_2} > \left( R^{-\sigma} \left(\frac{K_1}{a_1}\right)^{1-\sigma-\frac{1}{m}} \right)^{\frac{1}{1-\sigma}}. \quad (7)
\end{aligned}$$

Note that $1-\sigma-\frac{1}{m}>0$ because this condition is equivalent to $\sigma\theta<1$. Since $a_1 \le a_2$, we have $a_1/a_2 \le 1$. Moreover, since $R \ge 1$ and $K_1/a_1 \le 1$, we have $\left(R^{-\sigma}(K_1/a_1)^{\frac{(1-\sigma)m-1}{m}}\right)^{\frac{1}{1-\sigma}} \le 1$.

Proof of (iv): The sign of the derivative of $(K/a_1)^{\frac{1}{m}} f(K)$ is the same as the sign of the derivative of $kh(k)$, defined as in (5). Since the numerator of the derivative $h'(k)$ is derived as in (6),

$$\begin{aligned}
&[kh(k)]' = h(k) + kh'(k) = h(k)\left(1 + k\frac{h'(k)}{h(k)}\right)\\
=&h(k)\left(1 + k\frac{\sigma}{1-\sigma}R^{\frac{\sigma}{1-\sigma}}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})^{-1}(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})^{-1}\left(\zeta R^{\frac{\sigma(\theta-1)}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}-1} - k^{\frac{1}{1-\sigma}-1}\right)\right)\\
=&h(k)\left(1 + \frac{\sigma}{1-\sigma}R^{\frac{\sigma}{1-\sigma}}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})^{-1}(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})^{-1}\left(\zeta R^{\frac{\sigma(\theta-1)}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}} - k^{\frac{1}{1-\sigma}}\right)\right).
\end{aligned} \tag{8}$$

Therefore, $s_0^*$ is increasing in $K$ if and only if the sign of the parenthesis is positive. Note also that

$$\zeta R^{\frac{\sigma(\theta-1)}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}} > k^{\frac{1}{1-\sigma}} \iff \zeta > R^{-\frac{\sigma(\theta-1)}{1-\sigma}}K^{\frac{1-\sigma\theta}{m(1-\sigma)}}.$$

This expression is the same as (7). If this inequality holds, (8) is positive. ■

# 4 Proof of Proposition 8

**Lemma 13** *(Beta-Delta Saving Rule) The optimal saving rules for the β-δ model are given by*

$$s_1^{bd} = B_1(I_1 + Rs_0) \qquad \textit{where } B_1 : \frac{1}{1+(\beta\delta R^\sigma)^{\frac{1}{\sigma-1}}}$$

*and*

$$s_0^{bd} = \frac{I_0 - (\beta\delta R B_1)^{\frac{1}{\sigma-1}} I_1}{1+(\beta\delta R B_1)^{\frac{1}{\sigma-1}} R}, \qquad \textit{where } B_0 := R^\sigma \frac{\left((\beta\delta)^{\frac{1}{\sigma-1}}R^{\frac{1}{\sigma-1}}\right)^\sigma + \delta}{\left(1+(\beta\delta)^{\frac{1}{\sigma-1}}R^{\frac{\sigma}{\sigma-1}}\right)^\sigma}.$$

**Proof.** Self 1 solves

$$\max_{s_1} \ (I_1 + Rs_0 - s_1)^\sigma + \beta\delta(Rs_1)^\sigma.$$

The FOC is

$$\sigma(I_1 + Rs_0 - s_1)^{\sigma-1} = \beta\delta\sigma R^\sigma {s_1}^{\sigma-1}$$

yielding the saving rule

$$s_1^{bd} = \frac{I_1 + Rs_0}{1+(\beta\delta)^{\frac{1}{\sigma-1}}R^{\frac{\sigma}{\sigma-1}}}.$$

Self 0 solves

$$
\begin{aligned}
&\max_{s_0}\ (I_0 - s_0)^\sigma + \beta[\delta(I_1 + Rs_0 - s_1^{bd})^\sigma + \delta^2(Rs_1^{bd})^\sigma] \\
\iff &\max_{s_0}\ (I_0 - s_0)^\sigma + \beta[\delta(I_1 + Rs_0 - \frac{I_1 + Rs_0}{1 + (\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}})^\sigma + \delta^2(R\frac{I_1 + Rs_0}{1 + (\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}})^\sigma] \\
\iff &\max_{s_0}\ (I_0 - s_0)^\sigma + \beta[\delta\left((I_1 + Rs_0)\frac{(\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}}{1 + (\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}}\right)^\sigma + \delta^2\left(R\frac{I_1 + Rs_0}{1 + (\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}}\right)^\sigma] \\
\iff &\max_{s_0}\ (I_0 - s_0)^\sigma + \beta\delta\left(\frac{I_1 + Rs_0}{1 + (\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}}\right)^\sigma \left[\left((\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}\right)^\sigma + \delta R^\sigma\right] \\
\iff &\max_{s_0}\ (I_0 - s_0)^\sigma + \beta\delta(I_1 + Rs_0)^\sigma B_0,
\end{aligned}
$$

where

$$
B_0 := \frac{\left((\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}\right)^\sigma + \delta R^\sigma}{\left(1 + (\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}\right)^\sigma} = R^\sigma \frac{\left((\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{1}{\sigma-1}}\right)^\sigma + \delta}{\left(1 + (\beta\delta)^{\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}\right)^\sigma}.
$$

The FOC is

$$
\sigma(I_0 - s_0)^{\sigma-1} = \beta\delta R\sigma(I_1 + Rs_0)^{\sigma-1} B_1
$$

yielding

$$
s_0^{bd} = \frac{I_0 - (\beta\delta R B_1)^{\frac{1}{\sigma-1}} I_1}{1 + (\beta\delta R B_1)^{\frac{1}{\sigma-1}} R}.
$$

■

**Lemma 14** *The Sophisticated CCE and Sophisticated β-δ models are not observationally equivalent.*

**Proof.** Comparing saving rules it must be that $A_t = B_t$ for $t = 0, 1$. The condition $A_1 = B_1$ is equivalent to

$$
\beta\delta = \left(\frac{K}{a_1}\right)^{\frac{1}{m}}.
$$

Given this, $A_0 = B_0$ is equivalent to

$$
\delta + (\beta\delta R)^{\frac{\sigma}{\sigma-1}} = \left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} + (\beta\delta R)^{\frac{\sigma\theta}{\sigma-1}}\right)^{\frac{1}{\theta}}.
$$

Therefore, we have a closed-form solution which must be nonconstant in $R$:

$$
\delta(R) = \left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} + \left(\frac{K}{a_1}\right)^{\frac{\sigma\theta}{m(\sigma-1)}} R^{\frac{\sigma\theta}{\sigma-1}}\right)^{\frac{1}{\theta}} - \left(\frac{K}{a_1}\right)^{\frac{\sigma}{m(\sigma-1)}} R^{\frac{\sigma}{\sigma-1}},
$$

$$\beta(R) = \left(\frac{K}{a_1}\right)^{\frac{1}{m}} \frac{1}{\delta(R)}.$$

Since $x^{\frac{1}{\theta}}$ is strictly increasing, $\delta(R) > 0$ and so the solution is well-defined. ■

# 5 Proof of Proposition 9

**Lemma 15** *(Naive CCE Saving Rule) Assume* $0 < \sigma \frac{m}{m-1} < 1$ *and suppose that both self 0 and self 1 are cognitively constrained at their respective optimal consumption path. The optimal saving rules for the sophisticated CCE model are given by*

$$s_1^* = A_1(I_1 + Rs_0) \qquad \textit{where } A_1 := \frac{1}{1 + [\left(\frac{K}{a_1}\right)^{\frac{1}{m}} R^{\sigma}]^{\frac{1}{\sigma-1}}},$$

*and*

$$s_0^* = \frac{I_0 - (\left(\frac{K}{a_1}\right)^{\frac{1}{m}} \tilde{A}_0 R)^{\frac{1}{\sigma-1}} I_1}{1 + (\left(\frac{K}{a_1}\right)^{\frac{1}{m}} \tilde{A}_0 R)^{\frac{1}{\sigma-1}} R} \qquad \textit{where } \tilde{A}_0 := R^{\sigma} \frac{\left(\left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R\right)^{\frac{\sigma\theta}{\sigma\theta-1}} + \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}}\right)^{\frac{1}{\theta}}}{\left(1 + \left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R^{\sigma\theta}\right)^{\frac{1}{\sigma\theta-1}}\right)^{\sigma}}.$$

**Proof.** The calculations follow those in Lemma 11. The solution for $s_1^*$ is as in Lemma 11. Solve for self 0's optimal consumption assuming that the solution satisfies $\delta_1 c_1^{\sigma\theta} + \delta_2 c_2^{\sigma\theta} > m^{\theta} K$.

$$\max_{s_0, s_1} (I_0 - s_0)^{\sigma} + (K)^{\frac{1}{m}} [\delta_1 (I_1 + Rs_0 - s_1)^{\sigma\theta} + \delta_2 (Rs_1)^{\sigma\theta}]^{\frac{1}{\theta}}.$$

The FOC for $s_1$ is

$$\sigma\theta\delta_1 (I_1 + Rs_0 - s_1)^{\sigma\theta-1} = \delta_2 \sigma\theta R^{\sigma\theta} {s_1}^{\sigma\theta-1}$$

yielding $s_1 = \tilde{A}_1(I_1 + Rs_0)$ where $\tilde{A}_1 = \left(1 + \left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R^{\sigma\theta}\right)^{\frac{1}{\sigma\theta-1}}\right)^{-1}$. Self 0's optimization problem is reduced to

$$\max_{s_0} (I_0 - s_0)^{\sigma} + K^{\frac{1}{m}} [\delta_1 (1 - \tilde{A}_1)^{\sigma\theta} (I_1 + Rs_0)^{\sigma\theta} + \delta_2 (R\tilde{A}_1)^{\sigma\theta} (I_1 + Rs_0)^{\sigma\theta}]^{\frac{1}{\theta}}$$

$$\iff \max_{s_0} (I_0 - s_0)^{\sigma} + (K a_1^{-1})^{\frac{1}{m}} [(1 - \tilde{A}_1)^{\sigma\theta} + \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} (R\tilde{A}_1)^{\sigma\theta}]^{\frac{1}{\theta}} (I_1 + Rs_0)^{\sigma}$$

$$\iff \max_{s_0} (I_0 - s_0)^{\sigma} + (K a_1^{-1})^{\frac{1}{m}} \tilde{A}_0 (I_1 + Rs_0)^{\sigma},$$

where

$$
\begin{aligned}
\tilde{A}_0 =& [(1-\tilde{A}_1)^{\sigma\theta} + \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} (R\tilde{A}_1)^{\sigma\theta}]^{\frac{1}{\theta}} \\
=& \left( \left( \frac{\left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R^{\sigma\theta}\right)^{\frac{1}{\sigma\theta-1}}}{1+\left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R^{\sigma\theta}\right)^{\frac{1}{\sigma\theta-1}}} \right)^{\sigma\theta} + \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} \left( \frac{R}{1+\left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R^{\sigma\theta}\right)^{\frac{1}{\sigma\theta-1}}} \right)^{\sigma\theta} \right)^{\frac{1}{\theta}} \\
=& R^{\sigma} \frac{\left( \left( \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R \right)^{\frac{\sigma\theta}{\sigma\theta-1}} + \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} \right)^{\frac{1}{\theta}}}{\left(1+\left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R^{\sigma\theta}\right)^{\frac{1}{\sigma\theta-1}}\right)^{\sigma}},
\end{aligned}
$$

which yields the time 0 saving rule

$$
s_0^* = \frac{I_0 - (\left(\frac{K}{a_1}\right)^{\frac{1}{m}} \tilde{A}_0 R)^{\frac{1}{\sigma-1}} I_1}{1+(\left(\frac{K}{a_1}\right)^{\frac{1}{m}} \tilde{A}_0 R)^{\frac{1}{\sigma-1}} R}.
$$

■

The marginal propensity to save in period 1 is given by

$$
\frac{ds_1^*}{d(I_1+Rs_0)} = \frac{1}{1+\left(\frac{K}{a_1}\right)^{\frac{1}{m}\frac{1}{\sigma-1}} R^{\frac{\sigma}{\sigma-1}}}.
$$

Marginal propensity to saving in period 0 is given by

$$
\frac{ds_0^*}{dI_0} = \frac{1}{1+(\left(\frac{K}{a_1}\right)^{\frac{1}{m}} R\tilde{A}_0)^{\frac{1}{\sigma-1}} R}.
$$

We need to show that

**Lemma 16** $\frac{ds_1^*}{d(I_1+Rs_0)} > \frac{ds_0^*}{dI_0}$ *for any* $I_0, I_1 > 0$.

**Proof.** Since $\frac{ds_1^*}{d(I_1+Rs_0)} > \frac{ds_0^*}{dI_0} \iff \tilde{A}_0 > 1$, it suffices to show that $\tilde{A}_0 > 1$. Compute that

$$
\tilde{A}_0 = R^{\sigma} \frac{\left( \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} + \left( \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R \right)^{\frac{\sigma\theta}{\sigma\theta-1}} \right)^{\frac{1}{\theta}}}{\left(1+\left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R^{\sigma\theta}\right)^{\frac{1}{\sigma\theta-1}}\right)^{\sigma}} = R^{\sigma} \left(\frac{a_1}{a_2}\right)^{\frac{1}{m}} \left(1+\left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}} R^{\sigma\theta}\right)^{\frac{1}{\sigma\theta-1}}\right)^{\frac{1-\sigma\theta}{\theta}}
$$

$$= R^\sigma \left(\frac{a_1}{a_2}\right)^{\frac{1}{m}} \frac{\left(1+\left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m}} R^\sigma\right)^{\frac{\theta}{1-\sigma\theta}}\right)^{\frac{1-\sigma\theta}{\theta}}}{\left(\frac{a_1}{a_2}\right)^{\frac{1}{m}} R^\sigma} = \left(1+\left(\left(\frac{a_1}{a_2}\right)^{\frac{1}{m}} R^\sigma\right)^{\frac{\theta}{1-\sigma\theta}}\right)^{\frac{1-\sigma\theta}{\theta}} > 1.$$

■

# 6 Proof of Proposition 10

Consider an arbitrary stream $z \in X$. The first order conditions of the optimization problem yields that for any stream $z$ the optimal $D_z$ must satisfy

$$\frac{u(z_t)}{u(z_{t'})} = \frac{\varphi_t'(D_z(t))}{\varphi_{t'}'(D_z(t'))} \qquad \forall t, t' > 0 \ s.t.\ u(z_t), u(z_{t'}) > 0,$$

and the capacity constraint $\sum_{t>0} \varphi_t(D_z(t)) \leq K$. Since, $D_z(t) = 0$ for any $t > 0$ s.t. $u(z_t) = 0$, to establish our result we can wlog focus on $z$ for which $u(z_t) > 0$ for all $t > 0$. Consider the following *truncated optimization problem* where we treat $D_z(\tau)$ as fixed:

$$\max_{(D(t))_{0<t\neq\tau}} \{\left[\sum_{0<t\neq\tau} D(t)u(z_t) - \sum_{0<t\neq\tau} \varphi_t(D(t))\right] + [D_z(\tau)u(z_\tau) - \varphi_\tau(D_z(\tau))]\}$$
$$\text{s.t.} \sum_{0<t\neq\tau} \varphi_t(D(t)) \leq K_z^{trunc} := K - \varphi_\tau(D_z(\tau)).$$

Define $K_z^{trunc} := K - \varphi_\tau(D_z(\tau))$. Since the terms with $D_z(\tau)$ are constants, the problem is equivalent to

$$\max_{(D(t))_{0<t\neq\tau}} \{\sum_{0<t\neq\tau} D(t)u(z_t) - \sum_{0<t\neq\tau} \varphi_t(D(t))\} \text{ s.t. } \sum_{0<t\neq\tau} \varphi_t(D(t)) \leq K_z^{trunc}.$$

It is clear that the solution $(D_z^{trunc}(t))_{0<t\neq\tau}$ to this problem coincides with the solution $D_z$ (restricted to $0 < t \neq \tau$) to the original problem with $z$. We use these observations to prove the proposition.

Take any $x$ and $\epsilon^\tau$ and wlog suppose $u(x_t) > 0$ for all $t$ and $\epsilon > 0$. Denote $y = x + \epsilon^\tau$. By the preceding, when restricted to $0 < t \neq \tau$, $D_x$ and $D_y$ solve a truncated optimization problem with different the same objective function but different constraints. We show that $x$ has a more relaxed problem, that is, $K_y^{trunc} \leq K_x^{trunc}$ .

Suppose by way of contradiction that $K_y^{trunc} > K_x^{trunc}$. This implies that $\varphi_\tau(D_y(\tau)) < \varphi_\tau(D_x(\tau))$ which implies $D_y(\tau) < D_x(\tau)$. By the FOC ratios, we also have that for any $0 < t \neq \tau$

$$\frac{\varphi_t'(D_y(t))}{\varphi_t'(D_y(\tau))} = \frac{u(x_t)}{u(x_\tau + \epsilon)} < \frac{u(x_t)}{u(x_\tau)} = \frac{\varphi_t'(D_x(t))}{\varphi_\tau'(D_x(\tau))}.$$

But then $D_y(\tau) < D_x(\tau)$ implies $D_y(t) < D_x(t)$ for all $0 < t \neq \tau$. This contradicts the optimality of $D_y$ in the original problem for $y$, since $D_x$ is feasible, $\sum_{t>0} \varphi_t(D_x(t)) \leq K$, and improves the value of the objective function,

$$\left[\sum_{0<t\neq\tau} D_x(t)u(x_t) - \sum_{0<t\neq\tau} \varphi_t(D_x(t))\right] + \left[D_x(\tau)u(x_\tau + \epsilon) - \varphi_\tau(D_x(\tau))\right]$$

$$> \left[\sum_{0<t\neq\tau} D_y(t)u(x_t) - \sum_{0<t\neq\tau} \varphi_t(D_y(t))\right] + \left[D_y(\tau)u(x_\tau + \epsilon) - \varphi_\tau(D_y(\tau))\right].$$

Having established that $K_y^{trunc} \leq K_x^{trunc}$ we can now prove the result. The inequality directly implies $D_{x+\epsilon^\tau}(\tau) \geq D_x(\tau)$. Because $x$ has a more relaxed truncated problem, it must be that $D_{x+\epsilon^\tau}(t) \leq D_x(t)$ for all $0 < t \neq \tau$.

# 7 Proof of Proposition 11

Consider the CE agent. If there is one task, then by assumption the agent completes the task. Assume the induction hypothesis that the agent would complete $n-1$ tasks with deadline $T_{n-1} = 2(n-1) - 2 = 2n - 4$. Suppose that there are $n$ tasks to be completed with deadline $T_n = 2n - 2$. If the agent does not do a task at time 0 then her problem from the next period on is that of $n-1$ tasks to be completed in $T_{n-1}$ periods, all of which she will complete given the induction hypothesis. Her time 0 problem compares the discounted utility of doing vs not doing a task:

$$U(n \text{ tasks}) = \left[u(0) + \sum_{t=2,4,\cdots,2n} D_{u(0)}(t)u(0)\right] + D_{u(R)}(1)u(R) + \sum_{t=3,\cdots,2n+1} D_{u(R)}(t)u(R),$$

$$U(n-1 \text{ tasks}) = \left[u(r) + \sum_{t=2,4,\cdots,2n} D_{u(0)}(t)u(0)\right] + \sum_{t=1,3,\cdots,2n+1} D_{u(R)}(t)u(R).$$

However,

$$U(n \text{ tasks}) - U(n-1 \text{ tasks}) = -u(r) + D_{u(R)}(1)u(R).$$

But $-u(r) + D_{u(r)}(1)u(R)$ is the net utility of doing one task where there is one to be done at $T = 0$, which is positive by hypothesis. Therefore $U(n \text{ tasks}) \geq U(n-1 \text{ tasks})$, and the CE agent would complete $n$ tasks with deadline $T = 2n - 2$.

We show how the above proof breaks down for the CCE agent. Since she may be cognitively constrained we write the discount functions in their more general form:

$$U(n \text{ tasks}) = \left[u(0) + \sum_{t=2,4,\cdots,2n} D_n(t)u(0)\right] + D_n(1)u(R) + \sum_{t=3,\cdots,2n+1} D_n(t)u(R),$$

$$U(n-1 \text{ tasks}) = \left[u(r) + \sum_{t=2,4,\cdots,2n} D_{n-1}(t)u(0)\right] + \sum_{t=1,3,\cdots,2n+1} D_{n-1}(t)u(R).$$

Note that the $n$ stream has a positive reward at $t = 1$ while the $n-1$ stream does not. If the agent is not cognitively constrained at $n$ then she is not cognitively constrained at $n-1$ and by the argument in the CE case we see that she will complete $n$ tasks. On the other hand, if the agent is cognitively constrained at $n$, then regardless of whether she is constrained at $n-1$, the model implies $D_n(t) \leq D_{n-1}(t)$ for all $t$ (BY PROP) since the cognitive resources have to be spread over more periods in case $n$. Consequently,

$$\begin{aligned} &U(n \text{ tasks}) - U(n-1 \text{ tasks}) \\ =& -u(r) + D_{u(r)}(1)u(R) + \left[\sum_{t=2,4,\cdots,2n} [D_n(t) - D_{n-1}(t)]u(0) + \sum_{t=1,3,\cdots,2n+1} [D_n(t) - D_{n-1}(t)]u(R)]\right], \end{aligned}$$

where the term in the square brackets is negative (since $D_n(t) \leq D_{n-1}(t)$ and $u(r) > 0 = u(0)$). With an appropriate choice of parameters we can obtain $U(n \text{ tasks}) - U(n-1 \text{ tasks}) < 0$.

To demonstrate this, and to also show the possibility of cycles, we construct an example with a Homogeneous CCE agent. Suppose there are $n = 3$ tasks to be done by period $T = 4$. It will be convenient to define

$$A_t(\beta) := \left(\frac{1}{ma_t}\right)^{\frac{1}{m-1}} u(\beta)^{\frac{m}{m-1}},$$

and

$$f(\alpha) = (mK)^{\frac{1}{m}} \{\alpha\}^{\frac{m-1}{m}}.$$

At $T = 4$, suppose the agent is constrained when considering the task. By the reduced form of the model, we therefore require that

$$A_1(R) = (ma_1)^{-\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} > mK.$$

Then the utility from doing the task is $f(A_1(R)) = (mK)^{\frac{1}{m}} \left\{ \left(\frac{1}{ma_1}\right)^{\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}}$. Thus at $T = 4$ constrained and agent does the task iff

$$f(A_1(R)) \geq u(r).$$

At $T = 2$ suppose that the agent is constrained when considering not doing the task, that is,

$$A_1(r) + A_3(R) = (ma_1)^{-\frac{1}{m-1}} u(r)^{\frac{m}{m-1}} + (ma_3)^{-\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} > mK.$$

This implies that she is also constrained when considering the task, that is,

$$A_1(R) + A_3(R) = (ma_1)^{-\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} + (ma_3)^{-\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} > mK.$$

Then at $T = 2$ she does not do the task iff

$$\begin{aligned}
&(mK)^{\frac{1}{m}} \left\{ \left(\frac{1}{ma_1}\right)^{\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} + \left(\frac{1}{ma_3}\right)^{\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}} \\
&\quad < u(r) + (mK)^{\frac{1}{m}} \left\{ \left(\frac{1}{ma_3}\right)^{\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}} \\
&\iff f\left(A_1(R) + A_3(R)\right) - f\left(A_3(R)\right) < u(r).
\end{aligned}$$

At $T = 0$ suppose she is constrained when considering not doing the task:

$$A_2(r) + A_5(R) = (ma_2)^{-\frac{1}{m-1}} u(r)^{\frac{m}{m-1}} + (ma_5)^{-\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} > mK.$$

This implies that she is constrained when considering the task:

$$A_1(R)+A_2(r)+A_5(R) = (ma_1)^{-\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} + (ma_2)^{-\frac{1}{m-1}} u(r)^{\frac{m}{m-1}} + (ma_5)^{-\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} > mK.$$

Then at $T = 0$ she does the task iff

$$\begin{aligned}
&(mK)^{\frac{1}{m}} \left\{ \left(\frac{1}{ma_1}\right)^{\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} + \left(\frac{1}{ma_2}\right)^{\frac{1}{m-1}} u(r)^{\frac{m}{m-1}} + \left(\frac{1}{ma_5}\right)^{\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}} \\
&\quad \geq u(r) + (mK)^{\frac{1}{m}} \left\{ \left(\frac{1}{ma_2}\right)^{\frac{1}{m-1}} u(r)^{\frac{m}{m-1}} + \left(\frac{1}{ma_5}\right)^{\frac{1}{m-1}} u(R)^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}} \\
&\iff f\left(A_1(R) + A_2(r) + A_5(R)\right) - f\left(A_2(r) + A_5(R)\right) \geq u(r).
\end{aligned}$$

Therefore we need the following inequalities to hold: those that involve being constrained,

$$\begin{aligned}
A_1(R) &> mK, \\
A_1(r) + A_3(R) &> mK, \\
A_2(r) + A_5(R) &> mK,
\end{aligned}$$

and those that involve choice:

$$f(A_1(R)) - f(0) \geq u(r),$$

$$\begin{aligned}
f\left(A_1(R) + A_3(R)\right) - f\left(A_3(R)\right) &< u(r), \\
f\left(A_1(R) + A_2(r) + A_5(R)\right) - f\left(A_2(r) + A_5(R)\right) &\geq u(r).
\end{aligned}$$

Assume that $A_5 \approx 0$ (by taking $a_5$ arbitrarily large). Then a sufficient condition for the "constraint inequalities" to be satisfied is that $A_2(r) > mK$, that is,

$$\left(\frac{1}{ma_2}\right)^{\frac{1}{m-1}} u(r)^{\frac{m}{m-1}} > mK.$$

Among the "choice inequalities", since $f$ is concave, the third implies the first, so the latter can be dropped. We therefore need to show that the following inequalities can be satisfied numerically:

$$\left(\frac{1}{ma_2}\right)^{\frac{1}{m-1}} u(r)^{\frac{m}{m-1}} > mK,$$

$$(mK)^{\frac{1}{m}} \{A_1(R) + A_3(R)\}^{\frac{m-1}{m}} - (mK)^{\frac{1}{m}} \{A_3(R)\}^{\frac{m-1}{m}} < u(r),$$

$$(mK)^{\frac{1}{m}} \{A_1(R) + A_2(r)\}^{\frac{m-1}{m}} - (mK)^{\frac{1}{m}} \{A_2(r)\}^{\frac{m-1}{m}} > u(r).$$

Assume $m = 2$ and insert the definition of $A$ in these inequalities, in which case they take the form:

$$\left(\frac{1}{2a_2}\right) u(r)^2 > 2K,$$

$$\left\{\left(\frac{1}{a_1}\right) + \left(\frac{1}{a_3}\right)\right\}^{\frac{1}{2}} - \left\{\left(\frac{1}{a_3}\right)\right\}^{\frac{1}{2}} < \frac{u(r)}{u(R)K^{\frac{1}{2}}},$$

$$\left\{\left(\frac{1}{a_1}\right) + \left(\frac{1}{a_2}\right)\frac{u(r)^2}{u(R)^2}\right\}^{\frac{1}{2}} - \left\{\left(\frac{1}{a_2}\right)\frac{u(r)^2}{u(R)^2}\right\}^{\frac{1}{2}} > \frac{u(r)}{u(R)K^{\frac{1}{2}}}.$$

We have taken $a_5 \to \infty$. Arbitrarily fix $a_1 < .. < a_4$. Take any $\gamma$ that lies strictly between

$$\left\{\left(\frac{1}{a_1}\right) + \left(\frac{1}{a_3}\right)\right\}^{\frac{1}{2}} - \left\{\left(\frac{1}{a_3}\right)\right\}^{\frac{1}{2}} < \left\{\left(\frac{1}{a_1}\right) + 0\right\}^{\frac{1}{2}} - \{0\}^{\frac{1}{2}},$$

and take a sufficiently small $\theta > 0$ s.t.

$$\left\{\left(\frac{1}{a_1}\right) + \left(\frac{1}{a_2}\right)\theta^2\right\}^{\frac{1}{2}} - \left\{\left(\frac{1}{a_2}\right)\theta^2\right\}^{\frac{1}{2}} > \gamma.$$

Let $K = \left(\frac{\theta}{\gamma}\right)^2$. Pick $u(r) > 0$ satisfying the first inequality, $\left(\frac{1}{2a_2}\right) u(r)^2 > 2K$. Choose $u(R)$ so that $\theta = \frac{u(r)}{u(R)} < 1$. Then we have

$$\left\{\left(\frac{1}{a_1}\right) + \left(\frac{1}{a_3}\right)\right\}^{\frac{1}{2}} - \left\{\left(\frac{1}{a_3}\right)\right\}^{\frac{1}{2}} < \gamma < \left\{\left(\frac{1}{a_1}\right) + \left(\frac{1}{a_2}\right)\theta^2\right\}^{\frac{1}{2}} - \left\{\left(\frac{1}{a_2}\right)\theta^2\right\}^{\frac{1}{2}}$$

and in particular, all the desired inequalities are satisfied.

This establishes the claim that the agent may complete all tasks but not consecutively. To show that cycling is possible, assume that $a_t \to \infty$ for all $t \geq 5$. Therefore the agent's "horizon" is effectively 4 periods. Repeating the above example with arbitrary $n$ and $T = 2n - 2$ therefore establishes that cycling is possible and completes the proof.

# 8 Proof of Proposition 12

It is important to first note that under the given assumption (that is, the agent would complete 1 task today when the deadline is $T = 2$), it must be that the agent would complete the task at $T = 2$. If not, then the utilities for the $t = 0$ self are

$$U(\text{complete 1 task at time 0}) = 0 + D_{u(R)}(1)u(R) + D_{u(r)}(2)u(r) + 0$$

$$U(\text{complete 1 task at time 2}) = u(r) + 0 + D_{u(r)}(2)u(r) + 0,$$

and the agent compares $D_{u(R)}(1)u(R)$ with $u(r)$, just as the $t = 0$ self does, and therefore will not do the task, contradicting the given assumption. Therefore self $T = 2$ will complete the task if there is 1 to be done.

Consider the CE agent, and suppose there is 1 task to be completed by time $T = 2$. Then the agent compares

$$U(\text{complete 1 task at time 0}) = u(0) + D_{u(R)}(1)u(R) + D_{u(r)}(2)u(r) + 0$$

$$U(\text{do not complete 1 task at time 0}) = u(r) + 0 + 0 + D_{u(R)}(3)u(R),$$

where our opening observation is used to note that if the agent does not do the task at time 0 then the time 2 self will wish to complete the task in the deadline period 2. By assumption, the agent would prefer to do the task at time 0.

Then the Separability property of the CE model and an induction argument yields that the agent will do 1 task immediately regardless of the deadline. To see this, consider the case where the deadline is $T = 4$. The problem of the $t = 2$ self is identical to the above problem and so she will prefer to do the task, and the utility calculations for the $t = 0$ self are the same as above except that there is a common term $D_{u(r)}(4)u(r) + 0$ (since self 4 does not have any task to do) appended to both utilities. Following this line of reasoning, an induction argument yields that 1 task will be completed immediately regardless of the deadline.

Turn to the CCE agent. Suppose by way of contradiction that the agent would do the task immediately for any $T \geq 2$. Then for all such $T$, the $t = 0$ self finds that $U(\text{1 task at time 0}) > U(\text{1 task at time 2})$ where

$$U(\text{complete 1 task at time 0}) = 0 + D_{0,T}(1)u(R) + D_{0,T}(2)u(r) + 0 + \sum_{t=4,\cdots,T} D_{0,T}(t)u(r),$$

$$U(\text{complete 1 task at time 2}) = u(r) + 0 + 0 + D_{2,T}(3)u(R) + \sum_{t=4,\cdots,T} D_{t,T}(t)u(r),$$

where $D_{\tau,T}$ denotes the optimal discount function for the stream where the task is done at $\tau$ and the deadline is $T$. The hypothesis that the agent would always do the task immediately implies

$$D_{0,T}(1)u(R) + D_{0,T}(2)u(r) \geq u(r) + D_{2,T}(3)u(R)$$

for all $T$. However, if we consider a model with time-independent cost function $\varphi_t(d) = d^m$, then for $\tau = 0, 2$, the discount function satisfies $D_{\tau,T}(t) = D_{\tau,T}(t')$ for all $t, t' > 0$ that have the same reward. Denote $D_{\tau,T}(t) = d_{r,\tau,T}$ for periods $4, 6, ...$ where the reward is $r$. Then the capacity constraint implies:

$$\frac{T-4}{2}\varphi(d_{r,\tau,T}) = \sum_{t=4,6,\cdots,T} \varphi(D_{\tau,T}(t)) \leq \sum_{t=1}^{T+1} \varphi_t(D_{\tau,T}(t)) \leq K.$$

As $T \to \infty$ it must therefore be that $D_{0,T}, D_{2,T} \to 0$ and therefore the above inequality cannot hold for all $T$, a contradiction.

# References


[1] Noor, J. and N. Takeoka (2023): "Constrained Optimal Discounting, " Working paper.