Supplementary Appendix for "Constrained Optimal Discounting"*

Jawwad Noor   Norio Takeoka

May 30, 2026

Abstract

This supplementary appendix for Noor and Takeoka [4] presents an addtional difference between our model and the beta-delta model, compares our model with Wakai model [5], and provides the proofs omitted in the paper.

1 Consumption-Saving Application

1.1 CCE-DU Observational Non-Equivalence

Previous literature on the beta-delta model (Laibson [2], Barro [1]) has shown that the beta-delta model (with beta strictly less than 1 and log utility) is observationally equivalent to the standard exponential discounting model when the data consists only of consumption-saving behavior. In this section, we ask if the consumption-savings profile of a Sophisticated CCE agent can be replicated by a Sophisticated beta-delta model (with beta allowed to equal 1, so as to nest the standard exponential discounting model). It is readily shown that:

Proposition. *Suppose $0 < \sigma\frac{m}{m-1} < 1$ and consider the consumption profiles of a Sophisticated CCE agent with $K_0 = K_1$ that is constrained along the optimal consumption paths at any two distinct $R, R' > 1$. There does not exist a Sophisticated beta-delta model (with utility index u) that can simultaneously match these consumption profiles.*

*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 Numbers JP19KK0308 and JP25K00619. 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.

Since the standard exponential discounting model is a special case of the Sophisticated beta-delta model, the result establishes observational non-equivalence with the standard model as well. The reason is that the nonseparability of the CCE model comes into play as R varies. In the proof, we show that any calibrating beta-delta model must strictly change with R. While the algebra in the proof is greatly simplified by assuming that the beta-delta model has the same u as the CCE model, we expect the non-equivalence to hold even if we have freedom to choose the utility index for the calibrating model.

1.2 Proof

Lemma 1. *(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

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

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

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

$\square$

Lemma 2. *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. $\square$

2 Proofs in Section 4 (Consumption-Savings)

2.1 Proof of Proposition 4

Lemma 3. *(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*

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

and

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

Proof. Write $0 < \sigma\theta < 1$, where $\theta = \frac{m}{m-1}$. Let $\gamma_t = (a_t m)^{-\frac{1}{m-1}}$. By Proposition 3, 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,

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

where $\delta_t = (a_t^{-1})^{\frac{1}{m-1}}$. Note that $\delta_2 \le \delta_1 \le 1$. 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

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

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

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

□

Lemma 4. *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

$$h(k) := \frac{(1 + \zeta R^{\frac{\sigma\theta}{1-\sigma}} k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}}}{(1 + R^{\frac{\sigma}{1-\sigma}} k^{\frac{1}{1-\sigma}})^{\sigma}}. \tag{2}$$

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

$$
\begin{aligned}
&\frac{\sigma}{1-\sigma}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}}(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})^{\sigma} \\
&\times\left((1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})^{-1}\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}-1}-(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})^{-1}R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}-1}\right) \\
\iff &\frac{\sigma}{1-\sigma}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}}(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})^{\sigma} \\
&\times\frac{R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}-1}\zeta(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})-R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}-1}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})}{(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})} \\
\iff &\frac{\sigma}{1-\sigma}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}}(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})^{\sigma} \\
&\times\frac{\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}-1}-R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}-1}}{(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})} \\
\iff &\frac{\sigma}{1-\sigma}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})^{\frac{1}{\theta}-1}(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})^{\sigma-1}\left(R^{\frac{\sigma}{1-\sigma}}(\zeta R^{\frac{\sigma(\theta-1)}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}-1}-k^{\frac{1}{1-\sigma}-1}\right). \qquad (3)
\end{aligned}
$$

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

$$
\begin{aligned}
&\zeta R^{\frac{\sigma(\theta-1)}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}-1}>k^{\frac{1}{1-\sigma}-1} \\
\iff &\zeta>R^{-\frac{\sigma(\theta-1)}{1-\sigma}}k^{\frac{1-\sigma\theta}{1-\sigma}} \\
\iff &\left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}}>R^{-\frac{\sigma}{1-\sigma}\frac{1}{m-1}}\left(\frac{K_1}{a_1}\right)^{\frac{(1-\sigma)m-1}{m(m-1)(1-\sigma)}} = \left(R^{-\sigma}\left(\frac{K_1}{a_1}\right)^{\frac{(1-\sigma)m-1}{m}}\right)^{\frac{1}{(m-1)(1-\sigma)}} \\
\iff &\frac{a_1}{a_2}>\left(R^{-\sigma}\left(\frac{K_1}{a_1}\right)^{1-\sigma-\frac{1}{m}}\right)^{\frac{1}{1-\sigma}}. \qquad (4)
\end{aligned}
$$

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 (2). Since the numerator of the derivative $h'(k)$ is derived as in (3),

$$
\begin{aligned}
[kh(k)]' &= h(k) + kh'(k) = h(k)\left(1 + k\frac{h'(k)}{h(k)}\right) \\
&= h(k)\left(1 + k\frac{\sigma}{1-\sigma}R^{\frac{\sigma}{1-\sigma}}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})^{-1}(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})^{-1}\left(\zeta R^{\frac{\sigma(\theta-1)}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}-1} - k^{\frac{1}{1-\sigma}-1}\right)\right) \\
&= h(k)\left(1 + \frac{\sigma}{1-\sigma}R^{\frac{\sigma}{1-\sigma}}(1+\zeta R^{\frac{\sigma\theta}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}})^{-1}(1+R^{\frac{\sigma}{1-\sigma}}k^{\frac{1}{1-\sigma}})^{-1}\left(\zeta R^{\frac{\sigma(\theta-1)}{1-\sigma}}k^{\frac{\sigma\theta}{1-\sigma}} - k^{\frac{1}{1-\sigma}}\right)\right)
\end{aligned} \tag{5}
$$

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 (4). If this inequality holds, (5) is positive. $\square$

2.2 Proof of Proposition 5

Lemma 5. *(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 3. The solution for s_1^* is as in Lemma 3. 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

$$
\begin{aligned}
&\max_{s_0}\ (I_0-s_0)^\sigma + K^{\frac{1}{m}}[\delta_1(1-\tilde{A}_1)^{\sigma\theta}(I_1+Rs_0)^{\sigma\theta} + \delta_2(R\tilde{A}_1)^{\sigma\theta}(I_1+Rs_0)^{\sigma\theta}]^{\frac{1}{\theta}} \\
\Longleftrightarrow\quad &\max_{s_0}\ (I_0-s_0)^\sigma + (Ka_1^{-1})^{\frac{1}{m}}[(1-\tilde{A}_1)^{\sigma\theta} + \left(\frac{a_1}{a_2}\right)^{\frac{1}{m-1}}(R\tilde{A}_1)^{\sigma\theta}]^{\frac{1}{\theta}}(I_1+Rs_0)^\sigma \\
\Longleftrightarrow\quad &\max_{s_0}\ (I_0-s_0)^\sigma + (Ka_1^{-1})^{\frac{1}{m}}\tilde{A}_0(I_1+Rs_0)^\sigma,
\end{aligned}
$$

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

$$
s_0^* = \frac{I_0 - (\left(\frac{K}{a_1}\right)^{\frac{1}{m}}\tilde{A}_0R)^{\frac{1}{\sigma-1}}I_1}{1+(\left(\frac{K}{a_1}\right)^{\frac{1}{m}}\tilde{A}_0R)^{\frac{1}{\sigma-1}}R}.
$$

□

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 6. *$\frac{ds_1^*}{d(I_1+Rs_0)} > \frac{ds_0^*}{dI_0}$ 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

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

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

□

2.3 Proof of Proposition 6

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:

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

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

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

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 ϵ^τ 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 $0<t\neq\tau$.

3 Proofs in Section 5 (Task Completion)

3.1 Proof of Proposition 7

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 Proposition 8 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

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

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,

$$A_1(R) > mK,$$

$$A_1(r) + A_3(R) > mK,$$
$$A_2(r) + A_5(R) > mK,$$

and those that involve choice:

$$f(A_1(R)) - f(0) \geq u(r),$$
$$f\left(A_1(R) + A_3(R)\right) - f\left(A_3(R)\right) < u(r),$$
$$f\left(A_1(R) + A_2(r) + A_5(R)\right) - f\left(A_2(r) + A_5(R)\right) \geq u(r).$$

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:

$$\left(\frac{1}{ma_2}\right)^{\frac{1}{m-1}} u(r)^{\frac{m}{m-1}} > mK,$$
$$(mK)^{\frac{1}{m}} \left\{A_1(R) + A_3(R)\right\}^{\frac{m-1}{m}} - (mK)^{\frac{1}{m}} \left\{A_3(R)\right\}^{\frac{m-1}{m}} < u(r),$$
$$(mK)^{\frac{1}{m}} \left\{A_1(R) + A_2(r)\right\}^{\frac{m-1}{m}} - (mK)^{\frac{1}{m}} \left\{A_2(r)\right\}^{\frac{m-1}{m}} > u(r).$$

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 γ 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.

3.2 Proof of Proposition 8

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 $U(1 \text{ task at time } 0) > U(1 \text{ task at time } 2)$ where

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

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

where $D_{\tau,T}$ denotes the optimal discount function for the stream where the task is done at τ 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.

4 Comparison with Wakai (2008)

4.1 Implications of Wakai (2008) in the Task Completion

Assume that $R > r > 0$. For $n > 0$, we compare two streams

$$x^n = (0, R, r, 0, z^n), \text{ and } y^n = (r, 0, 0, R, z^n),$$

where z^n is given as an n-repetition of $(r, 0)$, that is, $z^n = (r, 0, r, 0, \cdots, r, 0)$. Let $x^0 = (0, R, r, 0, 0)$ and $y^0 = (r, 0, 0, R, 0)$ with convention.

Let $X = C^{T+2}$ be the space of consumption streams. A generic element is denoted by $x = (x_0, x_1, \cdots, x_T, x_{T+1})$.

Consider a Wakai model [5]: for any $t \in \{0, 1, \cdots, T\}$.

$$U_t(x_t, \cdots, x_{T+1}) = \min_{\delta \in [\underline{\delta}_t, \overline{\delta}_t]} \{(1-\delta)u(x_t) + \delta U_{t+1}(x_{t+1}, \cdots, x_{T+1})\},$$

with convention $U_{T+1} = u$.

Proposition. *Let $T \geq 4$. Assume that $\succsim$ over X is represented by Wakai model U_0. If*

$$\frac{\underline{\delta}_0}{\overline{\delta}_0}\frac{\underline{\delta}_3}{\overline{\delta}_3} \geq \frac{\underline{\delta}_1}{\overline{\delta}_1}\frac{\underline{\delta}_2}{\overline{\delta}_2}, \tag{6}$$

then $x^0 \succ y^0$ implies $x^n \succ y^n$ for all $n > 0$.

Condition (6) holds if:

- Stationarity: $\underline{\delta}_t$ and $\overline{\delta}_t$ are independent of t,
- Discounted utility: $\underline{\delta}_t = \overline{\delta}_t$ for all t,
- Discount factors in periods 1 and 2 are less stable than those in periods 0 and 3, that is, the spreads of $[\underline{\delta}_1, \overline{\delta}_1]$ and $[\underline{\delta}_2, \overline{\delta}_2]$ are relatively larger than those of $[\underline{\delta}_0, \overline{\delta}_0]$ and $[\underline{\delta}_3, \overline{\delta}_3]$.

Proof. Step 1: $W_n = U_4(z^n) < u(r)$. In period T, the agent faces a stream $(r, 0)$. The corresponding utility is $U_T(r, 0) = (1 - \overline{\delta}_T)u(r) < u(r)$.
Next consider period $T - 1$. The agent faces a stream $(0, r, 0)$. The corresponding utility is $U_{T-1}(0, r, 0) = \underline{\delta}_{T-1}(1 - \overline{\delta}_T)u(r) < u(r)$.

By induction, consider period $t \geq 4$ and suppose $U_{t+1}(z^n_{t+1}, \cdots, z^n_{T+1}) < u(r)$. When the agent faces z^n, the payoff in period t is either $u(r)$ or zero. If the payoff in period t is $u(r)$, then $U_t(z^n_t, \cdots, z^n_{T+1}) = \min_{[\underline{\delta}_t, \overline{\delta}_t]}(1 - \delta)u(r) + \delta U_{t+1}(z^n_{t+1}, \cdots, z^n_{T+1}) < u(r)$. If the payoff in period t is zero, then $U_t(z^n_t, \cdots, z^n_{T+1}) = \underline{d}_t U_{t+1}(z^n_{t+1}, \cdots, z^n_{T+1}) < u(r)$. Hence, we have $U_t(z^n_t, \cdots, z^n_{T+1}) < u(r)$, as desired.

Step 2: Evaluation of x^n. Since $W_n > 0$, $\underline{\delta}_3$ is used in period 3, and $U_3(0, z^n) = \underline{\delta}_3 W_n$. Since $u(r) > W_n$ by Step 1, $\overline{\delta}_2$ is used in period 2, and $U_2(r, 0, z^n) = (1 - \overline{\delta}_2)u(r) + \overline{\delta}_2\underline{\delta}_3 W_n$. Since $u(R) > u(r) > (1 - \overline{\delta}_2)u(r) + \overline{\delta}_2\underline{\delta}_3 W_n$, $\overline{\delta}_1$ is used in period 1, and $U_1(R, r, 0, z^n) = (1 - \overline{\delta}_1)u(R) + \overline{\delta}_1((1 - \overline{\delta}_2)u(r) + \overline{\delta}_2\underline{\delta}_3 W_n)$. Finally, since $U_1(R, r, 0, z^n) \geq 0$,

$$U_0(x^n) = \underline{\delta}_0((1 - \overline{\delta}_1)u(R) + \overline{\delta}_1((1 - \overline{\delta}_2)u(r) + \overline{\delta}_2\underline{\delta}_3 W_n)).$$

Step 3: Evaluation of y^n. Since $u(r) > W_n$, $\overline{\delta}_3$ is used in period 3, and $U_3(R, z^n) = (1 - \overline{\delta}_3)u(R) + \overline{\delta}_3 W_n$. Moreover, since this is a positive value, $U_1(0, 0, R, z^n) = \underline{\delta}_1\underline{\delta}_2((1 - \overline{\delta}_3)u(R) + \overline{\delta}_3 W_n)$. Therefore, we have

$$U_0(y^n) = \begin{cases} (1 - \overline{\delta}_0)u(r) + \overline{\delta}_0\underline{\delta}_1\underline{\delta}_2((1 - \overline{\delta}_3)u(R) + \overline{\delta}_3 W_n) & \text{if } u(r) \geq \underline{\delta}_1\underline{\delta}_2((1 - \overline{\delta}_3)u(R) + \overline{\delta}_3 W_n), \\ (1 - \underline{\delta}_0)u(r) + \underline{\delta}_0\underline{\delta}_1\underline{\delta}_2((1 - \overline{\delta}_3)u(R) + \overline{\delta}_3 W_n) & \text{if } u(r) \leq \underline{\delta}_1\underline{\delta}_2((1 - \overline{\delta}_3)u(R) + \overline{\delta}_3 W_n). \end{cases} \tag{7}$$

Let $\overline{W}$ be the number satisfying $u(r) = \underline{\delta}_1\underline{\delta}_2((1 - \overline{\delta}_3)u(R) + \overline{\delta}_3\overline{W})$, or $\overline{W} = (u(r)/(\underline{\delta}_1\underline{\delta}_2) - (1 - \overline{\delta}_3)u(R))/\overline{\delta}_3$. If $W_n \leq \overline{W}$, we have the upper case of (7), while if $W_n \geq \overline{W}$, we have the lower case of (7).

Step 4: The result.

Note that $U_0(x^n) - U_0(y^n)$ is a linear function of W_n. Let $D_U(W)$ denote the difference $U_0(x^n) - U_0(y^n)$ when $W_n = W$ and $U_0(y^n)$ corresponds to the upper case of (7). Similarly, let $D_L(W)$ denote the difference $U_0(x^n) - U_0(y^n)$ when $W_n = W$ and $U_0(y^n)$ corresponds to the lower case of (7).

In particular, note that $U_0(x^0) - U_0(y^0)$ corresponds to either $D_U(0)$ or $D_L(0)$, depending on whether $U_0(y^0)$ corresponds to either the upper or the lower case of (7).

Case (a): $\overline{W} \geq W_n \geq 0$. Since $W_n \leq \overline{W}$ and $0 \leq \overline{W}$, both $U_0(y^n)$ and $U_0(y^0)$ correspond to the upper case of (7). From Steps 2 and 3,

$$U_0(x^n) - U_0(y^n) = D_U(W_n) = D_U(0) + (\underline{\delta}_0\overline{\delta}_1\overline{\delta}_2\underline{\delta}_3 - \overline{\delta}_0\underline{\delta}_1\underline{\delta}_2\overline{\delta}_3)W_n. \tag{8}$$

Since $x^0 \succ y^0$, we have $D_U(0) = U_0(x^0) - U_0(y^0) > 0$. Moreover, $\underline{\delta}_0\overline{\delta}_1\overline{\delta}_2\underline{\delta}_3 - \overline{\delta}_0\underline{\delta}_1\underline{\delta}_2\overline{\delta}_3 \geq 0$ by assumption. Thus, from (8), $U_0(x^n) - U_0(y^n) > 0$.

Case (b): $W_n \geq 0 \geq \overline{W}$. Since $W_n \geq \overline{W}$ and $0 \geq \overline{W}$, both $U_0(y^n)$ and $U_0(y^0)$ correspond to the lower case of (7). From Steps 2 and 3,

$$U_0(x^n) - U_0(y^n) = D_L(W_n) = D_L(0) + \underline{\delta}_0(\overline{\delta}_1\overline{\delta}_2\underline{\delta}_3 - \underline{\delta}_1\underline{\delta}_2\overline{\delta}_3)W_n. \tag{9}$$

Since $x^0 \succ y^0$, we have $D_L(0) = U_0(x^0) - U_0(y^0) > 0$. Moreover, the assumption implies

$$\frac{\underline{\delta}_3}{\overline{\delta}_3} \geq \frac{\underline{\delta}_1}{\overline{\delta}_1}\frac{\underline{\delta}_2}{\overline{\delta}_2}$$

or equivalently,

$$\overline{\delta}_1\overline{\delta}_2\underline{\delta}_3 \geq \underline{\delta}_1\underline{\delta}_2\overline{\delta}_3.$$

Thus, from (9), $U_0(x^n) - U_0(y^n) > 0$.

Case (c): $W_n \geq \overline{W} \geq 0$. Since $W_n \geq \overline{W}$ and $0 \leq \overline{W}$, $U_0(y^n)$ corresponds to the lower case, while $U_0(y^0)$ corresponds to the upper case of (7). From Steps 2 and 3,

$$\begin{aligned} U_0(x^n) - U_0(y^n) &= D_L(W_n) = D_L(0) + \underline{\delta}_0(\overline{\delta}_1\overline{\delta}_2\underline{\delta}_3 - \underline{\delta}_1\underline{\delta}_2\overline{\delta}_3)W_n \\ &= D_L(0) + \underline{\delta}_0(\overline{\delta}_1\overline{\delta}_2\underline{\delta}_3 - \underline{\delta}_1\underline{\delta}_2\overline{\delta}_3)\overline{W} + \underline{\delta}_0(\overline{\delta}_1\overline{\delta}_2\underline{\delta}_3 - \underline{\delta}_1\underline{\delta}_2\overline{\delta}_3)(W_n - \overline{W}) \\ &= D_U(0) + (\underline{\delta}_0\overline{\delta}_1\overline{\delta}_2\underline{\delta}_3 - \overline{\delta}_0\underline{\delta}_1\underline{\delta}_2\overline{\delta}_3)\overline{W} + \underline{\delta}_0(\overline{\delta}_1\overline{\delta}_2\underline{\delta}_3 - \underline{\delta}_1\underline{\delta}_2\overline{\delta}_3)(W_n - \overline{W}), \end{aligned} \tag{10}$$

where the last equality follows from the fact that $D_L(\overline{W}) = D_U(\overline{W})$. Since $x^0 \succ y^0$, we have $D_U(0) = U_0(x^0) - U_0(y^0) > 0$. Moreover, by assumptions, the second and third terms of (10) are non-negative. Thus, $U_0(x^n) - U_0(y^n) > 0$, as desired. □

4.2 An Example of Procrastination

Consider an example where condition (6) fails to hold. Let $u(c) = c$ and $R = 3$ and $r = 1$. Assume that $[\underline{\delta}_0, \overline{\delta}_0] = [0.1, 0.8]$, $[\underline{\delta}_1, \overline{\delta}_1] = [0.1, 0.1]$, $[\underline{\delta}_2, \overline{\delta}_2] = [0.4, 0.5]$, $[\underline{\delta}_3, \overline{\delta}_3] = [0.1, 0.3]$, and $[\underline{\delta}_4, \overline{\delta}_4] = [0.1, 0.1]$.

Then, $W_1 = (1-\delta_4)r = 0.9$.

$$\begin{aligned} U_0(x^0) = (0,3,1,0,0) &= \underline{\delta}_0((1-\overline{\delta}_1)u(R) + \overline{\delta}_1((1-\overline{\delta}_2)u(r))) \\ &= (0.1)(0.9)3 + (0.1)(0.1)(0.5)1 = 0.275. \end{aligned}$$

Since $\underline{\delta}_1\underline{\delta}_2(1-\overline{\delta}_3)u(R) = (0.1)(0.4)(0.7)3 = 0.084 < 1 = u(r)$,

$$\begin{aligned} U_0(y^0) = (1,0,0,3,0) &= (1-\overline{\delta}_0)u(r) + \overline{\delta}_0\underline{\delta}_1\underline{\delta}_2(1-\overline{\delta}_3)u(R) \\ &= (0.2) + (0.8)(0.1)(0.4)(0.7)3 = 0.2672. \end{aligned}$$

Hence, $x^0 \succ y^0$.

On the other hand,

$$\begin{aligned} U_0(x^1) = (0,3,1,0,1,0) &= \underline{\delta}_0((1-\overline{\delta}_1)u(R) + \overline{\delta}_1((1-\overline{\delta}_2)u(r) + \overline{\delta}_2\underline{\delta}_3 W_n)) \\ &= 0.275 + (0.1)(0.1)(0.5)(0.1)(0.9) = 0.275 + 0.00045 = 0.27545. \end{aligned}$$

Since $\underline{\delta}_1\underline{\delta}_2((1-\overline{\delta}_3)u(R) + \overline{\delta}_3 W_1) = (0.1)(0.4)((0.7)3 + (0.3)(0.9)) = 0.0948 < 1 = u(r)$,

$$\begin{aligned} U_0(y^1) = (1,0,0,3,1,0) &= (1-\overline{\delta}_0)u(r) + \overline{\delta}_0\underline{\delta}_1\underline{\delta}_2((1-\overline{\delta}_3)u(R) + \overline{\delta}_3 W_n) \\ &= 0.2672 + (0.8)(0.1)(0.4)(0.3)(0.9) = 0.00864 = 027584. \end{aligned}$$

Thus, we have a preference reversal, $y^1 \succ x^1$.

5 Proof of Necessity in Theorem 1

We prove the necessity of the axioms. Noor and Takeoka [4] show in Section 2.4 that the optimal discount function takes the form of equation (4) in Noor and Takeoka [4] if $U(x) - u(x_0) \leq mK$ and of equation (5) if $U(x) - u(x_0) > mK$. By Proposition 3 of Noor and Takeoka [4], the set of magnitude-sensitive streams is given by

$$X_m = \{x \in X \mid \sum_{t\geq 1} \gamma(t)u(x_t)^{\frac{m}{m-1}} \leq mK\}. \tag{11}$$

These results were established for the version of the model in the deterministic setting but the arguments are identical for the lottery setting where $u(x_t)$ is the expected utility of a lottery x_t.

Note that $D_x(t)$ is continuous in x. By equations (4) and (5) in Noor and Takeoka [4], $D_x(t)$ is strictly increasing in $u(x_t)$ in X_m, and it is constant on a ray in $X \setminus X_m$. It is obvious that $\succsim$ that U represents satisfies Order, Continuity, C-Monotonicity, Impatience, Present Equivalents, and Risk Preference. Weak Homotheticity requires $\alpha U(x) \geq U(\alpha \circ x)$, which follows from equations (4) and (5) in Noor and Takeoka [4]. Since $D_x(t)$ depends only on $u(x_t)$ on X_{ms}, U is additively separable on this subdomain, which implies that every $x \in X_m$ is in fact separable. MS-Separability is therefore implied, and we also see that $X_m = X_{ms}$. The second part of MS-Homogeneity is implied by the same argument as in Theorem 7 of Noor and Takeoka [2].

To show PBS, notice that the CCE representation is additively separable between period 0 and period t for all $t > 0$. Thus, for all x, $U(0, x_{-0}) = U(x) - u(x_0)$. Take any $x \in X_m = X_{ms}$ and y with $U(0, y_{-0}) \leq U(0, x_{-0})$. Since $U(y) - u(y_0) = U(0, y_{-0}) \leq U(0, x_{-0}) = U(x) - u(x_0) \leq mK$, we have $U(y) - u(y_0) \leq mK$, that is, $y \in X_{ms}$. If $x \notin X_{ms}$, $U(0, x_{-0}) = U(x) - u(x_0) > mK$. Since $U(0, \alpha \circ x_{-0}) \to 0$ as $\alpha \to 0$, there exists some $\alpha \in (0,1)$ such that $U(0, \alpha \circ x_{-0}) \leq U(0, x_{-0})$ and $U(0, \alpha \circ x_{-0}) = U(\alpha \circ x) - u(\alpha \circ x_0) \leq mK$, that is, $\alpha \circ x \in X_{ms}$.

Finally, we show the following property:

Lemma 7. $\succsim$ *satisfies Monotonicity.*

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

Since $u(x_t)^{\frac{m}{m-1}} \geq u(y_t)^{\frac{m}{m-1}}$ for all t, we have $\sum \gamma(t) u(x_t)^{\frac{m}{m-1}} \geq \sum \gamma(t) u(y_t)^{\frac{m}{m-1}}$. If x and y are magnitude sensitive, we have the desired result. Suppose however that either x or y is not magnitude sensitive. Since $\sum \gamma(t) u(x_t)^{\frac{m}{m-1}} \geq \sum \gamma(t) u(y_t)^{\frac{m}{m-1}}$, we have either (a) neither x nor y is magnitude sensitive, or (b) y is magnitude sensitive but x is not. If case (a) holds, the representation implies that

$$\begin{aligned}\sum_{t\geq 1} D_x(t) u(x_t) =& (mK)^{\frac{1}{m}} \left\{ \sum_{t\geq 1} \gamma(t) u(x_t)^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}} \\ \geq& (mK)^{\frac{1}{m}} \left\{ \sum_{t\geq 1} \gamma(t) u(y_t)^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}} = \sum_{t\geq 1} D_y(t) u(y_t),\end{aligned}$$

as desired.

Suppose that case (b) holds. Define $x(\alpha) = \alpha x + (1-\alpha) y$ for all $\alpha \in (0,1)$. Note that $u(x_t) \geq u(x_t(\alpha)) \geq u(y_t) \geq 0$. By continuity of the representation, there exists $\alpha^* \in (0,1)$ such that $x(\alpha)$ is not magnitude sensitive if $\alpha > \alpha^*$ and $x(\alpha)$ is magnitude sensitive if $\alpha \leq \alpha^*$. Since $\sum_{t\geq 1} \gamma(t) u(x_t(\alpha^*))^{\frac{m}{m-1}} = mK$, the representation implies

$\sum_{t\geq 1} D_x(t) u(x_t) = (mK)^{\frac{1}{m}} \left\{ \sum_{t\geq 1} \gamma(t) u(x_t)^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}}$
$\geq (mK)^{\frac{1}{m}} \left\{ \sum_{t\geq 1} \gamma(t) u(x_t(\alpha^*))^{\frac{m}{m-1}} \right\}^{\frac{m-1}{m}} = \sum_{t\geq 1} \gamma(t) u(x_t(\alpha^*))^{\frac{m}{m-1}}$
$\geq \sum_{t\geq 1} \gamma(t) u(y_t)^{\frac{m}{m-1}} = \sum_{t\geq 1} D_y(t) u(y_t)$. This completes the proof. □

References

[1] Barro, R. (1999): "Ramsey Meets Laibson in the Neoclassical Growth Model", *The Quarterly Journal of Economics* 114(4), pp. 1125–1152.

[2] Laibson, D. (1997): “Golden Eggs and Hyperbolic Discounting,” Quarterly Journal of Economics 112, pp. 443–77.

[3] Noor, J. and N. Takeoka (2022): “Optimal Discounting, ” *Econometrica* 90, pp. 585–623.

[4] Noor, J. and N. Takeoka (2023): “Constrained Optimal Discounting, ” Working paper.

[5] Wakai, K. (2008): “A Model of Utility Smoothing”, *Econometrica* 76, pp. 137–153.