KU LEUVEN

DEPARTMENT OF ECONOMICS

When it Sucks to be Sophisticated: On Present Bias, Elasticity of Intertemporal Substitution, Growth and Welfare

Carles Mañó-Cabello and Simon Elgersma

FACULTY OF ECONOMICS AND BUSINESS



DISCUSSION PAPER SERIES DPS 26.05

November 2026

When it Sucks to be Sophisticated: On Present Bias, Elasticity of Intertemporal Substitution, Growth and Welfare

Carles Mañó-Cabello*
Department of Economics
KU Leuven

Simon Elgersma†
Faculty of Economics and Business Economics
University of Groningen

November 10, 2026

Abstract

Does awareness of present bias improve welfare? We answer this in the AK growth model by comparing naive and sophisticated households. For an elasticity of intertemporal substitution (EIS) below one, sophistication raises welfare. At unit EIS the two agents' consumption, growth, and welfare coincide, and the welfare ordering reverses locally. For an EIS above one, a naive agent consumes less initially and accumulates more wealth, but whether this yields higher welfare than a sophisticated agent depends on how the discount function weights short- versus long-run impatience. We provide a testable condition on the discount function that determines the ordering. The choice of discount function is therefore not innocuous for the welfare implications of present bias.

Keywords: Present Bias; Sophistication; Naïveté; Intertemporal Choice.

JEL codes: D91, E21, E71, O40

1 Introduction

Agents often exhibit present-biased behavior in their decision making (Thaler, 1981, Frederick et al., 2002, DellaVigna, 2009, Ericson and Laibson, 2019). Thus, their behavior can be time-inconsistent. A long-standing tradition, going back to Strotz (1955), Phelps and Pollak (1968), Pollak (1968), O'Donoghue and Rabin (1999), has focused on two types of agents: one who is aware of her future impatience and can constrain her future self (sophistication), and one who incorrectly believes she will stick to her intended plans (naïveté).

This paper studies how (the lack of) awareness of present-bias influences savings, economic growth and welfare in the canonical AK-model.[1] Within this framework, we first show that the welfare functionals of naive and sophisticated agents coincide pointwise when the elasticity of intertemporal substitution (EIS, σ) equals unity.[2] We then establish a local reversal of the welfare ordering at $\sigma = 1$. After this, we provide a testable condition on the discount function that allows us to judge whether for any $\sigma \neq 1$ a naive or sophisticated agent achieves higher welfare. We also illustrate this condition with two commonly used present-biased discount functions. Thus, how harmful present-bias is for welfare depends on i) whether agents are naive or sophisticated, ii) the EIS and iii) the shape of the discount

*Corresponding author: Carles Mañó-Cabello. Department of Economics, KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium. Email: carles.manocabello@kuleuven.be.

†Faculty of Economics and Business Economics, University of Groningen, 9747 AE Groningen, The Netherlands.

[1] Any difference in outcomes is attributable to (the lack of) awareness of present-bias as the interest rate is constant.

[2] This happens because consumption paths of naive and sophisticated agents coincide pointwise at $\sigma = 1$ (Pollak, 1968).

function. The last point is important, since a discount function that attaches high weight to the future can result in sophisticated agents achieving higher welfare than naive agents when $\sigma > 1$.

Related literature: There has been an extensive discussion on saving behavior and growth rates under present-biased preferences (Barro, 1999, Strulik, 2015, Cabo et al., 2015). Reversal of consumption and growth rates has been studied in Kang and Kim (2024), Groneck et al. (2024), Acharya et al. (2026), Laibson et al. (2026), Maxted (2025), Maxted et al. (2025). Notably, Groneck et al. (2024) show in a cake-eating problem that a naive (sophisticated) agent has a larger Marginal Propensity to Consume (MPC) than a sophisticated (naive) agent when the EIS < 1 (> 1). Acharya et al. (2026) establish a similar MPC reversal in a continuous-time portfolio choice model studying the specific Instantaneous Gratification discounting framework (Harris and Laibson, 2013).

Fishman (2020) touches upon welfare in the AK growth model, but does not investigate the shape of the discount function. Gómez (2025) recently explored numerically the growth rate orderings between naive and sophisticated agents in the AK model, but restricts himself to the region EIS < 1 and a particular discount function. We extend his numerical analysis in two directions: we derive analytical results for EIS > 1 and allow for a general discount function. Additionally, we show welfare orderings between naive and sophisticated agents are *not monotone* and depend heavily on the shape of the discount function.

2 Model and preliminary results

We briefly present the main model and preliminary results. Following Strulik (2015), Cabo et al. (2015), and Gómez (2025), we consider a continuous-time economy populated with a continuum of homogeneous agents (i.e. households). We either assume the whole continuum of agents is sophisticated (labelled S), naive (labelled N) or exponentially discounting (labelled E). The representative agent maximizes the discounted stream of instantaneous utility,

$$J_t = \int_t^\infty \theta(s-t)\frac{c_t(s)^{1-\frac{1}{\sigma}}-1}{1-\frac{1}{\sigma}}ds, \tag{1}$$

where $c_t(s)$ is the consumption at time s decided at time t and σ is the EIS.[3] Let $j = s - t$ be the time distance from the present t to the future s and define $\theta(s-t)$ as the present-bias discount function.[4]

Definition 1. *A present-biased discount function is a function $\theta(j) > 0$, $\forall j$ which is C^1 on $[0,\infty)$ with $\dot\theta(j) < 0$, $\theta(0) = 1$, $\lim_{j\to\infty}\theta(j) = 0$, and $\Theta \equiv \int_0^\infty \theta(j)\,dj < \infty$. The instantaneous discount rate, $\rho(j) \equiv -\dot\theta(j)/\theta(j)$, is continuous and satisfies $\rho(j) > 0$ and $\dot\rho(j) \leq 0$ for all $j \geq 0$ with $\rho_\infty = \lim_{j\to\infty}\rho(j) > 0$.*

The production technology is of the AK type, $y(t) = Ak(t)$. Capital depreciates at the constant rate $\delta_k \geq 0$. Thus, the net return is constant and equal to $r = A - \delta_k$, and the resource constraint is

$$\dot k_t(s) = rk_t(s) - c_t(s), \quad \text{with} \quad k_t(t) = k_t \quad \text{given.} \tag{2}$$

Cabo et al. (2015) and Gómez (2025) show that the corresponding modified Ramsey rule for each type of the representative agent is given by:[5]

$$\gamma_i = \frac{\dot c_i(t)}{c_i(t)} = \sigma(r - d_i(\sigma,r)), \quad \text{with } d_i(\sigma,r) = \begin{cases}\hat\rho & \text{if } i = E,\\ \lambda_N(\sigma,r) & \text{if } i = N,\\ \lambda_S(\sigma,r) & \text{if } i = S,\end{cases} \tag{3}$$

[3] Unlike Strulik (2015), Cabo et al. (2015), Fishman (2020), Gómez (2025), we decide to write $1 - 1/\sigma$ for the power of consumption, so that the instantaneous utility function features the EIS directly.

[4] Definition 1 satisfies Assumptions 3 and 4 in Fishman (2020).

[5] One recovers the solution for the naive agent with the maximum principle. The sophisticated outcome is obtained as the subgame-perfect equilibrium of an intrapersonal game played between her current and all future selves (Cabo et al., 2015).

where γ_i ($i \in \{E, N, S\}$) is the agent-specific growth-rate of consumption and $d_i = d_i(\sigma, r)$ is the agent-specific effective rate of time preferences (ERTP) and $\hat{\rho}$ the discount rate under exponential discounting.[6] The ERTP can be conveniently expressed as weighted averages of the instantaneous discount rate $\rho(j)$ using weighting functions that depend on (i) the EIS σ and (ii) the agent's behavioral type (see Supplementary Appendix A.1).

As the ERTP differ between naive and sophisticated agents, the literature uses the Equivalent Present Value,

$$\int_0^\infty e^{-\hat{\rho} j} dj \stackrel{!}{=} \int_0^\infty \theta(j) dj, \quad \Longleftrightarrow \quad \hat{\rho} = \left[\int_0^\infty \theta(j) dj\right]^{-1} = \Theta^{-1}. \tag{4}$$

Under (4), we can attribute the welfare outcomes solely to the underlying behavioral assumptions.

As the interest rate is constant, so are the ERTP and the growth rate. Thus, we can express the consumption rule as a constant propensity out of capital, $c_i(t) = b_i k_i(t)$, $i \in \{N, S, E\}$. Using (2), the equilibrium growth rate satisfies $\gamma_i = r - b_i$. Combining this identity with $\gamma_i = \sigma(r - d_i)$ gives the MPC

$$b_i = (1 - \sigma) r + \sigma d_i. \tag{5}$$

Useful for the remainder of this paper is the following result (whenever b_S exists):

Proposition 1. *The ordering of MPCs, ERTP and growth rates is given by*

(i) *If* $\sigma < 1$, $b_S < b_N$, $\lambda_S < \lambda_N$ *and* $\gamma_S > \gamma_N$.

(ii) *If* $\sigma > 1$, $b_S > b_N$, $\lambda_S > \lambda_N$ *and* $\gamma_S < \gamma_N$.

(iii) *If* $\sigma = 1$, $b_S = b_N$, $\lambda_S = \lambda_N$ *and* $\gamma_S = \gamma_N$.

Proof. See Supplementary Appendix A.2. □

3 Local welfare reversal at the EIS=1

In this section we show that the welfare ordering between naive and sophisticated agents reverses at $\sigma = 1$, with welfare defined as the discounted utility stream under the corresponding optimal policies.

Proposition 2 (Welfare ordering of naive and sophisticated agents)**.** *Let θ be any present-biased discount function satisfying Definition 1 and define the equilibrium welfare of agent $i \in \{N, S\}$ as the time-t functional (1) evaluated along that agent's equilibrium consumption path $c_i^\star$ (i.e. generated by Equation (3))*

$$\mathcal{W}_i \;:=\; J_t[c_i^\star] \;:=\; \int_t^\infty \theta(s-t)\, u\big(c_i^\star(s)\big)\, ds \;=\; \int_0^\infty \theta(j)\, u\big(c_i^\star(t+j)\big)\, dj. \tag{6}$$

Define the welfare gap $\Delta\mathcal{W}(\sigma) := \mathcal{W}_S(\sigma) - \mathcal{W}_N(\sigma)$ and $T_1 := \int_0^\infty j\,\theta(j)\,dj$. Then:

1. *Both realized consumption paths and welfare of naive and sophisticated agents coincide at $\sigma = 1$: $\Delta\mathcal{W}(1) = 0$.*
2. *The welfare gap satisfies*

$$\left.\frac{\partial\, \Delta\mathcal{W}}{\partial \sigma}\right|_{\sigma=1} \;=\; \left(\tfrac{1}{\hat{\rho}^2} - T_1\right)\left(\partial_\sigma \lambda_S(1) - \partial_\sigma \lambda_N(1)\right) \;<\; 0.$$

Proof. See Supplementary Appendix A.3. □

Thus, marginally below (above) $\sigma = 1$ a sophisticated (naive) agent achieves higher welfare.

[6] We suppress dependence of the ERTP on r in the remainder of this paper to economize on notation.

4 A condition for global welfare reversal

Proposition 2 states that the welfare ordering between sophisticated and naive agents is guaranteed to reverse at $\sigma = 1$. We now show that this ordering is not necessarily global and depends on the shape of the discount function. It is useful to rewrite Equation (6) as a function of the MPC b_i, and $k(t)$.

$$\mathcal{W}(b_i) = \frac{1}{1-\frac{1}{\sigma}} \left(b_i\, k(t)\right)^{1-\frac{1}{\sigma}} \int_0^\infty \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b_i - r)j}\, dj - \frac{1}{1-\frac{1}{\sigma}} \frac{1}{\hat{\rho}}. \tag{7}$$

We then have:

Theorem 1 (Global welfare ordering)**.**

1. *Let $\sigma < 1$ and given that b_S exists, then $\mathcal{W}_S > \mathcal{W}_N$.*
2. *For a feasible $\sigma > 1$, consider the function $D(j, \sigma, r)$*

$$\mathcal{D}(j, \sigma, r) \equiv j \cdot \left(\rho(j) - \left(1 - \frac{1}{\sigma}\right) r\right) \tag{8}$$

If $D(j, \sigma, r)$ is nondecreasing in $j \in (0, \infty)$, every stationary point of $\mathcal{W}$ lies strictly below b_N, $\mathcal{W}$ is strictly decreasing on $[b_N, \infty)$, and $\mathcal{W}_N > \mathcal{W}_S$.

Proof. For the case $\sigma < 1$ see Fishman (2020). For $\sigma > 1$, see Supplementary Appendix A.5. □

Our results show the shape of the discount function is crucial in the region $\sigma > 1$, which is not touched upon in Fishman (2020).[7] We highlight the implications of Theorem 1 with two types of admissible discount functions: the Tsoukis discount function (c.f. Tsoukis et al. (2017), Cabo et al. (2020a)) and the convex combination of exponentials discount (CCED) function (c.f. Karp (2007), Marín-Solano and Navas (2009)).

Lemma 1.

1. *Consider the Tsoukis discount function, $\theta(j) = (1 + \delta j)^{-\psi/\delta} e^{-\rho j}$. For every $\sigma > 1$ satisfying $\sigma\rho > (\sigma - 1)r$, $\mathcal{D}(j, \sigma, r)$ is strictly increasing in j, thus $\mathcal{W}_N > \mathcal{W}_S$.*
2. *Consider the CCED function, $\theta(j) = ze^{-\rho_L j} + (1 - z)e^{-\rho_S j}$, $z \in (0, 1)$ with z and $(1 - z)$ the weights attached to the long-run rate ρ_L and short-run rate ρ_S where $\rho_L < \rho_S$. Then, there exists a set of mixture parameters (z, ρ_S, ρ_L) and a $\sigma > 1$ s.t D is decreasing and $W_N < W_S$.*[8]

Proof. See Supplementary Appendix A.6. □

For the Tsoukis discount function, Lemma 1 guarantees that the local ordering of Proposition 2 can be extended to a global ordering. With the CCED function, if the short-run and long-run impatience are sufficiently far apart, the discount rate falls sharply over intermediate delays. Although a sophisticated agent will save less than a naive agent (c.f. Proposition 1), whether that results in lower welfare depends on how much weight the discount function attaches to short- and long-run discount rates. The results for the CCED function are illustrated in Figure 1. Additionally, Table 1 shows consumption-equivalent variation and welfare crossings for a high interest rate scenario.[9] See Appendix A.7 for additional visualizations and simulations.

[7] We revisit Fishman's result in Appendix A.4 and show that additional restrictions on the discount function are required for his conclusions to hold.

[8] The discount rate of this mixture function has properties $\rho(0) = z\rho_L + (1 - z)\rho_S$ and $\lim_{j\to\infty} \rho(j) = \rho_L \equiv \rho_\infty$.

[9] The low-interest-rate case gives a wider range of EIS values for where the flip happens. For the case in Figure A.4, the rage of EIS where the sophisticated agent has higher welfare goes from 1.63 to 3.27.

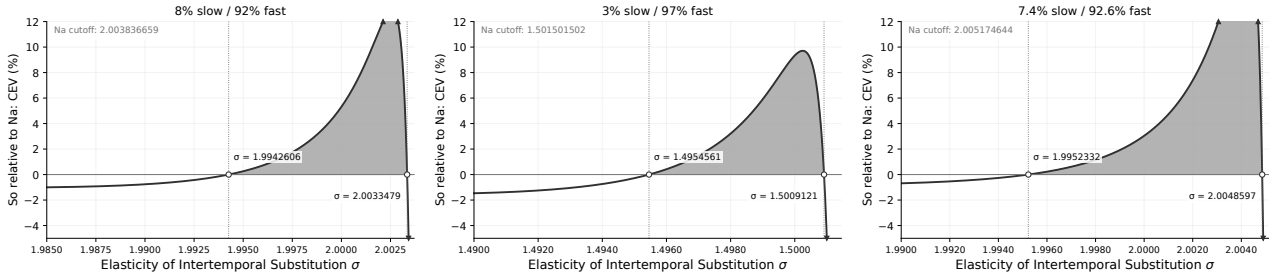


Figure 1: Welfare for different CCED functions.

Notes: "z% Slow" refers to the weight z attached to ρ_L. In the unshaded area naive agents are better off, and in the shaded area the sophisticated agents are better off.

Table 1: Consumption-Equivalent Variations and Equal-Welfare Crossings

					Equal-Welfare Crossings	
z	ρ_L	ρ_S	Selected σ	$\text{CEV}_{\text{S}\|\text{N}}$ (%)	σ_1	σ_2
0.080	0.0250478664	0.0421900000	2.0	+5.313018	1.994260596	2.003347854
0.030	0.0167000000	0.0418038710	1.5	+9.342818	1.495456054	1.500912083
0.074	0.0250645161	0.0420000000	2.0	+3.038192	1.995233165	2.004859700

Notes: This table reports the consumption-equivalent variation ($\text{CEV}_{\text{S}|\text{N}}$) and the welfare crossing thresholds for three parameterizations of the CCED function $\theta(j) = ze^{-\rho_L j} + (1-z)e^{-\rho_S j}$ at $r = 5\%$. $\text{CEV}_{\text{S}|\text{N}}(\sigma)$ is the percentage initial wealth adjustment required for a naive household to achieve the lifetime welfare of a sophisticated household, defined by $\mathcal{W}_{\text{N}}([1 + \text{CEV}_{\text{S}|\text{N}}]k_0) = \mathcal{W}_{\text{S}}(k_0)$. The columns σ_1 and σ_2 report the crossings where $\mathcal{W}_{\text{N}}(k_0) = \mathcal{W}_{\text{S}}(k_0)$; for any EIS $\sigma \in (\sigma_1, \sigma_2)$, the naive household experiences a strictly different welfare ranking than outside this interval.

5 Conclusion

This paper has shown that welfare ordering of naive and sophisticated households is not straightforward. A naive agent has a higher (lower) MPC than a sophisticated agent when $\sigma < 1$ (> 1), and consumption growth ordering among the two agents follows this reversal. However, the welfare ordering does not. Locally around $\sigma = 1$, the welfare ordering between naive and sophisticated agents reverses. Marginally above (below) $\sigma = 1$, a naive (sophisticated) agent achieves higher welfare. However, we show that sophisticated agents can achieve higher welfare even when $\sigma > 1$. Thus, the global welfare ordering depends on the shape of the discount function, making its choice not innocent.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process.

During the preparation of this work the authors used Claude and ChatGPT to aid with calculations, code, and writing. After using these services, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

This work was supported by the Research Foundation Flanders (FWO) under project DFG-E5810-PDO/24 MANO-CABELLO (1250025N; 3H240148), by the Spanish State Research Agency under projects PID2020-112509GB-100 and TED2021-130390B-I00. The funders had no role in the study design, analysis, interpretation, writing of the manuscript, or decision to submit the article for publication.

Acknowledgements

We would like to thank Pietro Peretto, Florian Wagener, Pim Heijnen, David Jimenez-Gomez, Alejandro Rivera, Marco Haan and Max Groneck for their comments and suggestions.

References

Acharya, S., D. Jimenez-Gomez, D. Rachinskii, and A. Rivera (2026): "Present-Bias and the Value of Sophistication: Splurging vs. Smoothing," *Games and Economic Behavior*. 2

Barro, R. J. (1999): "Ramsey meets Laibson in the neoclassical growth model," *The Quarterly Journal of Economics*, 114, 1125–1152. 2

Cabo, F., G. Martín-Herrán, and M. P. Martínez-García (2015): "Non-constant discounting and AK-type growth models," *Economics Letters*, 131, 54–58. 2, 9, 11

——— (2020a): "Non-constant discounting, social welfare and endogenous growth with pollution externalities," *Environmental and Resource Economics*, 76, 369–403. 4, 18, 20

——— (2020b): "Present bias and the inefficiency of the centralized economy: The role of the elasticity of intertemporal substitution," *Economic Modelling*, 93, 702–716. 9

DellaVigna, S. (2009): "Psychology and economics: Evidence from the field," *Journal of Economic Literature*, 47, 315–72. 1

Ericson, K. M. and D. Laibson (2019): "Intertemporal choice," in *Handbook of behavioral economics: Applications and foundations 1*, Elsevier, vol. 2, 1–67. 1

Fishman, R. (2020): "Welfare implications of naive and sophisticated saving," *Social Choice and Welfare*, 54, 623–638. 2, 4, 16, 17

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time discounting and time preference: A critical review," *Journal of Economic Literature*, 40, 351–401. 1

Gómez, M. A. (2025): "Sophistication drives higher growth than naïveté in the AK model with hyperbolic discounting: An exhaustive numerical exploration," *Economics Letters*, 112311. 2, 22, 23

Groneck, M., A. Ludwig, and A. Zimper (2024): "Who saves more, the naive or the sophisticated agent?" *Journal of Economic Theory*, 219, 105848. 2

Harris, C. and D. Laibson (2013): "Instantaneous gratification," *The Quarterly Journal of Economics*, 128, 205–248. 2

Kang, M. and E. Kim (2024): "Present bias and endogenous growth," *Economics Letters*, 236, 111556. 2

Karp, L. (2007): "Non-constant discounting in continuous time," *Journal of Economic Theory*, 132, 557–568. 4

Laibson, D., S. C. Lee, P. Maxted, A. Repetto, and J. Tobacman (2026): "Estimating discount functions with consumption choices over the lifecycle," *The Review of Financial Studies*, hhae035. 2

Mañó-Cabello, C. (2025): "Being Human: Endogenous Growth, Pollution and Natural Resources under Time-Inconsistent Preferences," *SSRN Electronic Journal*, working Paper version, updated November 5, 2025. 8

Marín-Solano, J. and J. Navas (2009): "Non-constant discounting in finite horizon: The free terminal time case," *Journal of Economic Dynamics and Control*, 33, 666–675. 4

——— (2010): "Consumption and portfolio rules for time-inconsistent investors," *European Journal of Operational Research*, 201, 860–872. 9

Maxted, P. (2025): "Present bias unconstrained: Consumption, welfare, and the present-bias dilemma," *The Quarterly Journal of Economics*, 140, 2963–3013. 2

Maxted, P., D. Laibson, and B. Moll (2025): "Present bias amplifies the household balance-sheet channels of macroeconomic policy," *The Quarterly Journal of Economics*, 140, 691–743. 2

O'Donoghue, T. and M. Rabin (1999): "Doing it now or later," *American Economic Review*, 89, 103–124. 1

Phelps, E. S. and R. A. Pollak (1968): "On second-best national saving and game-equilibrium growth," *The Review of Economic Studies*, 35, 185–199. 1

Pollak, R. A. (1968): "Consistent planning," *The Review of Economic Studies*, 35, 201–208. 1

Strotz, R. H. (1955): "Myopia and inconsistency in dynamic utility maximization," *The Review of Economic Studies*, 23, 165–180. 1

Strulik, H. (2015): "Hyperbolic discounting and endogenous growth," *Economics Letters*, 126, 131–134. 2

Thaler, R. (1981): "Some empirical evidence on dynamic inconsistency," *Economics letters*, 8, 201–207. 1

Tsoukis, C., F. Tournemaine, and M. Gillman (2017): "Hybrid Exponential-Hyperbolic Discounting and Growth Without Commitment," *The Manchester School*, 85, e45–e74. 4, 18, 20

A Supplementary Appendix

A.1 Admissibility and weighting functionals

Assumption 1 (Admissibility). *Define*

$$Z(\ell,\alpha,\sigma) := \int_0^\infty K(j;\ell,\alpha,\sigma)\,dj, \qquad N(\ell,\alpha,\sigma) := \int_0^\infty \rho(j)\,K(j;\ell,\alpha,\sigma)\,dj, \tag{A.1}$$

so one can write $\Psi(\ell,\alpha,\sigma) = N(\ell,\alpha,\sigma)/Z(\ell,\alpha,\sigma)$. *The triple* (ℓ,α,σ) *is admissible if* $0 < Z(\ell,\alpha,\sigma) < \infty$ *and* $0 < N(\ell,\alpha,\sigma) < \infty$. *Assume that for the parameter points of interest,*

$$\int_0^\infty j\,K(j;\ell,\alpha,\sigma)\,dj < \infty, \qquad \int_0^\infty |\ln\theta(j)|\,K(j;\ell,\alpha,\sigma)\,dj < \infty, \tag{A.2}$$

These conditions ensure we may differentiate $Z(.)$ *and* $N(.)$ *by moving derivatives inside the integrals.*

Given (ℓ,α,σ), define the kernel

$$K(j;\ell,\alpha,\sigma) := \theta(j)^\alpha \exp\big(-(1-\sigma)(r-\ell)j\big), \qquad j \in [0,\infty). \tag{A.3}$$

Then the weighted-average functional is

$$\Psi(\ell,\alpha,\sigma) = \frac{N(\ell,\alpha,\sigma)}{Z(\ell,\alpha,\sigma)}. \tag{A.4}$$

Define the induced probability density (weight)

$$\omega(j;\ell,\alpha,\sigma) := \frac{K(j;\ell,\alpha,\sigma)}{\int_0^\infty K(j;\ell,\alpha,\sigma)\,dj}, \qquad \text{where } \int_0^\infty \omega(j;\ell,\alpha,\sigma)\,dj = 1. \tag{A.5}$$

With this notation, we can define the Effective Rate of Time Preferences (ERTP) as the weighted average of the *instantaneous discount rate* $\rho(j)$,

$$\Psi(\ell,\alpha,\sigma) = \int_0^\infty \rho(j)\,\omega(j;\ell,\alpha,\sigma)\,dj =: E_{\omega(j;\ell,\alpha,\sigma)}[\rho(s)]. \tag{A.6}$$

The object Ψ can be interpreted as an expectation of ρ over the artificial density $\omega(.)$. This interpretation is particularly useful, as we can use the properties of expectations and covariances to identify the relative slopes of the growth rates of the naive and sophisticated agents.

Lemma 2 (Sufficient tail condition). *Assume there exist* $C > 0$, $\underline{\rho} > 0$, *and* $J \geq 0$ *such that* $\theta(j) \leq Ce^{-\underline{\rho}j}$ *for all* $j \geq J$. *Then* (ℓ,α,σ) *is admissible whenever*

$$(\sigma-1)(r-\ell) < \alpha\,\underline{\rho}.$$

In particular, for $\sigma > 1$ *the naive effective rate* $\lambda_N = \Psi(0,\sigma,\sigma)$ *is admissible if* $(\sigma-1)r < \sigma\,\underline{\rho}$, *or in terms of a maximum EIS possible, then* $\sigma < \bar{\sigma} \equiv r/(r-\underline{\rho})$, *where* $\bar{\sigma}$ *is the threshold above which, the integrals are not well defined (see also Mañó-Cabello (2025)).*

Proof. Fix (ℓ,α,σ) and suppose $(\sigma-1)(r-\ell) < \alpha\underline{\rho}$. For $j \geq J$ we have

$$K(j;\ell,\alpha,\sigma) \leq (Ce^{-\underline{\rho}j})^\alpha\, e^{(\sigma-1)(r-\ell)j} = C^\alpha e^{-(\alpha\underline{\rho}-(\sigma-1)(r-\ell))j}.$$

The exponent is strictly negative by the stated inequality in Lemma 2, i.e., $(\sigma-1)(r-\ell) < \alpha\,\underline{\rho}$, hence $\int_J^\infty K < \infty$. On $[0,J]$ the function K is bounded, so $\int_0^J K < \infty$. Thus $\int_0^\infty K < \infty$ and is strictly positive.

For the numerator, note $\rho(j)$ is continuous on $[0,J]$ hence bounded there, and $\rho(j) \geq 0$. On $[J,\infty)$, $\rho(j) \leq \rho(0)$ since ρ is nonincreasing. Therefore $\int_J^\infty \rho K \leq \rho(0)\int_J^\infty K < \infty$, and $\int_0^J \rho K < \infty$ by boundedness.

□

With the machinery defined above, we can express the ERTP in the following way.

Definition 2. *Let $\alpha \in \{\sigma, 1\}$ be a type index and $\ell \in \{0, \lambda_S\}$. We set $\alpha = \sigma$ and $\ell = 0$ for the naive agent, and $\alpha = 1$ and $\ell = \lambda_S$ for the sophisticated agent. As the kernel $K(j;\ell,\alpha,\sigma) \equiv \theta(j)^{\alpha} e^{-(1-\sigma)(r-\ell)j}$, then, one can define a weighting functional as*

$$\Psi(\ell,\alpha,\sigma) \equiv \int_0^{\infty} \rho(j) \underbrace{\frac{K(j;\ell,\alpha,\sigma)}{\int_0^{\infty} K(j;\ell,\alpha,\sigma)\, dj}}_{:=\omega(j;\ell,\alpha,\sigma)} dj, \tag{A.7}$$

where $\omega(j;\ell,\alpha,\sigma)$ can be interpreted as the induced density/weight. With this, one can define the corresponding ERTP for both types of agents, naive and sophisticated. The naive ERTP λ_N is given by:

$$\lambda_N = \Psi(0,\sigma,\sigma). \tag{A.8}$$

The sophisticated effective rate λ_S is given as the fixed point solution λ_S of:

$$\lambda_S = \Psi(\lambda_S, 1, \sigma). \tag{A.9}$$

In Definition 2, we link the effective rates for both naive and sophisticated agents to a weighting functional $\Psi(\cdot)$. This weighting functional features induced weights $\omega(\cdot)$, which have the standard properties of density weights. For a compact and coherent exposition of our results, we will use expectation and covariance operators in the remainder of this paper.[10] Note also that at $\sigma = 1$ both induced weights collapse to

$$\omega_0(j) := \frac{\theta(j)}{\int_0^{\infty} \theta(s)\, ds},$$

where we simplify notation $\omega_0(j) := \omega_0(j,\ell,\alpha,\sigma = 1)$ for the case of logarithmic utility, where one obtains $\lambda_N(\sigma = 1, r) = \lambda_S(\sigma = 1, r) = \hat{\rho}$: the equivalence result of λ_S and λ_N at $\sigma = 1$ (see Marín-Solano and Navas (2010), Cabo et al. (2015, 2020b)). At $\sigma = 1$, both functions $K(.)$ collapse to $\theta(j)$, as the exponential tilt disappears and $\alpha = \sigma = 1$ for the naive agent.

A.2 Proof of Proposition 1

In order to prove Proposition 1, we need a few results. The objects in Definition 2 have simple expressions in terms of the MPC $b_i = (1-\sigma)r + \sigma\lambda_i$ of Equation (5). Integrating $N(\ell,\alpha,\sigma)$ by parts gives $\alpha\, N(\ell,\alpha,\sigma) = 1 + (\sigma - 1)(r - \ell)\, Z(\ell,\alpha,\sigma)$. Substituting this into (A.8) and (A.9), gives the two effective rates are equivalent to

$$b_N = \left[\int_0^{\infty} \theta(j)^{\sigma} e^{(\sigma-1)rj}\, dj\right]^{-1}, \tag{A.10}$$

$$1 = b_S \int_0^{\infty} \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b_S - r)j}\, dj, \tag{A.11}$$

where b_N is defined explicitly, and the MPC of the sophisticated agents is defined implicitly as the solution of the fixed point b_S. It will be useful to define, whenever the integral converges, the RHS of the sophisticated expression,

$$F(b) := b \int_0^{\infty} \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\, dj,$$

so that a sophisticated equilibrium is a solution of $F(b_S) = 1$. Additionally, we need two Integration By Parts identities derived in Lemma 3.

[10] It is to be understood that these expectation and covariance operators are induced under their corresponding density weights $\omega(\cdot)$, as there is no stochasticity in our framework.

Lemma 3. *(i) For every x with $\rho_\infty + x > 0$,*

$$\int_0^\infty j\,(\rho(j)+x)\,\theta(j)\,e^{-xj}\,dj = \int_0^\infty \theta(j)\,e^{-xj}\,dj. \tag{A.12}$$

(ii) For every $\sigma > 0$ and real g with $\sigma\rho_\infty - g > 0$,

$$\int_0^\infty j\,(\sigma\rho(j)-g)\,\theta(j)^\sigma e^{gj}\,dj = \int_0^\infty \theta(j)^\sigma e^{gj}\,dj. \tag{A.13}$$

Proof. (i) Since $\dfrac{d}{dj}\Big[\theta(j)e^{-xj}\Big] = -\big(\rho(j)+x\big)\theta(j)e^{-xj}$, integration by parts gives

$$\int_0^\infty j\,(\rho(j)+x)\theta(j)e^{-xj}\,dj = -\Big[j\,\theta(j)e^{-xj}\Big]_{j=0}^{j\to\infty} + \int_0^\infty \theta(j)e^{-xj}\,dj.$$

Because $\rho(j) \geq \rho_\infty$, $\theta(j)e^{-xj} \leq e^{-(\rho_\infty+x)j}$, so the boundary term vanishes. (ii) Identical, using $\dfrac{d}{dj}\Big[\theta(j)^\sigma e^{gj}\Big] = -\big(\sigma\rho(j)-g\big)\theta(j)^\sigma e^{gj}$ and $\theta(j)^\sigma e^{gj} \leq e^{-(\sigma\rho_\infty - g)j}$. □

Now we can provide a proof for Proposition 1.

Proof of Proposition 1, existence and uniqueness. Intuitively, we ask whether $F(b)$ reaches 1 as b increases (existence), and whether this happens only once (uniqueness).

Differentiating under the integral sign (the derivative of the integrand is dominated by a multiple of $j\,e^{-\varepsilon j}$ for some $\varepsilon > 0$),

$$\frac{dF(b)}{db} = \int_0^\infty \theta(j)\,e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\,dj - \left(1-\frac{1}{\sigma}\right)\,b\int_0^\infty j\,\theta(j)\,e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\,dj. \tag{A.14}$$

By (A.12) with $x = \left(1-\frac{1}{\sigma}\right)(b-r)$,

$$\int_0^\infty \theta(j)\,e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\,dj = \int_0^\infty j\left(\rho(j) + \left(1-\frac{1}{\sigma}\right)(b-r)\right)\theta(j)\,e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\,dj.$$

Substituting the second display into the first, the two terms in b cancel and we obtain

$$\frac{dF(b)}{db} = \int_0^\infty j\left(\rho(j) - \left(1-\frac{1}{\sigma}\right)r\right)\theta(j)\,e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\,dj. \tag{A.15}$$

(i) Here $\left(1-\frac{1}{\sigma}\right) < 0$ and $r \geq 0$, so $\rho(j) - \left(1-\frac{1}{\sigma}\right)r > 0$ and F is strictly increasing wherever it is finite. Since $\theta(j) \leq e^{-\rho_\infty j}$, the integral converges for every $b \in \big(0, \bar{b}\big)$, where $\bar{b} := r + \frac{\sigma\rho_\infty}{1-\sigma}$, and $F(b) \to 0$ as $b \to 0^+$. This proves convergence below the cutoff. As $b \uparrow \bar{b}$ the integrand increases pointwise to $\theta(j)\,e^{\rho_\infty j}$, because $-\left(1-\frac{1}{\sigma}\right)\left(\bar{b}-r\right) = \rho_\infty$; by monotone convergence $F(b) \to \bar{b}\int_0^\infty \theta(j)\,e^{\rho_\infty j}\,dj \in (0, \infty]$. For $b > \bar{b}$ the integral diverges, because $\theta(j)\,e^{\rho_\infty j} = \exp\big(-\int_0^j(\rho(i)-\rho_\infty)\,di\big)$ decays more slowly than any exponential. The claims follow from continuity and strict monotonicity. (In the knife-edge case $\bar{b}\int_0^\infty \theta(j)\,e^{\rho_\infty j}\,dj = 1$ the equation is solved by $b = \bar{b}$ itself.)

(ii) Here $\rho(j) \geq \rho_\infty > \left(1-\frac{1}{\sigma}\right)r$ for all j, so the integrand in (A.15) is strictly positive and F is strictly increasing on $(0,\infty)$. As $b \to 0^+$, $F(b) \to 0$ because $\int_0^\infty \theta(j)e^{\left(1-\frac{1}{\sigma}\right)rj}dj < \infty$. As $b \to \infty$, $\theta(j) \geq e^{-\rho(0)j}$ gives $F(b) \geq \dfrac{b}{\rho(0) + \left(1-\frac{1}{\sigma}\right)(b-r)} \to \frac{1}{1-\frac{1}{\sigma}} > 1$. By continuity and monotonicity there is a unique b_S with $F(b_S) = 1$.

□

Proof of Proposition 1, ordering. Let $\sigma > 1$. For every $j \geq 0$,

$$\theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b_S - r)j} = \left(\theta(j)^{\sigma} e^{(\sigma-1)rj}\right)^{\frac{1}{\sigma}} \left(e^{-b_S j}\right)^{1-\frac{1}{\sigma}}.$$

Hölder's inequality with exponents σ and $\frac{1}{1-\frac{1}{\sigma}}$, applied to (A.11), gives

$$\begin{aligned} 1 &= b_S \int_0^{\infty} \left(\theta(j)^{\sigma} e^{(\sigma-1)rj}\right)^{\frac{1}{\sigma}} \left(e^{-b_S j}\right)^{1-\frac{1}{\sigma}} dj \\ &< b_S \left(\int_0^{\infty} \theta(j)^{\sigma} e^{(\sigma-1)rj}\, dj\right)^{\frac{1}{\sigma}} \left(\int_0^{\infty} e^{-b_S j}\, dj\right)^{1-\frac{1}{\sigma}} \\ &= b_S\, b_N^{-\frac{1}{\sigma}}\, b_S^{-\left(1-\frac{1}{\sigma}\right)} = \left(\frac{b_S}{b_N}\right)^{\frac{1}{\sigma}}, \end{aligned}$$

with strict inequality because the two functions are not proportional when $\rho(j)$ is not constant. Hence $b_S > b_N$, meaning that the marginal propensity to consume from the naive is smaller and she will save more.

Let $\sigma < 1$. For every $j \geq 0$, $\theta(j)^{\sigma} e^{(\sigma-1)rj} = \left(\theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b_S - r)j}\right)^{\sigma} \left(e^{-b_S j}\right)^{1-\sigma}$, and Hölder's inequality with exponents $\frac{1}{\sigma}$ and $\frac{1}{1-\sigma}$, applied to (A.10), gives

$$\frac{1}{b_N} < \left(\int_0^{\infty} \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b_S - r)j}\, dj\right)^{\sigma} \left(\int_0^{\infty} e^{-b_S j}\, dj\right)^{1-\sigma} = \left(\frac{1}{b_S}\right)^{\sigma} \left(\frac{1}{b_S}\right)^{1-\sigma} = \frac{1}{b_S}.$$

Hence $b_S < b_N$, which means that the marginal propensity to consume for the sophisticated is smaller, and sophisticated agents save more.

For $\sigma = 1$, see Cabo et al. (2015). □

A.3 Proof of Proposition 2

We first need a Lemma that orders the slopes of the ERTP at $\sigma = 1$. Figures A.5 and A.6 show the intuition.

Lemma 4. $\lambda_S'(1) - \lambda_N'(1) = -\operatorname{Cov}_{\omega_0}(\rho(j), \ln\theta(j) + \hat{\rho} j) = \frac{1}{2}\int_0^{\infty} g(j)^2 \rho(j) \omega_0(j)\, dj > 0$ *where* $\omega_0(j) = \frac{\theta(j)}{\Theta}$.

Proof. Here we prove Lemma 4 by demonstrating that the covariance $\operatorname{Cov}_{\omega_0}\left(\rho(j), g(j)\right) < 0$, where $g(j) \equiv \ln\theta(j) + \hat{\rho} j$. Recall that the benchmark density at $\sigma = 1$ is $\omega_0(j) := \theta(j) / \int_0^{\infty} \theta(s)\, ds = \hat{\rho}\, \theta(j)$, where the second equality uses $\hat{\rho} = [\int_0^{\infty} \theta(s)\, ds]^{-1}$ from (4).

Keep r fixed. Since $\rho_{\infty} > 0$, $\theta(j) \leq e^{-\rho_{\infty} j}$ and $|\ln\theta(j)| \leq \rho(0) j$. These bounds justify differentiation under the integral sign in neighborhoods of the relevant parameter points at $\sigma = 1$. Using $\Psi = N/Z$, the quotient rule gives, for $p \in \{\ell, \alpha, \sigma\}$,

$$\begin{aligned} \partial_p \Psi &= \frac{1}{Z}\int_0^{\infty} \rho(j) K(j)\, \partial_p \ln K(j)\, dj - \Psi \frac{1}{Z}\int_0^{\infty} K(j)\, \partial_p \ln K(j)\, dj \\ &= \operatorname{Cov}_{\omega}\left(\rho(j), \partial_p \ln K(j)\right). \end{aligned}$$

Since $\ln K = \alpha \ln\theta - (1-\sigma)(r-\ell) j$,

$$\partial_{\ell} \ln K = (1-\sigma) j, \qquad \partial_{\alpha} \ln K = \ln\theta(j), \qquad \partial_{\sigma} \ln K = (r-\ell) j.$$

In particular, $\partial_{\ell}\Psi(\hat{\rho}, 1, 1) = 0$. Hence the derivative with respect to ℓ of $\ell - \Psi(\ell, 1, \sigma)$ equals 1 at $(\ell, \sigma) = (\hat{\rho}, 1)$, so the implicit function theorem gives a differentiable sophisticated solution near $\sigma = 1$.

Applying the chain rule to $\lambda_N(\sigma) = \Psi(0,\sigma,\sigma)$ and $\lambda_S(\sigma) = \Psi(\lambda_S(\sigma),1,\sigma)$ gives

$$\begin{aligned}\lambda_N'(\sigma) &= \partial_\alpha \Psi(0,\sigma,\sigma) + \partial_\sigma \Psi(0,\sigma,\sigma),\\ \lambda_S'(\sigma) &= \partial_\ell \Psi(\lambda_S(\sigma),1,\sigma)\,\lambda_S'(\sigma) + \partial_\sigma \Psi(\lambda_S(\sigma),1,\sigma).\end{aligned}$$

At $\sigma = 1$, both weights equal ω_0 and both effective rates equal $\hat{\rho}$. Therefore,

$$\begin{aligned}\lambda_N'(1) &= \mathrm{Cov}_{\omega_0}\big(\rho(j), \ln\theta(j)\big) + r\ \mathrm{Cov}_{\omega_0}\big(\rho(j), j\big),\\ \lambda_S'(1) &= (r-\hat{\rho})\,\mathrm{Cov}_{\omega_0}\big(\rho(j), j\big).\end{aligned}$$

Subtracting yields

$$\lambda_S'(1) - \lambda_N'(1) = -\,\mathrm{Cov}_{\omega_0}\big(\rho(j), \ln\theta(j) + \hat{\rho} j\big).$$

It remains to show that this covariance is negative.

The proof proceeds in four steps.

1): We first start expanding the covariance definition. We use the fundamental definition of covariance for two functions, $\rho(j)$ and $g(j)$, distributed according to the probability density function $\omega_0(j)$. The standard definition is the expected value of the product of their deviations from their respective means:

$$\mathrm{Cov}_{\omega_0}\left(\rho(j), g(j)\right) = \mathbb{E}_{\omega_0}\Big[\big(\rho(j) - \mathbb{E}_{\omega_0}[\rho]\big)\big(g(j) - \mathbb{E}_{\omega_0}[g]\big)\Big]. \tag{A.16}$$

Because the expectation operator $\mathbb{E}_{\omega_0}[\cdot]$ is linear, we can expand this product:

$$\mathrm{Cov}_{\omega_0}\left(\rho(j), g(j)\right) = \mathbb{E}_{\omega_0}\Big[\big(\rho(j) - \mathbb{E}_{\omega_0}[\rho]\big)g(j)\Big] - \mathbb{E}_{\omega_0}\Big[\big(\rho(j) - \mathbb{E}_{\omega_0}[\rho]\big)\mathbb{E}_{\omega_0}[g]\Big]. \tag{A.17}$$

Notice that the second term contains the constant $\mathbb{E}_{\omega_0}[g]$, which can be pulled out of the expectation. What remains inside is the expected deviation of $\rho(j)$ from its own mean, which equals zero:

$$\mathbb{E}_{\omega_0}\big[\rho(j) - \mathbb{E}_{\omega_0}[\rho]\big] = \mathbb{E}_{\omega_0}[\rho] - \mathbb{E}_{\omega_0}[\rho] = 0. \tag{A.18}$$

Therefore, the entire second term vanishes, leaving us with an identity for covariance:

$$\mathrm{Cov}_{\omega_0}\left(\rho(j), g(j)\right) = \mathbb{E}_{\omega_0}\Big[\big(\rho(j) - \mathbb{E}_{\omega_0}[\rho]\big)g(j)\Big]. \tag{A.19}$$

Writing this expectation explicitly as an integral over the probability density $\omega_0(j)$, we arrive at our starting formula:

$$\mathrm{Cov}_{\omega_0}\left(\rho(j), g(j)\right) = \int_0^\infty \Big(\rho(j) - \mathbb{E}_{\omega_0}[\rho]\Big)g(j)\omega_0(j)\,dj. \tag{A.20}$$

To evaluate this, we first compute the expected value of the instantaneous discount rate, $\mathbb{E}_{\omega_0}[\rho]$, under the benchmark density $\omega_0(j) = \hat{\rho}\theta(j)$. Recall that the instantaneous discount rate is defined as $\rho(j) \equiv -\dot{\theta}(j)/\theta(j)$. Integrating this gives:

$$\mathbb{E}_{\omega_0}[\rho] = \int_0^\infty \rho(j)\omega_0(j)\,dj = \int_0^\infty \left(-\frac{\dot{\theta}(j)}{\theta(j)}\right)\big(\hat{\rho}\theta(j)\big)\,dj = -\hat{\rho}\int_0^\infty \dot{\theta}(j)\,dj. \tag{A.21}$$

By the Fundamental Theorem of Calculus, $\int_0^\infty \dot{\theta}(j)\,dj = \lim_{j\to\infty}\theta(j) - \theta(0) = 0 - 1 = -1$. Therefore $\mathbb{E}_{\omega_0}[\rho] = -\hat{\rho}(-1) = \hat{\rho}$. Substituting $\mathbb{E}_{\omega_0}[\rho] = \hat{\rho}$ back into equation (A.20), our covariance simplifies to:

$$\mathrm{Cov}_{\omega_0}\left(\rho(j), g(j)\right) = \int_0^\infty \big(\rho(j) - \hat{\rho}\big)g(j)\omega_0(j)\,dj. \tag{A.22}$$

2) Second, consider the derivative of our object $g(j) = \ln\theta(j) + \hat{\rho} j$ with respect to j:

$$g'(j) = \frac{\dot{\theta}(j)}{\theta(j)} + \hat{\rho} = -\rho(j) + \hat{\rho} = -\big(\rho(j) - \hat{\rho}\big). \tag{A.23}$$

Notice that the term $\big(\rho(j) - \hat{\rho}\big)$ in our integral is exactly the negative of $g'(j)$. We substitute this directly into equation (A.22):

$$\mathrm{Cov}_{\omega_0}\big(\rho(j), g(j)\big) = \int_0^\infty \big(-g'(j)\big) g(j)\omega_0(j)\, dj = -\int_0^\infty g'(j) g(j)\omega_0(j)\, dj. \tag{A.24}$$

3) We now use integration by parts. By the chain rule, we know that $g'(j)g(j) = \frac{1}{2}\frac{d}{dj}[g(j)^2]$. Substituting this into (A.24) gives:

$$\mathrm{Cov}_{\omega_0}\big(\rho(j), g(j)\big) = -\frac{1}{2}\int_0^\infty \frac{d}{dj}[g(j)^2]\omega_0(j)\, dj. \tag{A.25}$$

Now we apply integration by parts $\int u\, dv = [uv] - \int v\, du$, where $u = \omega_0(j)$ and $dv = d[g(j)^2]$. The derivative of the density is $\dot{\omega}_0(j) = \hat{\rho}\dot{\theta}(j)$. Since $\dot{\theta}(j) = -\rho(j)\theta(j)$, we have $du = -\rho(j)\big(\hat{\rho}\theta(j)\big)\, dj = -\rho(j)\omega_0(j)\, dj$. Applying the integration by parts formula yields:

$$\mathrm{Cov}_{\omega_0}\big(\rho(j), g(j)\big) = -\frac{1}{2}\big[g(j)^2\omega_0(j)\big]_0^\infty + \frac{1}{2}\int_0^\infty g(j)^2\big(-\rho(j)\omega_0(j)\big)\, dj. \tag{A.26}$$

4) The fourth step is to evaluate the boundary term $\big[g(j)^2\omega_0(j)\big]_0^\infty$. Note that at the lower bound $j = 0$, one can see that $g(0) = \ln(\theta(0)) + \hat{\rho}\cdot 0 = \ln(1) + 0 = 0$. Thus, the boundary is exactly zero at $j = 0$.

The upper bound $j \to \infty$ needs more work. Consider

$$\lim_{j\to\infty} g(j)^2\, \omega_0(j) = \lim_{j\to\infty} \big(\ln\theta(j) + \hat{\rho} j\big)^2\, \hat{\rho}\, \theta(j).$$

Expanding the square,

$$\big(\ln\theta(j) + \hat{\rho} j\big)^2\, \hat{\rho}\, \theta(j) = \hat{\rho}\, \theta(j)\, (\ln\theta(j))^2 \;+\; 2\hat{\rho}^2\, j\theta(j)\, \ln\theta(j) \;+\; \hat{\rho}^3\, j^2\, \theta(j).$$

We show that each term converges to 0. First, we study the log-squared term. Since $\lim_{j\to\infty}\theta(j) = 0^+$, set $x = \theta(j)$ and use $\lim_{x\to 0^+} x(\ln x)^2 = 0$ to obtain

$$\lim_{j\to\infty} \hat{\rho}\, \theta(j)\, (\ln\theta(j))^2 = 0.$$

Secondly, we study the polynomial term. Assumption 1 at $\sigma = 1$ requires $\int_0^\infty j\,\theta(j)\, dj < \infty$, and note that $j\theta(j) \geq 0$. Hence the tail integral must vanish:

$$\lim_{J\to\infty}\int_J^\infty j\,\theta(j)\, dj = 0, \qquad \text{so in particular} \qquad \lim_{J\to\infty}\int_J^{2J} j\,\theta(j)\, dj = 0.$$

Since θ is decreasing, for $j \in [J, 2J]$ we have $\theta(j) \geq \theta(2J)$, hence pointwise on $[J, 2J]$, $j\,\theta(j) \geq j\,\theta(2J)$. Integrating this inequality over $[J, 2J]$ gives

$$\int_J^{2J} j\,\theta(j)\, dj \;\geq\; \theta(2J)\int_J^{2J} j\, dj \;=\; \theta(2J)\left[\frac{j^2}{2}\right]_J^{2J} \;=\; \frac{3}{2}J^2\,\theta(2J).$$

Since the left-hand side converges to 0 as $J \to \infty$, we conclude that

$$\lim_{J\to\infty} J^2\,\theta(2J) = 0.$$

Letting $j := 2J$ yields $\frac{1}{4}j^2\theta(j) \to 0$, hence

$$\lim_{j\to\infty} j^2\,\theta(j) = 0 \quad\Longrightarrow\quad \lim_{j\to\infty} \hat{\rho}^3\, j^2\,\theta(j) = 0.$$

We now study the cross term in the middle. We bound it by the two terms above. Using $2|ab| \leq a^2 + b^2$ with $a = \hat{\rho} j$ and $b = |\ln \theta(j)|$, we get, multiplying by $\hat{\rho}\ \theta(j)$, which is positive,

$$2\hat{\rho}^2\, j\, \theta(j)\, |\ln \theta(j)| \ \leq\ \hat{\rho}^3 j^2 \theta(j) + \hat{\rho}\theta(j)(\ln \theta(j))^2.$$

From above, since $j^2\theta(j) \to 0$ and $\theta(j)(\ln \theta(j))^2 \to 0$, it follows that

$$\lim_{j\to\infty} 2\hat{\rho}^2\, j\, \theta(j)\ \ln \theta(j) = 0.$$

Therefore all three expanded terms vanish, and thus

$$\lim_{j\to\infty} g(j)^2\, \omega_0(j) = 0,$$

so the upper boundary contribution of $\left[g(j)^2\omega_0(j)\right]_0^\infty$ is zero. Thus, equation (A.26) simplifies to:

$$\mathrm{Cov}_{\omega_0}\left(\rho(j), \ln \theta(j) + \hat{\rho} j\right) = -\frac{1}{2}\int_0^\infty g(j)^2 \rho(j) \omega_0(j)\, dj. \tag{A.27}$$

For any present-biased agent, the instantaneous discount rate is strictly positive ($\rho(j) > 0$), and the probability density is strictly positive ($\omega_0(j) > 0$). Since $g(j)$ is not identically zero (except under exponential discounting $\theta(j) = e^{-\hat{\rho} j}$), the integrand $g(j)^2\rho(j)\ \omega_0(j)$ is non-negative and strictly positive on a set of positive measure, so the integral is strictly positive. Thus, the right-hand side of equation (A.27) is strictly negative, which implies that $\lambda_S'(1) - \lambda_N'(1) > 0$.

□

Proof of Proposition 2. Throughout this proof we keep two integrals separate. The *resource constraint* (r-discounted) pins down the level of consumption an agent can afford given her wealth. The *welfare functional* (1) (θ-discounted) evaluates her psychological satisfaction along the equilibrium stream. The market interest rate r is constant in the AK model; the discount θ is a feature of the agent's taste.

By the modified Ramsey rule (3), the equilibrium consumption growth rate $\gamma_i = \sigma(r - \lambda_i(\sigma, r))$, $i \in \{N, S\}$, is constant over time, since σ, r, and $\lambda_i(\sigma, r)$ are all time-invariant. The equilibrium consumption path is therefore exponential:

$$c_i^\star(s)\ =\ c_i^\star(t)\, e^{\gamma_i j}, \qquad i \in \{N, S\}.$$

Integrating the budget constraint (2) forward at the market rate r and applying the transversality condition $\lim_{s\to\infty} e^{-r(s-t)} k_t(s) = 0$ yields the present-value form of the resource constraint,[11]

$$\int_0^\infty e^{-rj}\, c_i^\star(t + j)\, dj\ =\ A_t\ :=\ k_t\ +\ H_t, \qquad \text{where} \qquad H_t := \int_0^\infty e^{-rj}\, w_t(t + j)\, dj,$$

where A_t is total wealth (financial k_t plus wealth coming from labor income H_t) and r is the rate at which the technology physically transforms saving into capital. Substituting the equilibrium exponential path,

$$c_i^\star(t)\ =\ b_i\, A_t, \qquad b_i\ :=\ r - \gamma_i\ =\ (1 - \sigma)\, r + \sigma\, \lambda_i. \tag{A.28}$$

Thus, $c_i^\star(s)\ =\ c_i^\star(t + j)\ =\ c_i^\star(t)\, e^{\gamma_i j} = b_i\, A_t\, e^{\gamma_i j}$, for $i \in \{N, S\}$. Total wealth A_t is common across agent types because H_t depends only on the exogenous wage stream $w_t(\cdot)$ and the market rate r; only b_i differs across agent types, through λ_i.

Welfare for agent i is given by the integral (1), which discounts utility along the equilibrium stream using the agent's psychological discount θ, and applying the change of variables $j = s - t$:

$$\mathcal{W}_i\ =\ \int_t^\infty \theta(s - t)\, u\big(c_i^\star(s)\big)\, ds\ =\ \int_0^\infty \theta(j)\, u\big(c_i^\star(t + j)\big)\, dj\ =\ \int_0^\infty \theta(j)\, u\big(b_i A_t\, e^{\gamma_i j}\big)\, dj.$$

[11] One can easily generalize our constraint to incorporate income coming from wages. Without the wage channel, it is straightforward to see that $A_t = k_t$.

For CRRA u, expanding $u(c_0 e^{\gamma j}) = (c_0^{1-1/\sigma} e^{(1-1/\sigma)\gamma j} - 1)/(1-1/\sigma)$ and using $\int_0^\infty \theta(j)\, dj = 1/\hat{\rho}$ from (4) yields

$$\mathcal{W}_i \;=\; \frac{(b_i\, A_t)^{1-1/\sigma}}{1-1/\sigma}\, M(\eta_i) \;-\; \frac{1}{(1-1/\sigma)\,\hat{\rho}}, \qquad M(\eta) := \int_0^\infty \theta(j)\, e^{\eta j}\, dj, \tag{A.29}$$

with $\eta_i := (1-1/\sigma)\,\gamma_i = (\sigma-1)(r-\lambda_i)$. At $\sigma = 1$, (A.29) collapses to the log limit

$$\mathcal{W}_i\big|_{\sigma=1} \;=\; \frac{\ln(b_i\, A_t)}{\hat{\rho}} \;+\; \gamma_i\, T_1, \qquad T_1 := \int_0^\infty j\, \theta(j)\, dj. \tag{A.30}$$

Trivially, the observational equivalence holds, as at $\sigma = 1$, $\lambda_N(1) = \lambda_S(1) = \hat{\rho}$, so $b_S = b_N = \hat{\rho}$ and $\gamma_S = \gamma_N = r - \hat{\rho}$. The two equilibrium paths therefore coincide pointwise, and (A.30) delivers $\mathcal{W}_S(1) = \mathcal{W}_N(1)$.

Now we study the local welfare reversal. View $\mathcal{W}_i$ as $\mathcal{W}\big(\sigma, \lambda_i(\sigma, r)\big)$, with r held fixed throughout. By the chain rule,

$$\frac{\partial \mathcal{W}_i}{\partial \sigma} \;=\; \frac{\partial \mathcal{W}}{\partial \sigma}\bigg|_{\lambda_i(\sigma)} \;+\; \frac{\partial \mathcal{W}}{\partial \lambda}\bigg|_{\sigma,\,\lambda_i(\sigma)} \cdot\, \partial_\sigma \lambda_i(\sigma).$$

At $\sigma = 1$ both effective rates collapse to $\lambda_N(1) = \lambda_S(1) = \hat{\rho}$, so $\partial\mathcal{W}/\partial\sigma$ is evaluated at the *same* point $(1, \hat{\rho})$ for both agents and cancels in the difference $\partial_\sigma \Delta\mathcal{W}\big|_{\sigma=1}$. Only the λ-channel survives:

$$\frac{\partial\, \Delta\mathcal{W}}{\partial \sigma}\bigg|_{\sigma=1} \;=\; \frac{\partial \mathcal{W}}{\partial \lambda}\bigg|_{(1,\hat{\rho})} \Big(\partial_\sigma \lambda_S(1) - \partial_\sigma \lambda_N(1)\Big). \tag{A.31}$$

Differentiating (A.30) in λ at $(1, \hat{\rho})$ (with $\partial b/\partial\lambda = \sigma = 1$ and $\partial\gamma/\partial\lambda = -\sigma = -1$, and A_t inside a logarithm so its level cancels):

$$\frac{\partial \mathcal{W}}{\partial \lambda}\bigg|_{(1,\hat{\rho})} \;=\; \frac{1}{\hat{\rho}} \cdot \frac{1}{\hat{\rho}} \;-\; T_1 \;=\; \frac{1}{\hat{\rho}^2} \;-\; T_1. \tag{A.32}$$

Recall that $\omega_0(j) := \hat{\rho}\, \theta(j)$ is the density and let $f^E(j) := \hat{\rho}\, e^{-\hat{\rho} j}$ be the EPV-equivalent exponential density. By definition, both integrate to 1 on $[0, \infty)$. Their ratio is

$$\frac{\omega_0(j)}{f^E(j)} \;=\; \theta(j)\, e^{\hat{\rho} j}, \qquad \frac{d}{dj} \ln\big[\theta(j)\, e^{\hat{\rho} j}\big] \;=\; \hat{\rho} - \rho(j),$$

which is increasing in j since $\rho(j)$ is decreasing under present bias. Because $\hat{\rho} = \mathbb{E}_{\omega_0}[\rho(j)] := \int_0^\infty \rho(j)\omega_0(j)dj \in (\rho(\infty), \rho(0))$ and $\rho(j)$ is non-constant, the log-ratio is strictly negative for small j and strictly positive for large j. Hence $\theta(j)\, e^{\hat{\rho} j}$ is U-shaped and the densities ω_0 and f^E single-cross at a unique point $j^* \in (0, \infty)$, with $\omega_0 < f^E$ on $[0, j^*)$ and $\omega_0 > f^E$ on (j^*, ∞).

Single-crossing of densities with equal mass implies a strict mean inequality. Specifically, since $\int_0^\infty (\omega_0 - f^E)\, dj = 0$ and $j < j^*$ on the first subinterval (where $\omega_0 - f^E < 0$) while $j > j^*$ on the second (where $\omega_0 - f^E > 0$),

$$\int_0^\infty j\, \big(\omega_0(j) - f^E(j)\big)\, dj \;>\; j^* \int_0^\infty \big(\omega_0(j) - f^E(j)\big)\, dj \;=\; 0,$$

with strict inequality by present bias. Therefore $\mathbb{E}_{\omega_0}[j] := \int_0^\infty j\, \omega_0(j)\, dj \;>\; \int_0^\infty j\, f^E(j)\, dj \;=\; 1/\hat{\rho}$. Since $T_1 = \int_0^\infty j\, \theta(j)\, dj = \frac{1}{\hat{\rho}} \int_0^\infty j\, \omega_0(j)\, dj$, this gives

$$T_1 \;>\; \frac{1}{\hat{\rho}^2}.$$

Substituting (A.32) into (A.31) and using Lemma 4 , $\partial_\sigma \lambda_S(1) - \partial_\sigma \lambda_N(1) > 0$, gives part (ii).

□

A.4 A comment on Theorem 1 in Fishman (2020)

In Appendix A of Fishman (2020), Lemma 4 states that for any hyperbolic discount factor $H(t)$ and any $x > 0$:

$$\frac{1}{\int_0^\infty H(t)e^{-xt}dt} < \frac{1}{\int_0^\infty H(t)dt} + x. \tag{A.33}$$

We show that inequality (A.33) holds in the opposite direction for all hyperbolic discount functions and all $x > 0$. As a consequence, the upper bound employed in Equation (42) of the proof of Proposition 3(c) in Fishman (2020) does not hold.

Proof. Let $H(t)$ be a hyperbolic discount factor satisfying Definition 1. For $x \geq 0$, define the Laplace-type transform, which is the element in one of the denominators:

$$J(x) \equiv \int_0^\infty H(t)e^{-xt}dt. \tag{A.34}$$

The left-hand side of (A.33) is $1/J(x)$, while the right-hand side is $1/J(0) + x$. Both sides coincide at $x = 0$. Differentiating the left-hand side with respect to x yields:

$$\frac{d}{dx}\left(\frac{1}{J(x)}\right) = -\frac{J'(x)}{[J(x)]^2} = \frac{\int_0^\infty H(t)e^{-xt}t\,dt}{\left(\int_0^\infty H(t)e^{-xt}dt\right)^2}. \tag{A.35}$$

Because exponential scaling preserves log-convexity (hyperbolicity), the function $\widetilde{H}(t) \equiv H(t)e^{-xt}$ is itself also a hyperbolic discount factor for any $x \geq 0$. Applying Lemma 3 of Fishman (2020) to $\widetilde{H}(t)$, we have:

$$\int_0^\infty \widetilde{H}(t)t\,dt > \left(\int_0^\infty \widetilde{H}(t)dt\right)^2 \iff \int_0^\infty H(t)e^{-xt}t\,dt > \left(\int_0^\infty H(t)e^{-xt}dt\right)^2. \tag{A.36}$$

Dividing by the right-hand side confirms that the derivative strictly exceeds unity for all $x \geq 0$:

$$\frac{d}{dx}\left(\frac{1}{J(x)}\right) > 1 = \frac{d}{dx}\left(\frac{1}{J(0)} + x\right). \tag{A.37}$$

Integrating both derivatives from 0 to $x > 0$ yields: $\frac{1}{J(x)} - \frac{1}{J(0)} = \int_0^x \frac{d}{ds}\left(\frac{1}{J(s)}\right)ds > \int_0^x 1\,ds = x$, which immediately establishes that for all $x > 0$:

$$\frac{1}{\int_0^\infty H(t)e^{-xt}dt} > \frac{1}{\int_0^\infty H(t)dt} + x. \tag{A.38}$$

□

To see this concretely, consider the CCED function: $H(t) = \frac{1}{2}e^{-t} + \frac{1}{2}e^{-2t}$, which has an initial discount rate $d(0) = 1.5$ strictly declining to $\lim_{t\to\infty} d(t) = 1 > 0$, satisfying all model assumptions. Evaluate both sides at $x = 1$. First evaluate the unweighted integral on the RHS: $J(0) = \int_0^\infty \left(\frac{1}{2}e^{-t} + \frac{1}{2}e^{-2t}\right)dt = \frac{1}{2} + \frac{1}{4} = \frac{3}{4} \implies \frac{1}{J(0)} = \frac{4}{3} \approx 1,333$. The right-hand side of (A.33) is: RHS $= \frac{1}{J(0)} + 1 = \frac{4}{3} + 1 = \frac{7}{3} \approx 2.333$. Now we evaluate the discounted integral on the LHS: $J(1) = \int_0^\infty \left(\frac{1}{2}e^{-2t} + \frac{1}{2}e^{-3t}\right)dt = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}$. The left-hand side of (A.33) is: LHS $= \frac{1}{J(1)} = \frac{12}{5} = 2.400$.

Comparing the two terms LHS $= 2.400 >$ RHS $= 2.333$, confirming that the original inequality (A.33) is strictly reversed. Note that one can also get the difference between the LHS and the RHS, as $LHS - RHS = x/(3 * (3 + 2x))$, which is strictly positive for all $x > 0$.

A.5 Proof of Theorem 1

Lemma 5. *For every $b > 0$ with $\rho_\infty + \left(1 - \frac{1}{\sigma}\right)(b - r) > 0$,*

$$\frac{d\mathcal{W}(b)}{db} = k(t)^{1-\frac{1}{\sigma}}\, b^{-\frac{1}{\sigma}} \left[\int_0^\infty \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\, dj - b \int_0^\infty j\, \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\, dj \right]. \tag{A.39}$$

Moreover, when $\rho(j)$ is strictly decreasing,

$$\left(\int_0^\infty \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\, dj \right)^2 < \int_0^\infty j\, \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\, dj. \tag{A.40}$$

Proof. (A.39) is the derivative of (7). For (A.40) (Fishman, 2020, Lemma 3): the function $j \mapsto \ln \theta(j) - \left(1 - \frac{1}{\sigma}\right)(b - r)j$ is convex (its derivative $-\rho(j) - \left(1 - \frac{1}{\sigma}\right)(b - r)$ is nondecreasing) and vanishes at $j = 0$, hence it is superadditive, so

$$\theta(i)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)i}\, \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j} \le \theta(i+j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)(i+j)},$$

Integrating over $(i, j) \in [0, \infty)^2$ and changing variables to $v = i + j$,

$$\int_0^\infty \int_0^\infty \theta(i+j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)(i+j)}\, di\, dj = \int_0^\infty v\, \theta(v)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)v}\, dv.$$

□

Proof of Theorem 1. For the case $\sigma < 1$ a proof can be found in Fishman (2020). We give the proof for $\sigma > 1$. Fix a stationary point b of $\mathcal{W}$, so that by (A.39)

$$b \int_0^\infty j\, \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\, dj = \int_0^\infty \theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\, dj. \tag{A.41}$$

Define the probability densities on $[0, \infty)$

$$p(j) := \frac{\theta(j)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}}{\int_0^\infty \theta(i)\, e^{-\left(1-\frac{1}{\sigma}\right)(b-r)i}\, di}, \qquad q(j) := \frac{\theta(j)^\sigma e^{(\sigma-1)rj}}{\int_0^\infty \theta(i)^\sigma e^{(\sigma-1)ri}\, di}.$$

Step 1: equal means of the delay elasticity. Divide (A.12), with $x = \left(1 - \frac{1}{\sigma}\right)(b - r)$, by the normalising integral of p:

$$\int_0^\infty j \left(\rho(j) - \left(1 - \frac{1}{\sigma}\right) r \right) p(j)\, dj + \left(1 - \frac{1}{\sigma}\right) b \int_0^\infty j\, p(j)\, dj = 1.$$

By (A.41), $b \int_0^\infty j\, p(j)\, dj = 1$, hence

$$\int_0^\infty j \left(\rho(j) - \left(1 - \frac{1}{\sigma}\right) r \right) p(j)\, dj = 1 - \left(1 - \frac{1}{\sigma}\right) = \frac{1}{\sigma}.$$

Note that for every $\sigma > 0$ and real g with $\sigma\rho_\infty - g > 0$, and with $g = (\sigma - 1)r$, dividing (A.13), by σ times the normalising integral of q:

$$\int_0^\infty j \left(\rho(j) - \left(1 - \frac{1}{\sigma}\right) r \right) q(j)\, dj = \frac{1}{\sigma}.$$

Therefore

$$\int_0^\infty j \left(\rho(j) - \left(1 - \frac{1}{\sigma}\right) r \right) (p(j) - q(j))\, dj = 0. \tag{A.42}$$

Step 2: $\int_0^\infty \theta(j)^\sigma e^{(\sigma-1)rj}dj \le \int_0^\infty \theta(j)e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}dj$. The proof proceeds by contradiction: Suppose that the left-hand side is strictly larger. Using $\theta(j) = \exp\big(-\int_0^j \rho(i)\,di\big)$,

$$\ln\frac{p(j)}{q(j)} = \ln\frac{\int_0^\infty \theta(i)^\sigma e^{(\sigma-1)ri}\,di}{\int_0^\infty \theta(i)\,e^{-\left(1-\frac{1}{\sigma}\right)(b-r)i}\,di} + (\sigma-1)\int_0^j \left(\rho(i) - \left(1-\frac{1}{\sigma}\right)r\right)di - \left(1-\frac{1}{\sigma}\right)b\,j.$$

The right-hand side is concave in j, because its derivative $(\sigma-1)\big(\rho(j) - \big(1-\frac{1}{\sigma}\big)\,r\big) - \big(1-\frac{1}{\sigma}\big)\,b$ is nonincreasing, and it is strictly positive at $j = 0$ under the supposition. A concave function that is positive at zero has at most one zero on $(0,\infty)$ and is negative beyond it. Hence $\{j : p(j) \ge q(j)\}$ is an interval $[0, j^*]$. If $j^* = \infty$, then $p \ge q$ everywhere with $p > q$ near zero, contradicting $\int_0^\infty p(j)dj = \int_0^\infty q(j)dj = 1$; so $j^* < \infty$, $p \ge q$ on $[0, j^*]$ and $p < q$ on (j^*, ∞). By (8),

$$\big(p(j) - q(j)\big)\left[j\left(\rho(j) - \left(1-\frac{1}{\sigma}\right)r\right) - j^*\left(\rho(j^*) - \left(1-\frac{1}{\sigma}\right)r\right)\right] \le 0 \qquad \text{for all } j > 0,$$

and the inequality is strict: $p - q$ vanishes only at j^*, and $j\big(\rho(j) - \big(1-\frac{1}{\sigma}\big)\,r\big)$ is not constant, being positive and tending to zero as $j \to 0^+$ (recall that $\rho(0) < \infty$). Integrating, and using $\int_0^\infty \big(p(j) - q(j)\big)dj = 0$,

$$\int_0^\infty j\left(\rho(j) - \left(1-\frac{1}{\sigma}\right)r\right)\big(p(j) - q(j)\big)\,dj < 0,$$

which contradicts (A.42).

Step 3: $b < b_N$. By Step 2, (A.40) and (A.41),

$$\frac{1}{b_N} = \int_0^\infty \theta(j)^\sigma e^{(\sigma-1)rj}\,dj \le \int_0^\infty \theta(j)\,e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\,dj < \frac{\int_0^\infty j\,\theta(j)\,e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\,dj}{\int_0^\infty \theta(j)\,e^{-\left(1-\frac{1}{\sigma}\right)(b-r)j}\,dj} = \frac{1}{b}.$$

Step 4: monotonicity on $[b_N, \infty)$. By Proposition 1(ii), b_S exists. At b_S, (A.11) and (A.40) give

$$b_S\int_0^\infty j\,\theta(j)e^{-\left(1-\frac{1}{\sigma}\right)(b_S-r)j}dj > b_S\left(\int_0^\infty \theta(j)e^{-\left(1-\frac{1}{\sigma}\right)(b_S-r)j}dj\right)^2 = \int_0^\infty \theta(j)e^{-\left(1-\frac{1}{\sigma}\right)(b_S-r)j}dj,$$

so $\dfrac{d\mathcal{W}(b)}{db} < 0$ at $b = b_S$. By Proposition 1, $b_S > b_N$, and by Step 3 the continuous function $\dfrac{d\mathcal{W}(b)}{db}$ has no zero on $[b_N, \infty)$. It is therefore negative on $[b_N, \infty)$, and $\mathcal{W}(b_N) > \mathcal{W}(b_S)$.

□

A.6 Proof of Lemma 1

1. Tsoukis discount function. With the Tsoukis et al. (2017), Cabo et al. (2020a) discount function, for every $\sigma > 1$ if $r > \rho$ and $\sigma \in (1, \frac{r}{r-\rho})$ we have (note that here $\rho_\infty = \rho$.).

$$\frac{d}{dj}\left[j\left(\rho + \frac{\psi}{1+\delta j} - \left(1-\frac{1}{\sigma}\right)r\right)\right] = \underbrace{\rho - \left(1-\frac{1}{\sigma}\right)r}_{>0 \text{ from } \sigma\rho > (\sigma-1)r} + \frac{\psi}{(1+\delta j)^2} > 0,$$

so (8) holds and $\mathcal{W}_N > \mathcal{W}_S$ on the whole admissible region. Together with Theorem 1, for this type of discount function, the welfare ranking of the agents switches exactly at $\sigma = 1$, wherever the sophisticated equilibrium exists; for $\sigma < 1$ it exists whenever $\psi \le \delta$. The reason why $\sigma\rho > (\sigma-1)r$ for the Tsoukis–Cabo discount function is that inside the naive effective rate, one has, $\theta(j)^\sigma e^{(\sigma-1)rj} = (1+\delta j)^{-\sigma\psi/\delta}e^{-[\sigma\rho-(\sigma-1)r]j}$, which converges if $\sigma\rho - (\sigma-1)r > 0$. Figure A.1 shows the $\mathcal{D}(j,\sigma,r)$ object and its derivative. Note that only the admissible cases have positive slopes, while the case for $r = 5\%$ and $\sigma = 2$ showing a decreasing $\mathcal{D}$ is non admissible.

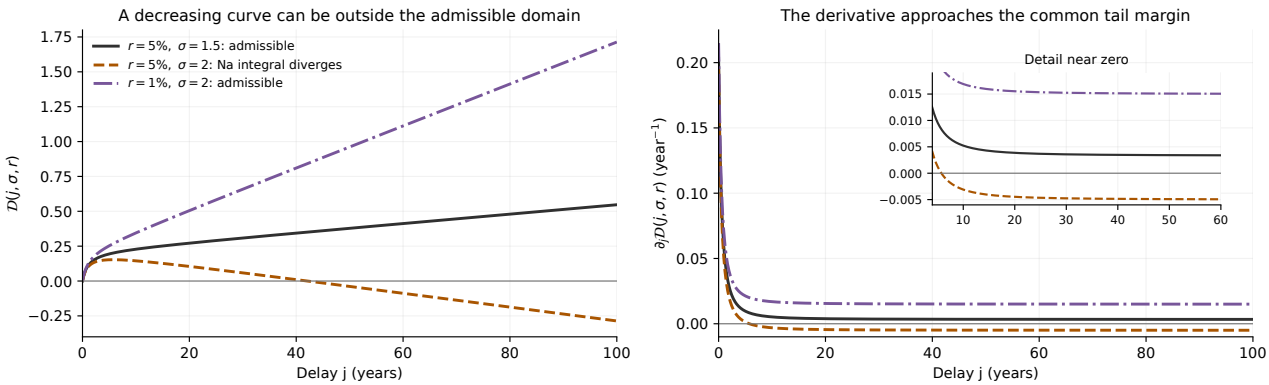


Figure A.1: Behavior of $\mathcal{D}(j,\sigma,r)$ for the Tsoukis-Cabo discount.

Notes: This figure shows the intuition for why the Tsoukis-Cabo discount delivers always higher welfare for the naive agent when $\sigma > 1$. The decreasing $\mathcal{D}$ is actually not admissible because $\sigma\rho < (\sigma-1)r$. Every curve uses $EPV = 25$, $\rho = 0.02$, $\psi = 0.2$, $\delta = 0.9261426339$.

2. Mixture of exponentials. For the convex combination of exponentials, the instantaneous discount rate is $\rho(j) = \frac{z\rho_L e^{-\rho_L j}+(1-z)\rho_S e^{-\rho_S j}}{ze^{-\rho_L j}+(1-z)e^{-\rho_S j}}$. Note that now $\rho_\infty = \rho_L$. Define the date-j weight on the high rate as $w(j) = \frac{(1-z)e^{-(\rho_S-\rho_L)j}}{z+(1-z)e^{-(\rho_S-\rho_L)j}}$. Then, one can write $\rho(j) = \rho_L + (\rho_S-\rho_L)w(j)$ and $\rho'(j) = -(\rho_S-\rho_L)^2 w(j)[1-w(j)]$. Therefore, from $\mathcal{D}(j,\sigma,r)$ defined in Theorem 1, one gets

$$\partial_j \mathcal{D}(j,\sigma,r) = \rho_L - \left(1-\frac{1}{\sigma}\right) r + (\rho_S-\rho_L)w(j) - j(\rho_S-\rho_L)^2 w(j)[1-w(j)],$$

which has a negative final term. Importantly, even at intermediate delays, it can be large enough to make the derivative negative even when the tail condition holds, i.e., $\rho_L - \left(1-\frac{1}{\sigma}\right) r > 0$. That is the crucial distinction from the Tsoukis–Cabo case: a positive limiting derivative does not guarantee positive derivatives at intermediate dates (there may be periods where it is decreasing, which implies the condition in Theorem 1does not hold).[12]

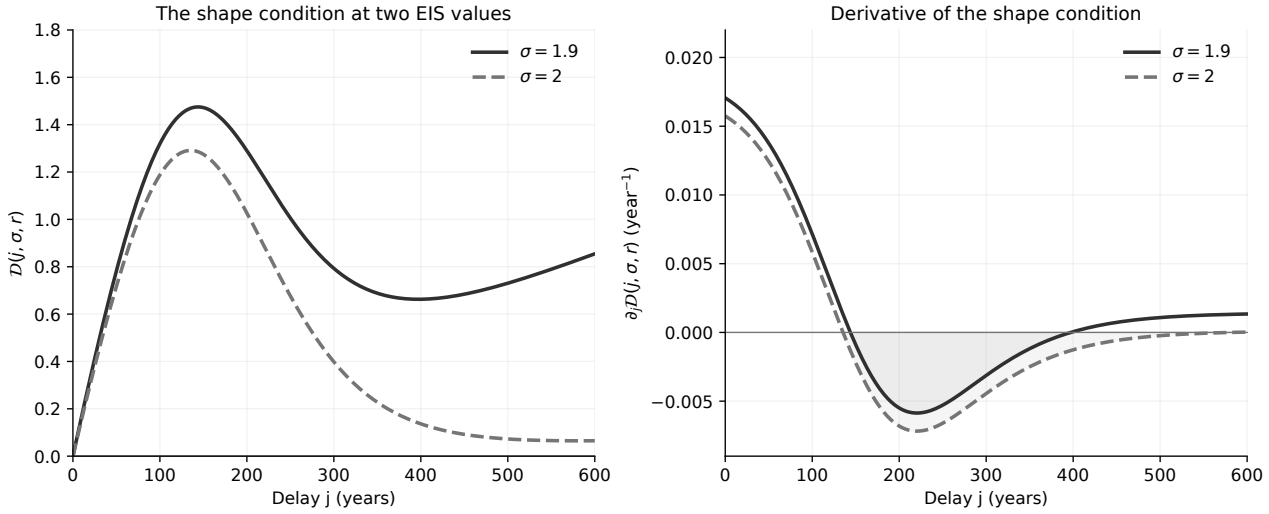


Figure A.2: Behavior of $\mathcal{D}(j,\sigma,r)$ for the convex combination of exponentials.

Notes: This figure shows a counterexample of why the CCED function can deliver higher welfare for the sophisticated agent when $\sigma > 1$. Parameters used are $\hat{\rho} = 0.04, r = 0.05, \rho_L = 0.0251, \rho_S = 0.042, z = \frac{251}{3380}$.

[12]The theorem establishes the implication $\mathcal{D}$ nondecreasing $\implies$ $\mathcal{W}_N > \mathcal{W}_S$. When $\mathcal{D}$ decreases somewhere, the theorem no longer determines the ranking.

To establish the welfare reversal, consider the first parameterization in Table 1, with exact EPV matching: $r = 0.05, \quad \sigma = 2, \quad z = 0.08, \quad \rho_S = 0.04219, \quad \rho_L = \frac{z}{25-(1-z)/\rho_S} = \frac{8438}{336875}$. Then $\frac{z}{\rho_L} + \frac{1-z}{\rho_S} = 25, \quad 2\rho_L - r > 0$. Thus $\hat{\rho} = 0.04$ and the strict tail condition holds. At $\sigma = 2$, expanding the square of the mixture in (A.10) gives

$$\frac{1}{b_N} = \frac{z^2}{2\rho_L - r} + \frac{2z(1-z)}{\rho_L + \rho_S - r} + \frac{(1-z)^2}{2\rho_S - r}.$$

The sophisticated equation (A.11) becomes

$$1 = b_S \left[\frac{z}{\rho_L + (b_S - r)/2} + \frac{1-z}{\rho_S + (b_S - r)/2} \right].$$

Equivalently, b_S is the unique positive root of

$$b_S^2 + 2(2z-1)(\rho_S - \rho_L)b_S - (2\rho_L - r)(2\rho_S - r) = 0.$$

These expressions give $b_N \approx 0.009998898365$, and $b_S \approx 0.028912620344$. For common initial capital $k(t) = 1$, (7) simplifies to

$$\mathcal{W}(b) = 2\sqrt{b} \left[\frac{z}{\rho_L + (b-r)/2} + \frac{1-z}{\rho_S + (b-r)/2} \right] - 50.$$

Substituting the two propensities gives $\mathcal{W}_N \approx -38.5384002164$, and $\mathcal{W}_S \approx -38.2378620420$, and therefore $\mathcal{W}_S - \mathcal{W}_N \approx 0.3005381744 > 0$. Moreover, $\partial_j \mathcal{D}(200, 2, 0.05) \approx -0.00692395481 < 0$, so $\mathcal{D}$ decreases over a nonempty interval.

Figures A.1 and A.2 show the intuition of Theorem 1 and Lemma 1. When $D(j, \sigma, r)$ is nondecreasing, as in the Tsoukis-Cabo discount (Figure A.1), the naive agent is better off than the sophisticated agent when $\sigma > 1$. However, under the CCED discount,the sophisticated agent can be better off, even for values of the EIS above one. In Figure A.2 that $D(j, \sigma, r)$ is decreasing for $\sigma = 1.9$ at the beginning, but in the long run increases again. However, for $\sigma = 2$, has a longer period where it decreases.

Finally, Figure A.3 (high interest rate and positive growth) and Figure A.4 (low interest rate and negative growth) summarize why these two discount specifications produce such different welfare outcomes. Under the generalized hyperbolic discounting of Tsoukis et al. (2017) and Cabo et al. (2020a), a welfare reversal above unit elasticity is not possible, because $D(j) \equiv j\big(\rho(j) - (1 - 1/\sigma)r\big)$ is strictly increasing across the entire admissible domain, ensuring that lifetime welfare $\mathcal{W}(b)$ remains strictly single-peaked and that naïveté globally dominates for all $\sigma > 1$.

In contrast, a convex combination of exponentials provides the flexibility required to overturn this result. When short-run and long-run impatience are sufficiently far apart, the discount rate falls sharply over intermediate delays, pulling $D(j)$ downward. This dip allows $\mathcal{W}(b)$ to develop a second local peak, enabling the sophisticated household to achieve higher welfare once σ is sufficiently large.

Panel (a) plots the instantaneous discount rate. The two functions satisfy the same EPV restriction, $\int_0^\infty \theta(j)\, dj = \int_0^\infty e^{-\hat{\rho} j}\, dj = 1/\hat{\rho}$, but distribute their discounting differently across delays. Equal EPV fixes one integral; it leaves considerable freedom in the shape of discounting. Panel (b) shows how the two exponentials discount gives two local maxima (it has three critical points). The naive and sophisticated propensities lie around the second peak. The sophisticated propensity b_S lies slightly beyond that peak, but its welfare remains higher than the naive agent's. Panel (c) shows that the Tsoukis–Cabo discount has one maximum. Both agents, naive and sophisticated chose propensities to the right of the maximum point: $b^* < b_N < b_S$. Moving from the naive propensity to the larger sophisticated propensity moves farther down the welfare curve. The naive agent therefore has higher welfare. Panel (d) varies σ while keeping the two discount calibrations and the common interest rate. Positive values mean that the sophisticated agent has higher welfare (red area). Negative values mean naive has higher welfare (blue area). The white circles identify equal-welfare crossings (when the function crosses the horizontal axis), and the black circle is the minimum for the Tsoukis-Cabo discount.

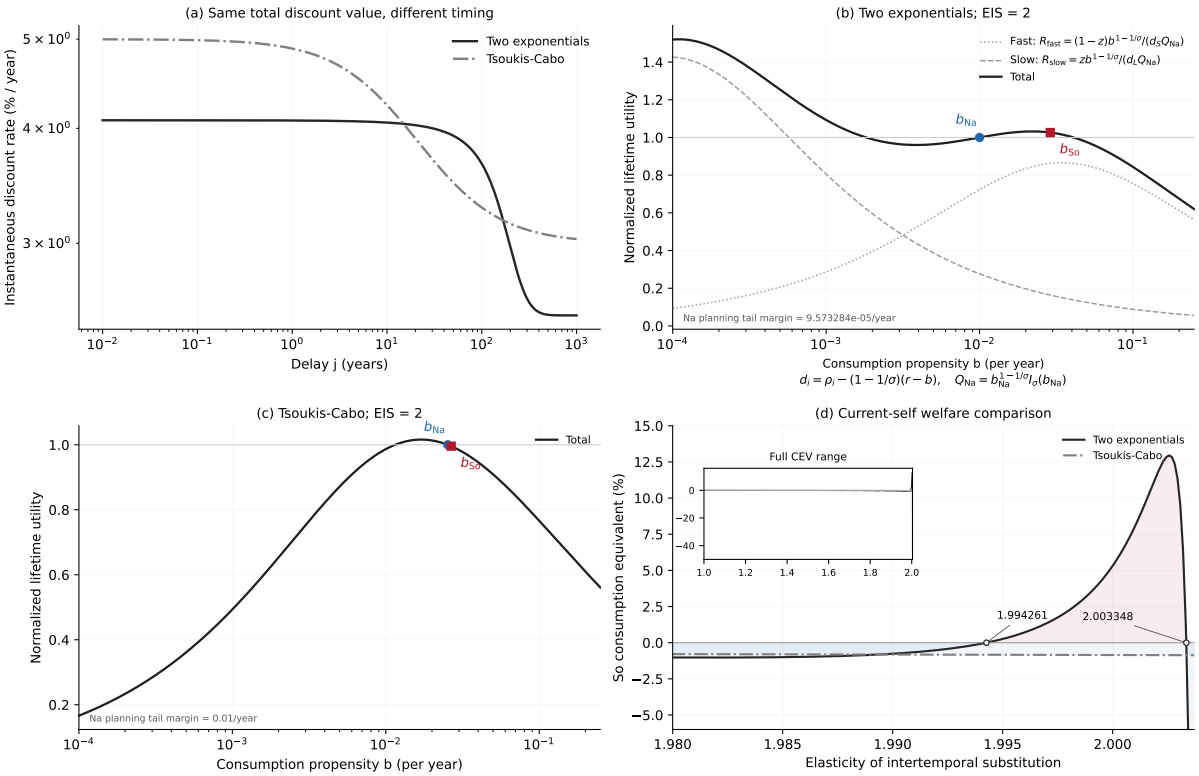


Figure A.3: Welfare comparisons, high interest rate

Notes: Panel (a) compares instantaneous discount rates. Panels (b) and (c) show welfare across different consumption propensities b. Panel (d) changes the EIS while preserving the discount calibrations and common return r. Neutral line styles distinguish the two families; red and blue areas identify Sophisticated and Naive welfare advantages respectively. Its two located mixture crossings above one are near 1.994261 and 2.003348.

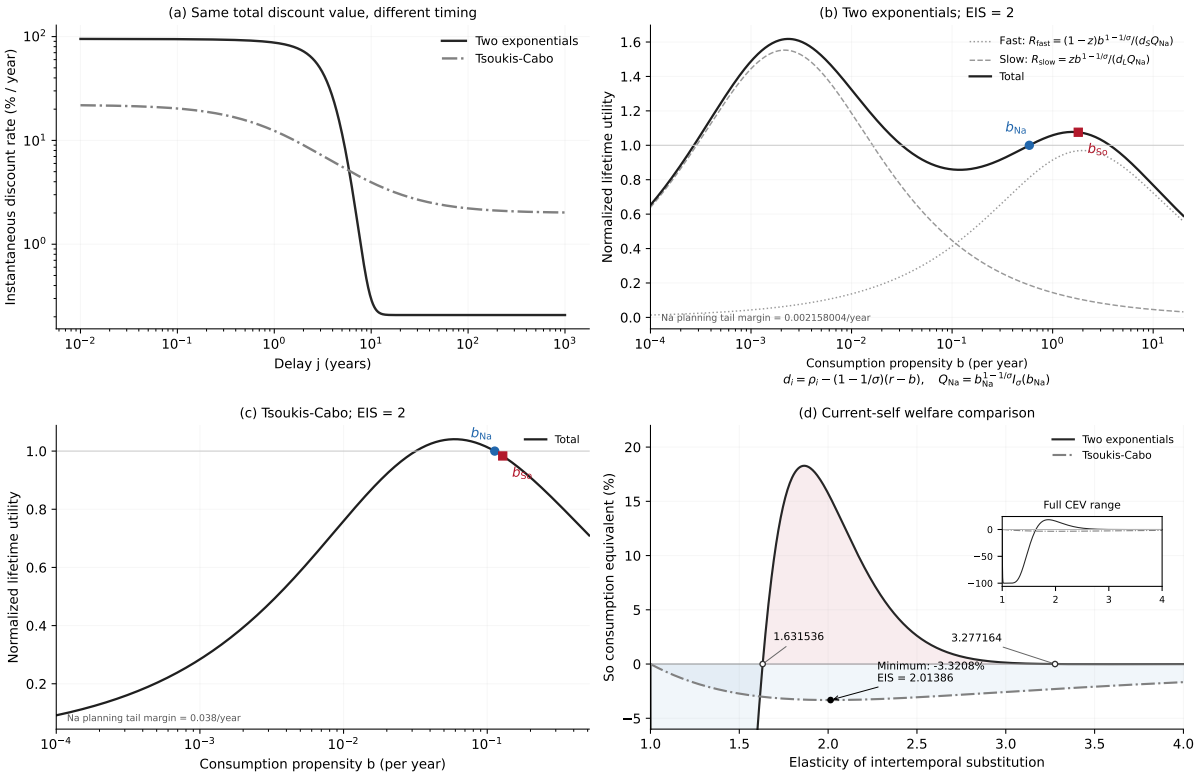
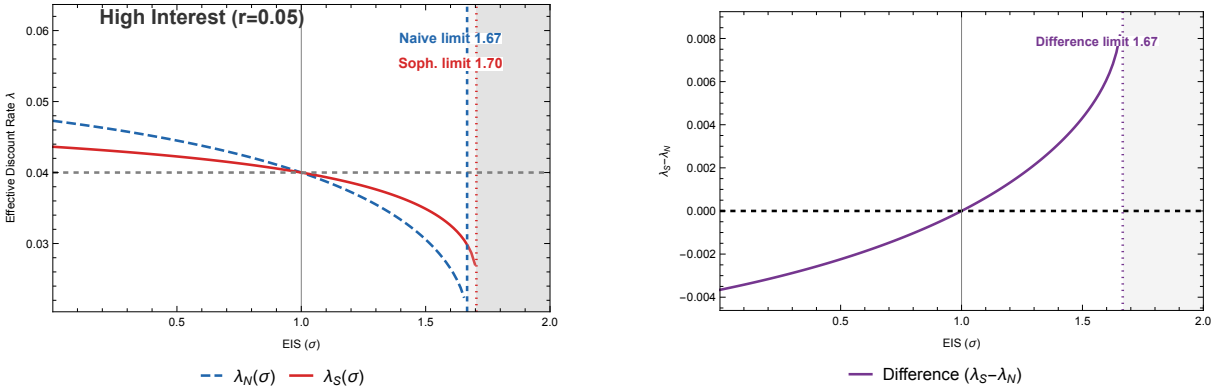


Figure A.4: Welfare comparisons, low interest rate

The two-exponential example has a positive region when EIS>1, including $\sigma = 2$, when sophisticated agent become the better off agent, and the sophisticated agent under Tsoukis–Cabo remains below zero over the displayed above-one range (meaning that sophisticated are worse off).

A.7 Other Illustrations

Figure A.5: Effective Discount Rates under positive growth ($r > \lambda_i$, High Interest)

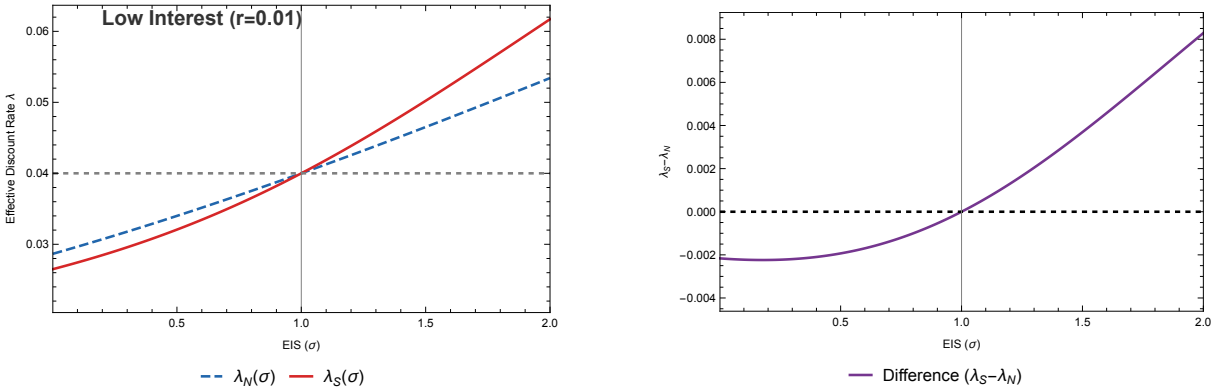


(a) Effective Discount Rates (λ)

(b) Difference ($\lambda_S - \lambda_N$)

Notes: Panel (a) compares the *effective discount rates* for naive, sophisticated, and exponential agents. Panel (b) isolates the difference between the sophisticated and naive effective rates. Under logarithmic utility ($\sigma = 1$), the effective discount rates collapse to the same constant, so the difference equals zero at $\sigma = 1$.

Figure A.6: Effective Discount Rates under negative growth ($r < \lambda_i$, Low Interest)



(a) Effective Discount Rates (λ)

(b) Difference ($\lambda_S - \lambda_N$)

Notes: Panel (a) compares the *effective discount rates* for naive, sophisticated, and exponential agents in a low interest rate environment. Panel (b) shows the difference $\lambda_S - \lambda_N$, which is strictly negative for $\sigma < 1$, crosses zero exactly at the logarithmic benchmark $\sigma = 1$, and turns strictly positive for $\sigma > 1$.

Figures A.5 and A.6 compare the effective discount rates across the admissible range of the EIS, using our Proposition 1. In both the high interest rate environment (Figure A.5) and the low interest rate environment (Figure A.6), the effective discount rates for the naive and sophisticated agents coincide at $\sigma = 1$ and diverge on either side, exactly as predicted in the text. For $\sigma < 1$, the sophisticated agent has a lower effective discount rate than the naive agent ($\lambda_S < \lambda_N$), consistent with the conjecture of Gómez (2025), whose exhaustive numerical exploration supports this ordering under hyperbolic discounting. For $\sigma > 1$, the ordering reverses: $\lambda_S > \lambda_N$, meaning the sophisticated agent discounts the future more heavily. The right-hand panels isolate the difference $\lambda_S - \lambda_N$, which crosses zero precisely at $\sigma = 1$.

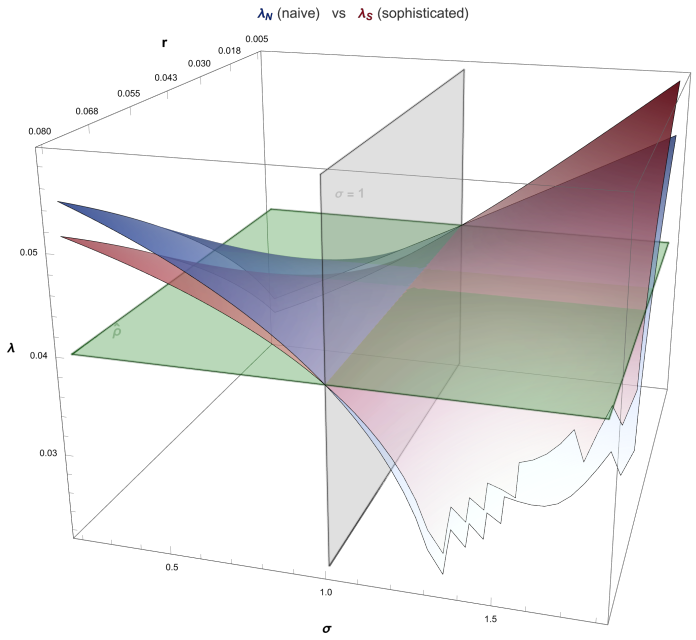


Figure A.7: Full Surface $\lambda_i(\sigma, r)$ over the (σ, r)-Plane

Notes: The figure shows the full effective rate of time preferences $\lambda_i(\sigma, r)$ jointly as a function of the EIS (σ) and the interest rate (r), for the two behavioral types: naive (λ_N, blue surface) and sophisticated (λ_S, red surface). The two surfaces intersect along the line $\sigma = 1$, where $\lambda_N(1, r) = \lambda_S(1, r) = \hat{\rho}$, and the ordering reverses across this threshold: $\lambda_S < \lambda_N$ for $\sigma < 1$ and $\lambda_S > \lambda_N$ for $\sigma > 1$, as established in Proposition 1. Figures A.5 and A.6 display 2D cross-sections of this surface at the fixed values $r = 0.05$ (high interest) and $r = 0.01$ (low interest), respectively. Discount-function specification: $\theta(j) = (1 + \delta j)^{-\psi/\delta} e^{-\rho j}$.

Gómez's numerical finding is thus confirmed analytically as one side of a more general result whose other side ($\sigma > 1$) had remained unexplored, and leads to the opposite conclusion.

Figure A.7 visualises the full surface $\lambda_i(\sigma, r)$ over the (σ, r)-plane and makes Proposition 1 geometrically transparent: the naive (blue) and sophisticated (red) surfaces intersect along the line $\sigma = 1$, where both rates collapse to $\hat{\rho}$, and the ordering reverses across this threshold for every admissible r. Figures A.5 and A.6 are 2D cross-sections of this surface at $r = 0.05$ and $r = 0.01$ respectively.

Copyright © 2026 @ the author(s). Discussion papers are in draft form. This discussion paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.