Who Saves More, the Naive or the Sophisticated Agent?*

Max Groneck† Alexander Ludwig‡ Alexander Zimper§

December 9, 2022

Abstract

We consider a class of additively time-separable life-cycle consumption-savings models with iso-elastic per period power utility featuring resistance to inter-temporal substitution of θ with linear consumption policy functions. The utility maximization problem is dynamically inconsistent for almost all specifications of effective discount factors. Pollak (1968) shows that the savings behavior of a sophisticated and a naive agent is identical with logarithmic utility ($\theta = 1$). We extend this result by showing that the sophisticated agent saves in any period a greater fraction of her wealth than the naive agent if and only if $\theta \geq 1$, irrespective of the discount function.

JEL Classification: D15, D91, E21.

Keywords: Life-Cycle Model; Discount Functions; Dynamic Inconsistency, Epstein-Zin-Weil Preferences

---

*We thank Daniel Harenberg, François Le Grand, Peter Wakker and seminar participants at the University of Alabama for helpful comments and discussions. Alex Ludwig gratefully acknowledges financial support by Research Center SAFE, funded by the State of Hessen initiative for research LOEWE.

†Department of Economics, Econometrics and Finance; University of Groningen; Netspar; AJSPH; P.O. Box 800 9700AV Groningen, The Netherlands: E-mail m.groneck@rug.nl

‡Goethe University Frankfurt, ICIR, SAFE & CEPR; Theodor-W.-Adorno-Platz 3; 60629 Frankfurt am Main, Germany; E-mail: mail@alexander-ludwig.com

§Department of Economics; University of Pretoria; Private Bag X20; Hatfield 0028; South Africa; E-mail: alexander.zimper@up.ac.za

1 Introduction

How time preferences and beliefs about the future affect consumption savings decisions is a classical economic question. The workhorse model of inter-temporal allocation is the life-cycle model of Modigliani and Brumberg (1954) and Ando and Modigliani (1963). Standard deterministic models assume an additively separable per period utility function, where future utility is discounted exponentially (Samuelson 1937). Models with survival beliefs express those as additive probability measures. *Standard discounting* as the combination of exponential time-discounting with additive survival beliefs results in a dynamically consistent life-cycle model in which the future selves of the economic agent have no incentives to deviate from her ex ante optimal consumption and savings plan. This paper instead analyzes life-cycle models with arbitrary effective discount factors. In this setup the generic case is *dynamic inconsistency*, that is, the optimal consumption plan from the perspective of some ex-ante agent does—for almost all specifications of discount factors—not coincide with the optimal consumption plan from the perspective of some ex-post agent.

We follow the literature since Strotz (1956) and Pollak (1968) and compare a naive agent—who does not foresee that her future selves will deviate from the current self's optimal consumption savings plan—with a sophisticated agent—who is aware of these deviating incentives. Our main research question is as follows:

> *How does the underlying effective discounting process impact on the question whether the naive or the sophisticated agent saves a greater fraction of her wealth in any given time-period?*

From an economic policy perspective this question is relevant because governments worldwide look for ways to induce more prudent savings behavior. If it was the case that a sophisticated agent will always save more than a naive agent under an empirically relevant discounting scenario such as, e.g., hyperbolic time discounting (Laibson 1997), awareness campaigns about people's dynamic inconsistencies may usefully complement financial incentives schemes. Closer to our own theoretical interests is the comparison between our economic intuition and the formal implications of life-cycle models. Intuitively, we had expected that the nature of effective discounting processes would have some impact on the question whether a naive agent saves more than her naive counterpart or vice versa.

We address our research question within a class of life-cycle models with a per period iso-elastic power utility function in which the consumption policy functions are linear in total wealth. We denote the concavity parameter of the per period utility function by θ so that $1/\theta$ measures the inter-temporal elasticity of substitution (IES). For this class of models our answer to the posed research question proves that our intuition was wrong:

Irrespective of the effective discount process, the sophisticated agent will save more than her naive counterpart in any given time period if and only if the IES *is smaller than one.*

Thus, the question which agent type saves a greater fraction of her wealth in any given time period is exclusively determined by the IES. This holds regardless of whether in any future period the naive agent consumes—or plans to consume—more or less than her sophisticated counterpart. We regard this finding as surprising because the effective discount process is the sole reason for why the model might be dynamically inconsistent to begin with.

Our result extends Pollak (1968)'s finding that for an IES of unity the naive and the sophisticated agent save exactly the same fraction of their wealth in every period irrespective of their effective discount function. This knife-edge case is often taken as a reference point for interpreting consumption behavior in models with a presence bias.[1] In these models, the sophisticated agent's life-cycle savings decision is shaped by two opposing forces—on the one hand, (i) restricting her future selves decisions because they understand they will over-consume and, on the other hand, (ii) providing sufficient resources to finance her consumption needs and thereby to smooth consumption. One may therefore conclude that the answer to our research question is obvious for models with a presence bias: With a low value of the inter-temporal elasticity of substitution (IES) relative to a value of one, the sophisticated agent cares more about inter-temporal consumption smoothing and thus it is plausible that motive (ii) is stronger than motive (i) and vice-versa for a high value of the IES.

Even in these models with a presence bias, however, this intuition provides only limited guidance because it relates to an *intra-personal* comparison of the consumption behavior of the sophisticated agent and not to the *inter-personal* comparison across the two types of agents. In fact, our main result shows that this intuition is misleading because irrespective of the shape

[1] A related literature studies observational (in)equivalence results to models with exponential discounting, where $\theta = 1$ serves as a reference point, cf., e.g., Strulik (2015), Cabo et al. (2016) and Cabo et al. (2020).

of the discount function—i.e., whether it induces a *presence* or a *future* bias—the sophisticated agent saves more than the naive agent if and only if the IES is less than one.

We start with a cake eating problem with finite horizon $2 \leq T < \infty$. In this model, we assume that households discount the future with age-dependent effective discount factors $\rho_{h,t}$, with $\rho_{h,t} > 0$ and $\rho_{t,t} = 1$, where h is the current age of the household and t is the age with the respective future period consumption delivery. We are agnostic about the nature of the discount process. In deterministic models it reflects pure time-discounting and in models with survival uncertainty a combination of pure time-discounting and survival beliefs.[2] Because the discount factors of the h-old agent, $h = 0, ..., T-1$, can be any strictly positive real-numbers, our life cycle model is very general and it encompasses relevant extensions of the standard model such as (quasi-)hyperbolic time-discounting models (cf. Phelps and Pollak 1968; Laibson 1997; 1998; O'Donoghue and Rabin 1999; Harris and Laibson 2001) and Choquet expected utility or/and Prospect theory life-cycle models with non-additive subjective survival beliefs (cf. Bleichrodt and Eeckhoudt 2006; Ludwig and Zimper 2013; Drouhin 2015; Groneck, Ludwig, and Zimper 2016; Grevenbrock, Groneck, Ludwig, and Zimper 2021 and references therein). To make this latter point explicit, one can show that the discount factors of an h-old Choquet expected utility decision maker are $\rho_{h,t} = \beta_{h,t}\nu_{h,t}$, where $\beta_{h,t}$ stands for pure time-discounting between present age h and future age t and $\nu_{h,t}$ stands for the decision maker's non-additive belief to survive from age h to age t.[3]

To deal with the generic case of dynamic inconsistency, we solve the life-cycle model for the realized consumption paths of a sophisticated and a naive agent, respectively. Despite the fact that both agents share the same preferences, their realized consumption paths result from very different optimization problems. The sophisticated agent chooses her per-period consumption as if she plays a strategic game against her future selves. In contrast, the naive agent chooses her per-period consumption under the misperception that her future selves will stick to the consumption plan that is optimal from her ex ante perspective.

Denoting by m_h^i, for $i \in \{n, s\}$ the marginal propensity to consume out of total wealth w_h

[2] Compare, e.g., Halevy (2008), Epper, Fehr-Duda, and Bruhin (2011), Saito (2011), Chakraborty, Halevy, and Saito (2020) who discuss the relationship between pure time-preferences and preferences under uncertainty or/and risk.

[3] A formal proof can be found in the earlier version of this paper (Groneck, Ludwig, and Zimper 2021).

at age h of the $n$aive and the $s$ophisticated agent, respectively, we derive as our main result Theorem 1, which states that for all (arbitrary) specifications of the effective discount factors[4], (i) $\theta < 1$ implies $m_h^n \leq m_h^s$ and $\theta > 1$ implies $m_h^n \geq m_h^s$. This result is directly derived from a comparison of the respective MPCs, thus it holds globally and hinges on the linearity of consumption policy functions. For the special case of a three-period model ($T = 2$) we further show analytically that the concavity parameter θ additionally determines the strength[5] of the response of the sophisticated relative to the naive agent's savings decision in response to *any* (local) departure from *standard discounting*. This straightforward first-order condition argument thus directly illustrates for the three-period model that only the concavity parameter θ but not the effective discount process drives our main finding. Beyond determining the degree of inter-temporal smoothing through the $IES = \frac{1}{\theta}$, the concavity parameter θ therefore plays an additional role for the *inter-personal* comparison of savings behavior between the naive and the sophisticated agent. In a quantitative illustration we further show that the differences in MPCs across the two types of agents can be large for plausible values of θ.

Finally, we extend our main finding in Theorem 1 to models with return risk, a portfolio choice and homothetic Epstein-Zin-Weil (EZW) preferences (Epstein and Zin 1989; Epstein and Zin 1991; Weil 1989), again for the multi-period model where $T \geq 2$. This also encompasses models with risky human capital (Krebs 2003) and thus may also allow for self imposed borrowing constraints if there exists a strictly positive probability of loosing the wealth endowment from a risky investment. Our main result therefore holds in a broad class of economic models. What is decisive is the linearity of consumption policy functions. With this extension we also formally establish that it is the value of the deviation of the IES from unity which governs the relative consumption-savings response of the two agents to deviations from geometric discounting and not risk aversion.[6]

[4]At ages $h \in \{T-1, T\}$ we always have $m_h^n = m_h^s$ irrespective of the value of θ.

[5]The impact of θ on the inter-personal comparison between both types is non-monotonic. We show for $\theta \geq 1$ that the relative size of the reaction of a sophisticated over a naive agent to *any* deviation from standard discounting increases up to a threshold value $\bar{\theta} > 1$ whereby it decreases afterwards. This non-monotonicity in $\theta \geq 1$ arises because the naive and the sophisticated agent have identical MPCs for $\theta = 1$ as well as in the limit $\theta = \infty$.

[6]In this extension, θ takes (at least) a triple role as a consumption smoothing parameter, a relative smoothing parameter of the naive and sophisticated savings response to deviations from standard discounting and as a parameter partially controlling precautionary savings behavior. For the latter point see Kimball and Weil (2009)

The remainder of our analysis proceeds as follows. Section 2 solves the model for the realized consumption path of the sophisticated agent as well as for the planned versus realized consumption paths of the naive agent. Section 3 formally defines dynamic consistency versus inconsistency of our life-cycle model in terms of the realized versus planned MPCs of the naive agent. Section 4 comprehensively answers our research question. Section 5 presents the extension to EZW preferences, and Section 6 concludes. All formal proofs are relegated to the Appendix.

2 The Life-Cycle Model

We study an additively time-separable life-cycle model with final period $T \geq 1$ such that every h-old agent's life-time utility over the consumption stream $(c_h, c_{h+1} ..., c_T) \in \mathbb{R}_{>0}^{T-h+1}$ is given as

$$U_h (c_h, c_{h+1} ..., c_T) = \sum_{t=h}^{T} \rho_{h,t} u (c_t), \tag{1}$$

where the age-dependent effective discount factors must only satisfy $\rho_{h,t} > 0$ and $\rho_{t,t} = 1$. There exists an initial amount of total wealth $w_0 > 0$ that the agent can spend over her life-cycle so that the budget constraint becomes

$$w_{t+1} = w_t - c_t \geq 0 \quad \text{for } t \in \{0, 1, \ldots, T-1\}. \tag{2}$$

We restrict attention to period-utility functions belonging to the family of iso-elastic power utility functions, that is, $u(c)$ must be differentiable on $\mathbb{R}_{>0}$ such that

$$u(c) = \frac{c^{1-\theta}}{1-\theta} \tag{3}$$

and $u'(c) = c^{-\theta}$ for concavity parameter $0 < \theta < \infty$.

2.1 Optimal Consumption Plan

For fixed period consumption c_t and wealth w_t let

$$c_t = m_t w_t$$

and Krueger, Ludwig and Villalvazo (2021).

where m_t denotes the agent's *marginal propensity to consume* (MPC). Because the optimal period consumption is linear in total wealth for power period utility functions, it will sometimes be convenient to consider MPCs rather than absolute consumption levels.[7] Expressed in terms of MPCs for the periods $h+1,...,T$ and period h wealth lifetime utility (1) of the h-old agent from consumption stream $(c_h,...,c_T)$ becomes

$$U_h\left(c_h;m_{h+1},...,m_T,w_h\right)=u(c_h)+\sum_{t=h+1}^{T}\rho_{h,t}u\left(\left(w_h-c_h\right)m_t\prod_{j=h+1}^{t-1}\left(1-m_j\right)\right). \tag{4}$$

Next we derive the MPCs that would maximize this utility function from the perspective of the h-old agent. In what follows, we denote by $m_h^{*,h}:[0,1]^{T-h}\rightarrow[0,1]$ the function that gives us, for any given argument

$$\left(m_{h+1},...,m_T\right)\in[0,1]^{T-h},$$

the (unique) MPC that maximizes through the absolute consumption level

$$c_h^{*,h}=m_h^{*,h}\left(m_{h+1},...,m_T\right)w_h$$

the utility function (4) over all admissible consumption levels c_h. In game-theoretic terms, $m_t^{*,h}\left(m_{h+1},...,m_T\right)$ would correspond to the best reply/response of an h-old agent who assumes that her future selves will be choosing $\left(m_{h+1},...,m_T\right)$ as their respective MPCs.

Proposition 1. *The MPCs $m_t^{*,h}$ that are optimal from the perspective of the h-old agent for fixed $m_{t+1},...,m_T$ are given as*

$$m_t^{*,h}\left(m_{t+1},...,m_T\right)=\begin{cases}1 & \text{for } t=T\\ \dfrac{1}{1+\left(\sum_{s=t+1}^{T}\frac{\rho_{h,s}}{\rho_{h,t}}\left(m_s\prod_{j=t+1}^{s-1}(1-m_j)\right)^{1-\theta}\right)^{\frac{1}{\theta}}} & \text{for } h\leq T-1\end{cases} \tag{5}$$

For $T\geq 2$ our life-cycle model will be, generically, dynamically inconsistent in the sense that for almost all specifications of discount factors there is some t-old agent with $t>h$, where $h\leq T-2$, who will have a strict incentive to deviate from a consumption plan that would be optimal from the perspective of the h-old agent. To solve for models that might be dynamically inconsistent, the literature distinguishes between the two extreme cases of a naive versus a

[7] Linearity of consumption policy functions in models with a deterministic labor income stream and no borrowing constraints is a well-established result in the consumption literature, cf., e.g., Deaton (1992).

sophisticated agent (cf. O'Donoghue and Rabin 1999). The remainder of this section defines both types of agents in terms of the optimal MPCs of Proposition 1.

2.2 Sophisticated versus Naive Saving Choices

In game-theoretic terms, $m_t^{*,h} : [0,1]^{T-h} \to [0,1]$ given by (5) is the h-old agent's best response function according to which she chooses for a given wealth level w_t the utility maximizing consumption level

$$c_t^{*,h} = m_t^{*,h}\left(m_{t+1}, ..., m_T\right) w_t$$

for the t-old agent whereby she assumes that the agents who are older than t choose

$$\left(m_{t+1}, ..., m_T\right) \in [0,1]^{T-t}$$

as their respective MPCs. In what follows we distinguish between an agent who is either sophisticated or naive throughout her whole life-cycle. Whereas the h-old sophisticated agent chooses a best response against the actual savings behavior of all her future selfs, the h-old naive agent chooses a best response against her most preferred savings behavior of her future selfs—which may or may not coincide with the actual savings behavior of these future selfs.

2.2.1 The Sophisticated Agent

Definition 1. *We speak of a "sophisticated agent" if and only if this agent correctly anticipates at every age h her future behavior.*

Denote by m_t^s the realized MPC of the t-old sophisticated agent. Expressed in terms of the optimal MPCs, the sophisticated agent solves through backward induction at every age $h \geq 0$ the problem

$$m_h^s = m_h^{*,h}\left(m_{h+1}^s, ..., m_T^s\right).$$

This gives us, by Proposition 1, the following recursive characterization of the realized MPCs of the sophisticated agent.

Proposition 2. *The realized MPCs of the sophisticated agent are given as follows:*

$$m_h^s = \begin{cases} 1 & \text{for } h = T \\ \frac{1}{1+\left(\rho_{h,h+1}\zeta_{h+1}^h\right)^{\frac{1}{\theta}}} & \text{for } h \leq T-1 \end{cases} \tag{6}$$

where ζ_t^h is the slope of the continue value function of self h in period t and is recursively defined as

$$\zeta_t^h = \begin{cases} 1 & \text{for } t = T \\ {m_t^s}^{1-\theta} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \left(1 - m_t^s\right)^{1-\theta} \zeta_{t+1}^h & \text{for } t \leq T-1 \end{cases}$$

Solving the model for the sophisticated agent through backward induction is equivalent to solving an extensive form game for the unique subgame-perfect Nash equilibrium where the agents of different ages are different players who can choose MPCs at each information node. The only way how an agent can influence through her chosen MPC the future consumption path in her favor is by restricting the budget, i.e., wealth level, of her future selfs. The MPC m_0^s—being a best response of the 0-old agent against the correctly anticipated MPCs of her future selfs—is therefore a function in m_t^s, for $t \geq h$. On the other hand, the MPCs of future agents do not depend on previously chosen MPCs. This is a consequence of the fact that optimal MPCs are independent of wealth levels for iso-elastic power period utility functions.

To interpret the marginal propensities to consume against the literature on hyperbolic discounting it is instructive to derive a *variant* of the *generalized Euler equation* (Harris and Laibson 2001), which as shown in the Appendix follows from the expressions for marginal propensities to consume given in Proposition 2 as

$$u_c(c_h^s) = \rho_{h,h+1} \left(m_{h+1}^s + \frac{\rho_{h,h+2}}{\rho_{h,h+1}\rho_{h+1,h+2}} \frac{\zeta_{h+2}^h}{\zeta_{h+2}^{h+1}} \left(1 - m_{h+1}^s\right) \right) u_c(c_{h+1}^s) \tag{7}$$

Equation (7) is the deterministic model analogue to the "generalized Euler equation with adjustment factor" we derived in a model with idiosyncratic productivity risk in Groneck, Ludwig and Zimper (2016). It reflects two effects on the consumption growth rate from dynamically inconsistent preferences. The first is through term

$$\frac{\rho_{h,h+2}}{\rho_{h,h+1}\rho_{h+1,h+2}} \neq 1 \text{ (in general)},$$

which in the familiar quasi-hyperbolic time discounting model is equal to the inverse of the short-run discount factor. The second is through the ratio

$$\frac{\zeta_{h+2}^h}{\zeta_{h+2}^{h+1}} \neq 1 \text{ (in general)},$$

which captures the difference in the marginal valuation of wealth in period $h+2$ from the perspective of the sophisticated agent h and her next period counterpart $h+1$. In the quasi-hyperbolic time discounting model $\frac{\zeta_{h+2}^{h}}{\zeta_{h+2}^{h+1}} = 1$, because continuation value functions from period $h+2$ onwards are the same for sophisticated agents h and $h+1$ in that model.

2.2.2 The Naive Agent

Definition 2. *We speak of a "naive agent" if and only if this agent assumes at every age h that her optimal consumption plan from the perspective of age h is also optimal from the perspective of all her future selfs $t > h$.*

In contrast to the sophisticated agent, the h-old naive agent bases her savings decision on a—possibly incorrect—assumption about her future behavior. Put differently, the naive agent completely ignores the possibility that her future selfs might have strict incentives to deviate from her optimal consumption path. Expressed in terms of the optimal MPCs of Proposition 1, the h-old naive agent's *planned* MPCs for $t \geq h$ are characterized as

$$m_t^{n,h} = m_t^{*,h}\left(m_{t+1}^{n,h}, ..., m_T^{n,h}\right). \tag{8}$$

Let $m_h^n = m_h^{n,h}$ be the realized MPCs of the h-old naive agent, which are stated in the next

Proposition 3. *The realized MPCs of the naive agent are given as follows:*

(i) *Recursive characterization:*

$$m_h^n = \begin{cases} 1 & \text{for } h = T \\ \frac{1}{1+\left(\sum_{t=h+1}^{T} \rho_{h,t}\left(m_t^{n,h} \prod_{j=h+1}^{t-1}\left(1-m_j^{n,h}\right)\right)^{1-\theta}\right)^{\frac{1}{\theta}}} & \text{for } h \leq T-1 \end{cases}$$

with planned MPCs

$$m_t^{n,h} = \begin{cases} 1 & \text{for } t = T \\ \frac{1}{1+\frac{\rho_{t,t+1}^{\frac{1}{\theta}}}{m_{t+1}^{n,h}}} = \frac{1}{1+\sum_{k=t+1}^{T}\left(\frac{\rho_{h,k}}{\rho_{h,t}}\right)^{\frac{1}{\theta}}} & \text{for } t \leq T-1 \end{cases}$$

(ii) *Closed form:*

$$m_h^n = \frac{1}{1+\sum_{t=h+1}^{T}\left(\rho_{h,t}\right)^{\frac{1}{\theta}}} \text{ for } h \leq T-1.$$

3 Dynamic Consistency versus Inconsistency

There exists a large behavioral and decision-theoretic literature which convincingly argues that human decision making is typically prone to violations of dynamic consistency. Dynamic inconsistencies arise, for example, within the following three modeling classes: (i) deterministic models with a presence bias induced by hyperbolic or quasi-hyperbolic time-discounting (Laibson 1997; 1998; O'Donoghue and Rabin 1999); (ii) non-deterministic models with expected utility maximizing agents who violate Bayes' rule[8]; (iii) non-deterministic models with Choquet expected utility decision makers (Schmeidler 1989; Gilboa 1987) or/and Prospect theory decision makers (Tversky and Kahneman 1992; Wakker and Tversky 1993; Wakker 2010) who form conditional non-additive beliefs that may or may not be updated in accordance with some Bayesian update rule (Gilboa and Schmeidler 1993; Eichberger, Grant, and Kelsey 2007; 2012). Within the context of life-cycle models with time-discounting and survival uncertainty, our model with arbitrary effective discount factors contains these modeling classes as special cases.

We formally define dynamic consistency versus dynamic inconsistency of the life-cycle model in terms of possible discrepancies between the planned and the realized MPCs of the naive agent. It will be analytical insightful to define these concepts with respect to the agent's age.

Definition 3.

(i) *We say that the model is "dynamically consistent at age h" if and only if*

$$m_t^{n,h} = m_t^n \text{ for all } t \geq h+1.$$

(ii) *Conversely, we say that the model is "dynamically inconsistent at age h" if and only if*

$$m_t^{n,h} \neq m_t^n \text{ for some } t \geq h+1.$$

The model is always dynamically consistent at ages $h \in \{T, T-1\}$. For $h \leq T-2$ we obtain, by Proposition 3, the following equivalent characterization of dynamic consistency in terms of discount factors.

[8]The economic literature which considers violations of Bayesian updating includes Rabin and Schrag (1999); Rabin (2002); Epstein (2006); Epstein, Noor, and Sandroni (2008); Mullainathan, Schwartzstein, and Shleifer (2008); Gennaioli and Shleifer (2010); Ortoleva (2012). Bayesian updating of additive probability measures is, through the *law of iterated expectations*, equivalent to dynamic consistency of expected utility preferences (cf., e.g., Epstein and Le Breton 1993; Epstein and Schneider 2003; Ghirardato 2002; Siniscalchi 2011).

Proposition 4. *The life-cycle model is dynamically consistent at age* $h \in \{0, ..., T-2\}$ *if and only if, for all* $t \in \{h+1, T-1\}$,

$$\sum_{k=t+1}^{T} \left(\frac{\rho_{h,k}}{\rho_{h,t}} \right)^{\frac{1}{\theta}} = \sum_{k=t+1}^{T} \left(\rho_{t,k} \right)^{\frac{1}{\theta}}. \tag{9}$$

Equation (9) is for every t generically violated over the space of all discount factors so that our model is, for almost all values of discount factors, dynamically inconsistent at any age $h \leq T-2$.

Example 1. To give an illustrative example, let $T = 3$ and observe that dynamic consistency at age $h = 0$ is characterized through the following two equations:

$$m_2^{n,h} = m_2^n \qquad \Leftrightarrow \qquad \frac{\rho_{0,3}}{\rho_{0,2}} = \rho_{2,3}$$

and

$$m_1^{n,h} = m_1^n \qquad \Leftrightarrow \qquad \left(\frac{\rho_{0,2}}{\rho_{0,1}} \right)^{\frac{1}{\theta}} + \left(\frac{\rho_{0,3}}{\rho_{0,1}} \right)^{\frac{1}{\theta}} = \left(\rho_{1,2} \right)^{\frac{1}{\theta}} + \left(\rho_{1,3} \right)^{\frac{1}{\theta}}.$$

Whenever we find some discount factors that satisfy both equations, a small perturbation of factors would break down equality. That is, dynamic consistency is non-generic at $h = T-3$ because it breaks down for the perturbed values of discount factors in any open interval—with strictly positive Lebesgue measure—around the original values.□

The standard way to ensure that (9) holds, and thereby dynamic consistency of the model at age h, is to impose the following condition *standard discounting (SDC)* on discount factors:

Definition 4. *We say that the discount factors* satisfy condition SDC at age h *if and only if*

$$\frac{\rho_{h,t+1}}{\rho_{h,t}} = \rho_{t,t+1} \text{ for all } t \in \{h+1, T-1\}, \tag{10}$$

which is mathematically equivalent to

$$\frac{\rho_{h,k}}{\rho_{h,t}} = \rho_{t,k} \text{ for all } t \in \{h+1, T-1\} \text{ and all } k > t. \tag{11}$$

Proposition 5. *Suppose that the discount factors satisfy condition SDC at age h. Then the following holds:*

(i) *The model is dynamically consistent at age h.*

(ii) *The discount factors also satisfy condition SDC at all ages $h' > h$.*

(iii) *By (i) and (ii), the model is dynamically consistent at all ages $h' \geq h$.*

To put Proposition 5 into context, it is important to notice two aspects. First, Condition SDC at h is sufficient but not necessary for ensuring dynamic consistency at h. Second, if Condition SDC is violated at h although the model is dynamically consistent at h, we may encounter situations where the model is dynamically inconsistent at some age $h' > h$. Both possibilities are illustrated by the following example.

Example 1 revisited. Suppose that the discount factors violate Condition SDC (10) but satisfy

$$\frac{\rho_{0,2}}{\rho_{0,1}} = \rho_{1,3}\ , \ \frac{\rho_{0,3}}{\rho_{0,1}} = \rho_{1,2}, \ \frac{\rho_{0,3}}{\rho_{0,2}} = \rho_{2,3},$$

implying

$$\frac{\rho_{1,3}}{\rho_{1,2}} = \frac{\rho_{0,2}}{\rho_{0,3}} = \frac{1}{\rho_{2,3}}. \tag{12}$$

The model is dynamically consistent at 0 because of

$$m_2^{n,h} = m_2^n \qquad \Leftrightarrow \qquad \frac{\rho_{0,3}}{\rho_{0,2}} = \rho_{2,3}$$

and

$$m_1^{n,h} = m_1^n \qquad \Leftrightarrow \qquad \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} + \left(\frac{\rho_{0,3}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} \quad = \quad \left(\rho_{1,3}\right)^{\frac{1}{\theta}} + \left(\rho_{1,2}\right)^{\frac{1}{\theta}}.$$

However, since dynamic consistency at age 1 requires $\rho_{2,3} = \frac{\rho_{1,3}}{\rho_{1,2}}$ the model is, by (12), dynamically inconsistent at age 1 unless $\rho_{2,3} = 1$.□

While condition SDC at h is satisfied for all $h \geq 0$ for exponential time-discounting combined with Bayesian updating of additive survival beliefs, it does typically not hold for (i) (quasi-)hyperbolic time-discounting models, (ii) for models with additive survival beliefs that are not updated in accordance with Bayes' rule, and (iii) for models with non-additive survival beliefs that are not updated in accordance with the optimistic Bayesian rule (Gilboa and Schmeidler 1993).[9]

[9] For detailed formal arguments we refer the reader to the working paper version of this paper (Groneck, Ludwig, and Zimper 2021).

4 Who Saves a Greater Fraction of Their Wealth: The Naive or the Sophisticated Agent?

4.1 Point of Departure

We will see that the sophisticated and the naive agent's savings behavior will coincide at all ages if the life-cycle model is dynamically consistent at all ages (cf. Corollary 3 below). Of course, this finding is not surprising. Quite surprising, however, is the following relationship: Even if the life-cycle model is dynamically inconsistent, both types of agents exhibit the same savings behavior whenever the period-utility function is of the logarithmic form. This remarkable finding goes back to the seminal analysis in Pollak (1968).

Theorem 0 (Pollak 1968). *For all (arbitrary) specifications of the effective discount factors we have at every age h:*

$$\theta = 1 \textit{ implies } m_h^n = m_h^s.$$

It is straightforward to verify Pollak's Theorem directly by setting $\theta = 1$ in the MPCs of Propositions 2 and 3 to obtain

$$m_h^s = m_h^n = \begin{cases} 1 & \text{for } h = T \\ \frac{1}{1+\sum_{t=h+1}^{T} \rho_{h,t}} & \text{for } h \leq T-1. \end{cases}$$

4.2 Main Result

For general $\theta \neq 1$ it follows also from Propositions 2 and 3 that the MPCs of the T- and $T-1$-old agents coincide for the naive and sophisticated type such that

$$\begin{aligned} m_T^n &= m_T^s = 1, \\ m_{T-1}^n &= m_{T-1}^s = \frac{1}{1+\left(\rho_{T-1,T}\right)^{\frac{1}{\theta}}}. \end{aligned}$$

For any ages $h \leq T-2$, however, it is no longer obvious how the sophisticated and naive agent's savings behavior will compare whenever $\theta \neq 1$. Our main result extends Pollak's analysis to the whole class of iso-elastic power utility functions, i.e., to all concavity parameter values $\theta \neq 1$.

Theorem 1. *For all (arbitrary) specifications of the effective discount factors we have at every age $h \le T-2$:*

(i) $\theta < 1$ *implies* $m_h^n \le m_h^s$*;*

(ii) $\theta > 1$ *implies* $m_h^n \ge m_h^s$*.*

The proof of Theorem 1 is based on the following

Lemma 1. *Let* $h \le T-2$*.*

(i) $\theta < 1$ *implies* $m_h^n < m_h^s$ *if and only if* $m_t^{n,h} \neq m_t^s$ *for some* $t \ge h+1$*.*

(ii) $\theta > 1$ *implies* $m_h^n > m_h^s$ *if and only if* $m_t^{n,h} \neq m_t^s$ *for some* $t \ge h+1$*.*

(iii) $\theta \neq 1$ *and* $m_h^n = m_h^s$ *if and only if* $m_t^{n,h} = m_t^s$ *for all* $t \ge h+1$*.*

Our proof of Lemma 1 is based on an application of Jensen's inequality. We here sketch the core of the proof for a three period model ($T = 2$).[10] By comparing the analytical expressions of the marginal propensities to consume of the sophistic and the naive agent at age $h = T-2 = 0$, we show that $m_0^n \le m_0^s$ is equivalent to

$$\left(\frac{m_1^{n,0}}{m_1}\right)^{\theta} m_1 + \left(\frac{1-m_1^{n,0}}{1-m_1}\right)^{\theta} (1-m_1) \le 1,$$

where $m_1^{n,0}$ is the period 0 plan for the period 1 marginal propensity to consume of the naive agent. This condition is a weighted average of two function values around one—the function is evaluated at $\frac{m_1^{n,0}}{m_1}$ and $\frac{1-m_1^{n,0}}{1-m_1}$—on a convex function for $\theta > 1$ and on a concave function for $\theta < 1$. Therefore, the term is less than one so that $m_0^s < m_0^n$ if the function is concave ($\theta < 1$) and above one so that $m_0^s > m_0^n$ if the function is convex ($\theta > 1$).

Let us use the characterizations of Lemma 1 to identify further conditions such that the weak inequalities in Theorem 1 either become strict or hold with equality. At first, observe that $m_{T-1}^s = m_{T-1}^n$ implies

$$m_{T-1}^{n,h} \neq m_{T-1}^s \quad \Leftrightarrow \quad m_{T-1}^{n,h} \neq m_{T-1}^n \quad \Leftrightarrow \quad \frac{\rho_{h,T}}{\rho_{h,T-1}} \neq \rho_{T-1,T}, \tag{13}$$

which gives us by Lemma 1 the following (easy-to-check) sufficiency condition for strict inequalities.

[10] The proof for $T > 2$ is a straightforward extension of the main idea and is based on backward induction.

Corollary 1. *Let $h \leq T-2$. Whenever the discount factors satisfy inequality (13), we have:*

(i) *$\theta < 1$ implies $m_h^n < m_h^s$;*

(ii) *$\theta > 1$ implies $m_h^n > m_h^s$.*

Because inequality (13) holds generically, we can combine these strict inequalities with Theorem 0 by Pollak (1968) to obtain the following statement.

Corollary 2. *Let $h \leq T-2$. We have generically that*

$$m_h^n < (>)\, m_h^s \text{ if and only if } \theta < (>)\, 1.$$

Recall that we have defined dynamic consistency at age h as

$$m_t^{n,h} = m_t^n \text{ for all } t \geq h+1.$$

Proposition 6. *Let $h \leq T-2$ and $\theta \neq 1$. We have $m_h^n = m_h^s$ whenever the model is dynamically consistent at all ages $t \geq h$.*

Combining Proposition 5 with Proposition 6 gives us the following corollary.

Corollary 3. *Suppose that the discount factors satisfy Condition SDC (10) at age h. Then $m_{h'}^n = m_{h'}^s$ for all ages $h' \geq h$.*

Next, recall our definition of dynamic inconsistency at age h:

$$m_t^{n,h} \neq m_t^n \text{ for some } t \geq h+1.$$

Proposition 7. *If the model is dynamically inconsistent at age $h \leq T-2$, we have:*

(i) *$\theta < 1$ implies $m_t^n < m_t^s$ for some $t \geq h$;*

(ii) *$\theta > 1$ implies $m_t^n > m_t^s$ for some $t \geq h$.*

If the model is dynamically consistent at age h, we have, by Proposition 6, that $m_h^n = m_h^s$ provided the model satisfies the additional requirement that it is also dynamically consistent at all ages $h' \geq h+1$. Whenever the discount factors satisfy Condition SDC (10) at age h, this additional requirement is automatically satisfied (cf. Proposition 5 and Corollary 3). The following result clarifies that we cannot drop this additional requirement whenever Condition SDC (10) is violated; that is, dynamic consistency at age h alone is, in general, not sufficient to guarantee $m_h^n = m_h^s$.

Proposition 8. *Suppose that the model is dynamically consistent at age h but dynamically inconsistent at some age $h' \geq h+1$. Then we have:*

(i) $\theta < 1$ *implies* $m_h^n < m_h^s$*;*

(ii) $\theta > 1$ *implies* $m_h^n > m_h^s$.

4.3 Interpretation: The three-period model

Why is our main result exclusively driven by the concavity parameter θ and not by properties of the effective discount process? To offer some intuition, consider the three-period model ($T = 2$) for which the MPCs of the sophisticated and the naive agent coincide in the second and third period, i.e., $m_t^s = m_t^n = m_t$ for $t = 1, 2$. As a consequence, we can comprehensively describe any departure from *standard discounting* in the three-period model through a single parameter $\epsilon = \frac{\rho_{0,2}}{\rho_{0,1}\rho_{1,2}} - 1 \neq 0$ only.

Denote by $d_0^i = \frac{1-m_0^i}{m_0^i}$ for $i \in \{n, s\}$ the ratio of the marginal propensity to save to the marginal propensity to consume for the two agent types. Next define the measure for the relative reaction between the sophisticated and the naive agent to *any* departure from standard discounting as the ratio

$$y \equiv \frac{\frac{\partial d_0^s}{\partial \epsilon}}{\frac{\partial d_0^n}{\partial \epsilon}}.$$

In Appendix B.8 we formally derive the relationship

$$y \equiv \frac{\frac{\partial d_0^s}{\partial \epsilon}}{\frac{\partial d_0^n}{\partial \epsilon}} = \left(\frac{1+\epsilon(1-m_1)}{1+\epsilon}\right)^{\frac{1}{\theta}-1} > 1 \text{ if and only if } \theta > 1, \epsilon \neq 0, \tag{14}$$

because of $\frac{1+\epsilon(1-m_1)}{1+\epsilon} < 1$. This local argument shows that only $\theta \geq 1$ versus $\theta \leq 1$ matters for the question whether the sophisticated or the naive type reacts stronger to any (small) deviation $\epsilon \neq 0$ from standard discounting whereby the direction of this deviation is irrelevant.

Now focus on the case $\theta \geq 1$. For this case we show that

$$\left.\frac{\partial y}{\partial \theta}\right|_{\theta^*} > 0 \text{ if and only if } \theta^* < \bar{\theta} \tag{15}$$

for the threshold value $\bar{\theta} > 1$ which is uniquely determined as the root of a non-linear equation. For parameter values θ such that $1 \leq \theta \leq \bar{\theta}$ we thus obtain that the relative size of the sophisticated agent's reaction to any (small) deviation $\epsilon \neq 0$ from standard discounting increases whereas it decreases for $\theta \geq \bar{\theta}$.[11] This argument illustrates that the role of the concavity parameter θ, i.e., the $IES = \frac{1}{\theta}$, for the *inter-personal* comparison between the naive and the sophisticated agent is more complex than the role it plays as an *inter-temporal* smoothing parameter in models with standard discounting.

4.4 Quantitative Relevance

We illustrate the quantitative implications of our findings in Figure 1 for a stylized calibration of a standard hyperbolic time discounting model assuming estimates of θ regarded as plausible in the literature, cf. Bansal and Yaron (2004). For a typical "finance" calibration with $1/\theta = 2$ the percent differences are very large, and also for a standard "macro" calibration where $1/\theta = 0.5$, the implied percent differences in consumption range from non-negligible -1% at age $h = 0$ to 2.3% at age T.[12]

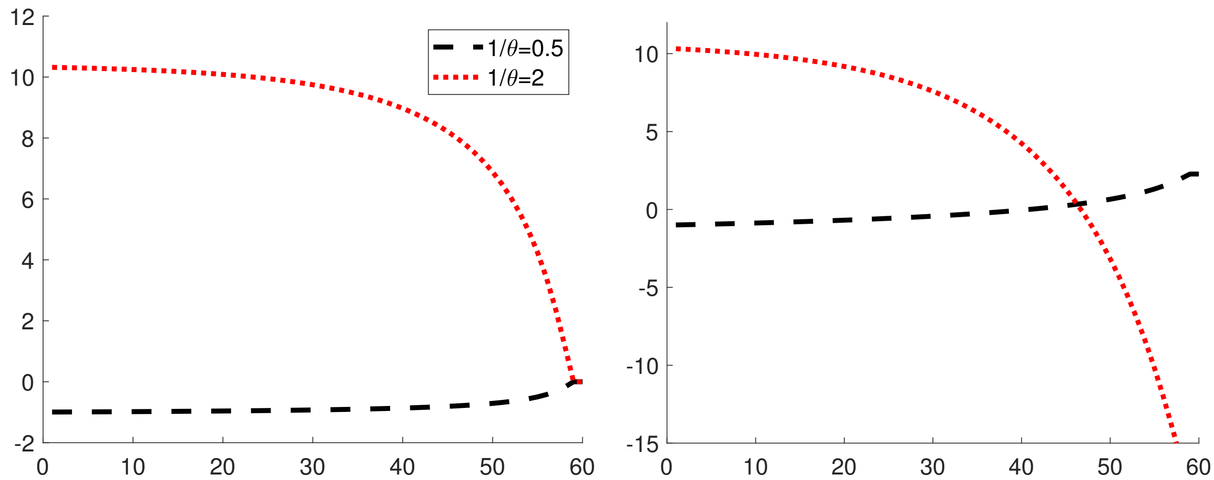
5 Extension

We extend our main result to a life-cycle model with random returns, a portfolio choice and Epstein-Zin-Weil (EZW) preferences (Epstein and Zin 1989; Epstein and Zin 1991; Weil 1989) with arbitrary discount factors. Our extension builds on fundamental insights of the seminal work by Merton (1969) and Samuelson (1969) that with homothetic preferences and serially uncorrelated returns the policy functions for consumption are linear in total wealth. The according expressions for the MPCs of the naive and the sophisticated agent are therefore analogous to those in our baseline model. It is then straightforward to establish that the backward recursive proof of Theorem 1 readily extends to this setup.

[11] Whenever $\theta \geq \bar{\theta}$ the differences between the sophisticated relative to the naive agent's saving reaction decreases because in the limit $\theta \to \infty$ the savings behavior of both types must coincide again.

[12] For the chosen calibration of the model, details of which are described in Appendix C, the value of $\bar{\theta}$, cf. Section 4.3, evaluated at age 0 is $\bar{\theta}_0 = 1.9344$.

Figure 1: Percent Differences in MPCs and Consumption Levels



(a) Percent Differences in MPCs

(b) Percent Differences in Consumption

Notes: Percentage differences between sophisticated and naive hyperbolic discounting agents' marginal propensity to consume $\left(\frac{m_h^s}{m_h^n} - 1\right) \cdot 100\%$ (Panel a) and level of consumption $\left(\frac{c_h^s}{c_h^n} - 1\right) \cdot 100\%$ (Panel b). See Appendix C for the calibration of the model.

5.1 Epstein-Zin-Weil Preferences with Arbitrary Discount Factors

The familiar Epstein and Zin (1989, 1991) aggregator with discount factor $\frac{\rho_{h,t+1}}{\rho_{h,t}}$ can be written as[13]

$$U_t^h = \frac{c_t^{1-\theta}}{1-\theta} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \frac{1}{1-\theta} \left(\mathbb{E}\left[\left((1-\theta) U_{t+1}^h \right)^{\frac{1-\sigma}{1-\theta}} \right] \right)^{\frac{1-\theta}{1-\sigma}} \quad \text{for all } t \geq h, \tag{16}$$

where parameter $\sigma > 0$ is a measure of risk-aversion whereas parameter θ is a measure of resistance to inter-temporal substitution. For the parametrization $\sigma = \theta$ the standard additively time-separable case with CRRA per period utility function (3) is nested as a special case.

We interpret the discount function $\rho_{h,t}$ as pure time discounting, thus $\rho_{h,t} = \beta_{h,t}$. As it is, there exists an ongoing discussion in the literature regarding interpretational issues of EZW life-cycle models with survival risks rather than with pure time-discounting only, which we sidestep here as this discussion is beyond the scope of the present paper (cf. Hugonnier et al.

[13] Our representation is an often used monotone transformation $U_t^h = \frac{1}{1-\theta} {V_t^h}^{1-\theta}$ of the standard textbook formulation $V_t^h = \left(c_t^{1-\theta} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \left(\mathbb{E}\left[{V_{t+1}^h}^{1-\sigma} \right] \right)^{\frac{1-\theta}{1-\sigma}} \right)^{\frac{1}{1-\theta}}$, which directly nests the additively separable case where $\sigma = \theta$.

2013; Córdoba and Ripoll 2017; Bommier et al. 2020; Bommier et al. 2021).[14] Also for the nested additively separable case where $\sigma = \theta$ we abstract from the possibility that dynamic inconsistency may be induced by non-additive subjective survival beliefs $\nu_{h,t}$, but again solely assume it through the pure time discounting function $\beta_{h,t}$.[15]

5.2 Random Return Process with Portfolio Choice

Let R_t be an independently (over time) distributed risky return factor governed by the additive probability measure π, where R_t takes weakly positive values π-almost surely.[16] As in Section 4.4 we think of an economy that is populated by naive and sophisticated agents so that R_t is an aggregate return process. Additionally, let R^f be a risk-free return factor such that $R^f < \mathbb{E}[R_t] = \int R_t d\pi$. The household chooses in period t to invest share α_t in stocks with next period risky return R_{t+1} and $1 - \alpha_t$ in bonds with risk-free return R^f. The stochastic portfolio return on the beginning of period t financial wealth holdings is accordingly $R^p_t = R^f + \alpha_{t-1}\left(R_t - R^f\right)$. Additionally, let e_t be a possibly time varying deterministic endowment income stream of the agent. In our Supplementary Appendix we show that the budget constraint in terms of total wealth $w_t = x_t + h_t$—where h_t is human wealth (the present value of labor income) and $x_t = x_t = a_t R^p_t(\alpha_{t-1}) + e_t$ is cash-on-hand—is given by

$$w_{t+1} = (w_t - c_t)\, R^p_{t+1}(\hat{\alpha}_t) \tag{17}$$

where

$$\hat{\alpha}_t = \alpha_t \frac{x_t - c_t}{w_t - c_t}. \tag{18}$$

Observe that a nested model variant is one without labor income and risky returns (with or without a portfolio choice), where households decumulate a given initial financial wealth endowment over the life-cycle. Furthermore, an alternative model giving rise to the same mathematical properties is one with risky labor income generated by risky returns to human capital h_t and

[14] In a nutshell, this discussion concerns the question whether homothetic EZW preferences that explicitly incorporate the utility of possible death can be consistent with the natural assumption that 'life is better than death' for parameter values $\sigma \neq \theta$, $\sigma \geq 1, \theta \geq 1$.

[15] Otherwise we would have to be explicit how to define the resolution of uncertainty with regard to the joint process of return and survival uncertainty.

[16] These are standard assumptions on the return process in the portfolio choice literature.

a linear human capital production function taking monetary human capital investments i_t as inputs, cf. Krebs (2003), and thus our results apply to a larger class of models.[17]

5.3 Main Results from Extension

The main results from this extension can be summarized in the following two corollaries, proven in the Supplementary Appendix.

Corollary 4. *Lemma 1 and thus Theorem 1 extend to the dynamically inconsistent EZW life-cycle model with arbitrary discount factors.*

Corollary 5. *Theorem 1 extends to portfolio shares in the dynamically inconsistent EZW life-cycle model with arbitrary discount factors such that:*

(i) $\theta < 1$ *implies* $\alpha_h^n \leq \alpha_h^s$ *for all* h *;*

(ii) $\theta > 1$ *implies* $\alpha_h^n \geq \alpha_h^s$ *for all* h *.*

6 Concluding Remarks

Pollak (1968) shows that—irrespective of the specification of discount factors—the sophisticated agent and her naive counterpart exhibit the same savings behavior whenever their period utility function is logarithmic. We extend Pollak's analysis to the class of all iso-elastic power utility functions by showing that the sophisticated agent saves in every period a greater fraction of her wealth than her naive counterpart if and only if the resistance to inter-temporal substitution is larger than one. As a generalization of the additively time-separable life-cycle model (1) we show that exactly the same relationship holds in a model with return risk, a portfolio choice and Epstein-Zin-Weil (EZW) preferences. We expect our findings to provide useful guidance for the interpretation of results in quantitative work where closed form solutions no longer arise but where the interpretation on the relative consumption responses of the two types of agents in θ may still hold approximately.

[17] This encompasses models with self imposed borrowing constraints if there exists a strict positive probability for the returns on human capital and on financial wealth to be zero.

Appendix

A Proof of Theorem 1

Our proof of Theorem 1 is based on the recursive presentations of the marginal propensities to consume of the sophisticated and the naive agent. The different implications for the cases $\theta < 1$ versus $\theta > 1$ result from a simple application of Jensen's inequality to strictly concave and strictly convex functions, respectively. Because the proof of Theorem 1 will be implied by the proof of Lemma 1, we prove, at first, Lemma 1.

Proof of Lemma 1. Part (i): We show for $h \in \{0, \dots, T-2\}$:

(i) $\theta < 1$ *implies* $m_h^n = m_h^s$ *if* $m_t^{n,h} = m_t^s$ *for all* $t \geq h+1$.

(ii) $\theta < 1$ *implies* $m_h^n < m_h^s$ *if* $m_t^{n,h} \neq m_t^s$ *for some* $t \geq h+1$.

Recall from (6) and (28) the following expressions for MPCs

$$m_h^s = \frac{1}{1 + \left(\rho_{h,h+1}\zeta_{h+1}^h\right)^{\frac{1}{\theta}}}$$

where

$$\zeta_t^h = m_t^{s^{1-\theta}} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \left(1 - m_t^s\right)^{1-\theta} \zeta_{t+1}^h \tag{19}$$

as well as

$$m_t^{n,h} = \frac{1}{1 + \left(\frac{\rho_{h,t+1}}{\rho_{h,t}}\right)^{\frac{1}{\theta}} m_{t+1}^{n,h^{-1}}}. \tag{20}$$

Using these expressions gives us at age $t = h$

$$\begin{aligned} m_h^n &\leq m_h^s \\ &\Leftrightarrow \\ \left(\rho_{h,h+1}\zeta_{h+1}^h\right)^{\frac{1}{\theta}} &\leq \left(\frac{\rho_{h,h+1}}{\rho_{h,h}}\right)^{\frac{1}{\theta}} m_{h+1}^{n,h^{-1}} \\ &\Leftrightarrow \\ m_{h+1}^{n,h^{\theta}}\zeta_{h+1}^h &\leq 1. \end{aligned}$$

Next, we appropriately transform ζ_t^h. To this purpose, notice from (20) that

$$\frac{\rho_{h,t+1}}{\rho_{h,t}} = \left(\frac{1 - m_t^{n,h}}{m_t^{n,h}}\right)^{\theta} m_{t+1}^{n,h^{\theta}}.$$

Using this in (19) we get recursively for $t = T-2, \dots, h+1$

$$\zeta_t^h = m_t^{s^{1-\theta}} + \left(\frac{1-m_t^{n,h}}{m_t^{n,h}}\right)^\theta (1-m_t^s)^{1-\theta} m_{t+1}^{n,h^\theta} \zeta_{t+1}^h$$

$$\Leftrightarrow \qquad m_t^{n,h^\theta} \zeta_t^h = \left(\frac{m_t^{n,h}}{m_t^s}\right)^\theta m_t^s + \left(\frac{1-m_t^{n,h}}{1-m_t^s}\right)^\theta (1-m_t^s) m_{t+1}^{n,h^\theta} \zeta_{t+1}^h. \tag{21}$$

The remainder of the proof proceeds by backward induction on (21) over $t = T-1, \dots, h+1$.

Claims: Firstly, we claim that, for all $t \in \{h+1, \dots, T-1\}$, $\theta < 1$ implies

$$m_t^{n,h^\theta} \zeta_t^h = 1 \tag{22}$$

if $m_t^{n,h} = m_t^s$ for all $t \geq h+1$.

Secondly, we claim that, for all $t \in \{h+1, \dots, T-1\}$, $\theta < 1$ implies

$$m_t^{n,h^\theta} \zeta_t^h < 1 \tag{23}$$

if $m_t^{n,h} \neq m_t^s$ for some $t \geq h+1$.

Base Case: Recall that $m_T^n = m_T^{n,h} = m_T^s = 1$. In period $t = T-1$ we have

$$m_{T-1}^{n,h^\theta} \zeta_{T-1}^h = \left(\frac{m_{T-1}^{n,h}}{m_{T-1}^s}\right)^\theta m_{T-1}^s + \left(\frac{1-m_{T-1}^{n,h}}{1-m_{T-1}^s}\right)^\theta \left(1-m_{T-1}^s\right).$$

Suppose, at first, that $m_{T-1}^{n,h} = m_{T-1}^s$. Then our first claim (22) is trivially satisfied for $t = T-1$ because of

$$m_t^{n,h^\theta} \zeta_t^h = 1$$

irrespective of the value of θ.

Suppose now that $m_{T-1}^{n,h} \neq m_{T-1}^s$, implying

$$\frac{m_{T-1}^{n,h}}{m_{T-1}^s} \neq \frac{1-m_{T-1}^{n,h}}{1-m_{T-1}^s}.$$

By the strict version of Jensen's inequality, we obtain for $\theta < 1$

$$\begin{aligned} m_{T-1}^{n,h^\theta} \zeta_{T-1}^h &= \left(\frac{m_{T-1}^{n,h}}{m_{T-1}^s}\right)^\theta m_{T-1}^s + \left(\frac{1-m_{T-1}^{n,h}}{1-m_{T-1}^s}\right)^\theta \left(1-m_{T-1}^s\right) \\ &< \left(\left(\frac{m_{T-1}^{n,h}}{m_{T-1}^s}\right) m_{T-1}^s + \left(\frac{1-m_{T-1}^{n,h}}{1-m_{T-1}^s}\right) \left(1-m_{T-1}^s\right)\right)^\theta \\ &= 1 \end{aligned}$$

because x^θ is strictly concave for $\theta < 1$. Consequently, our second claim (23) is satisfied for $t = T - 1$.

Backward Induction Step: Suppose that the first claim (22) has been proved for period $i+1$. That is, we have shown that $\theta < 1$ implies

$$m_{i+1}^{n,h^\theta} \zeta_{i+1}^h = 1 \tag{24}$$

if $m_t^{n,h} = m_t^s$ for all $t \geq i+1$. Rewrite (21) as

$$m_i^{n,h^\theta} \zeta_i^h = \underbrace{\left(\frac{m_i^{n,h}}{m_i^s}\right)^\theta m_i^s + \left(\frac{1-m_i^{n,h}}{1-m_i^s}\right)^\theta (1-m_i^s)}_{=\Lambda(m_i^{n,h}, m_i^s)} m_{i+1}^{n,h^\theta} \zeta_{i+1}^h.$$

By the same reasoning as in the base case, we have that $\theta < 1$ implies

$$\Lambda(m_i^{n,h}, m_i^s) \leq 1 \tag{25}$$

whereby this inequality is strict if and only if $m_i^{h,n} \neq m_i^s$. Since

$$x + y \leq 1 \text{ and } b \leq 1 \text{ implies } x + by \leq 1,$$

(24) together with (25) gives us the desired result that $\theta < 1$ implies

$$m_i^{n,h^\theta} \zeta_i^h = 1 \tag{26}$$

if $m_i^{h,n} = m_i^s$ whereas we have

$$m_i^{n,h^\theta} \zeta_i^h < 1$$

if $m_i^{h,n} \neq m_i^s$.

Next suppose that we have proved the second claim (23) for period $i+1$. That is, we have shown that $\theta < 1$ implies

$$m_{i+1}^{n,h^\theta} \zeta_{i+1}^h < 1$$

if $m_t^{n,h} \neq m_t^s$ for some $t \geq i+1$. Because of (25), we must have that

$$m_i^{n,h^\theta} \zeta_i^h < 1.$$

Combining both cases proves Part (i) of Lemma 1.□

Proof of Lemma 1. Part (ii): We show for $h \in \{0, \ldots, T-2\}$:

(i) $\theta > 1$ *implies* $m_h^n = m_h^s$ *if* $m_t^{n,h} = m_t^s$ *for all* $t \geq h+1$.

(ii) $\theta > 1$ *implies* $m_h^n < m_h^s$ *if* $m_t^{n,h} \neq m_t^s$ *for some* $t \geq h+1$.

The proof proceeds exactly as the proof of Part (i) of Lemma 1 whereby we prove the following two claims:

Firstly, for all $t \in \{h+1, \ldots, T-1\}$, $\theta > 1$ implies

$$m_t^{n,h^\theta} \zeta_t^h = 1$$

if $m_t^{n,h} = m_t^s$ for all $t \geq h+1$.

Secondly, for all $t \in \{h+1, \ldots, T-1\}$, $\theta > 1$ implies

$$m_t^{n,h^\theta} \zeta_t^h > 1 \tag{27}$$

if $m_t^{n,h} \neq m_t^s$ for some $t \geq h+1$.

The only difference to the proof of Part (i) is the reversed strict inequality in claim (27) which follows, by the strict version of Jensen's inequality, by strict convexity of x^θ for $\theta > 1$.□□

Proof of Theorem 1. To prove Part (i), we have to show that $\theta < 1$ implies $m_h^n \leq m_h^s$. Recall from the proof of Lemma 1(i) that

$$m_t^{n,h^\theta} \zeta_t^h \leq 1 \text{ for all } t \in \{T-2, \ldots, h+1\} \text{ implies } m_h^n \leq m_h^s.$$

Moreover, the proof of Lemma 1(i) had established that $\theta < 1$ implies either $m_t^{n,h^\theta} \zeta_t^h = 1$ or $m_t^{n,h^\theta} \zeta_t^h < 1$ for all $t \in \{T-2, \ldots, h+1\}$. An analogous argument applies to Part (ii) of Theorem 1.□□

B Additional Proofs and Derivations

B.1 Proof of Proposition 1

For $h = T$, we trivially have as optimal consumption $c_T^{*,T} = w_T$ with optimal MPC $m_T^{*,T} = 1$. For $h < T$, the optimal period h consumption $c_h^{*,h}$ from the perspective of the h-old agent is pinned down by the following FOC:

$$\left.\frac{dU_h(c_h; m_{h+1}, ..., m_T, w_h)}{dc_h}\right|_{c_h = c_h^{*,h}} = 0$$

$$\Leftrightarrow$$

$$u'(c_h^{*,h}) = \sum_{t=h+1}^{T} \rho_{h,t} u' \left(\left(w_h - c_h^{*,h} \right) m_t \prod_{j=h+1}^{t-1} (1-m_j) \right) \left(m_t \prod_{j=h+1}^{t-1} (1-m_j) \right),$$

which becomes for the power period utility function

$$\left(c_h^{*,h} \right)^{-\theta} = \left(w_h - c_h^{*,h} \right)^{-\theta} \sum_{t=h+1}^{T} \rho_{h,t} \left(m_t \prod_{j=h+1}^{t-1} (1-m_j) \right)^{1-\theta}.$$

Solving for $c_h^{*,h}$ results in

$$c_h^{*,h} = m_h^{*,h} \left(m_{h+1}, ..., m_T \right) w_h$$

such that the optimal period h MPC for fixed period $h+1, ..., T$ MPCs is given as

$$m_h^{*,h} \left(m_{h+1}, ..., m_T \right) = \frac{1}{1 + \left(\sum_{t=h+1}^{T} \rho_{h,t} \left(m_t \prod_{j=h+1}^{t-1} (1-m_j) \right)^{1-\theta} \right)^{\frac{1}{\theta}}}.$$

More generally, by the envelope theorem, the optimal period $t \geq h$ consumption from the perspective of the h-old agent given fixed values of $m_{t+1}, ..., m_T$ and wealth w_t is pinned down by

$$\rho_{h,t} \left(c_t^{*,h} \right)^{-\theta} = \left(w_t - c_t^{*,h} \right)^{-\theta} \sum_{s=t+1}^{T} \rho_{h,s} \left(m_t \prod_{j=t+1}^{s-1} (1-m_j) \right)^{1-\theta}.$$

B.2 Derivation of Equation (7)

The consumption growth rate of the sophisticated agent is given as

$$\frac{c_{h+1}^s}{c_h^s} = \frac{m_{h+1}^s w_{h+1}}{m_h^s w_h} = \frac{1-m_h^s}{m_h^s} m_{h+1}^s.$$

Using the expression for m_h^s and ζ_{h+1}^h from Proposition 2 we obtain

$$\begin{aligned}
\frac{c_{h+1}^s}{c_h^s} &= \rho_{h,h+1}^{\frac{1}{\theta}} \zeta_{h+1}^{h\frac{1}{\theta}} \\
&= \rho_{h,h+1}^{\frac{1}{\theta}} \left(m_{h+1}^{s\,1-\theta} + \frac{\rho_{h,h+2}}{\rho_{h,h+1}} \left(1 - m_{h+1}^s \right)^{1-\theta} \zeta_{h+2}^h \right)^{\frac{1}{\theta}} m_{h+1}^s \\
&= \rho_{h,h+1}^{\frac{1}{\theta}} \left(m_{h+1}^s + \frac{\rho_{h,h+2}}{\rho_{h,h+1}} \left(1 - m_{h+1}^s \right) \left(\frac{1-m_{h+1}^s}{m_{h+1}^s} \right)^{-\theta} \zeta_{h+2}^h \right)^{\frac{1}{\theta}} \\
&= \rho_{h,h+1}^{\frac{1}{\theta}} \left(m_{h+1}^s + \frac{\rho_{h,h+2}}{\rho_{h,h+1}} \left(1 - m_{h+1}^s \right) \left(\left(\rho_{h+1,h+2} \zeta_{h+2}^{h+1} \right)^{\frac{1}{\theta}} \right)^{-\theta} \zeta_{h+2}^h \right)^{\frac{1}{\theta}} \\
&= \rho_{h,h+1}^{\frac{1}{\theta}} \left(m_{h+1}^s + \frac{\rho_{h,h+2}}{\rho_{h,h+1} \rho_{h+1,h+2}} \frac{\zeta_{h+2}^h}{\zeta_{h+2}^{h+1}} \left(1 - m_{h+1}^s \right) \right)^{\frac{1}{\theta}}.
\end{aligned}$$

Noting that $u_c(c) = c^{-\theta}$ then gives (7).

B.3 Proof of Proposition 3

Mathematically equivalently, the h-old naive agent's *planned* MPCs are pinned down by the following FOCs for all t such that $h \leq t < T$:

$$\begin{aligned}
\rho_{h,t}(m_t^{n,h} w_t)^{-\theta} &= \rho_{h,t+1}\left(m_{t+1}^{n,h} w_{t+1}\right)^{-\theta} \\
&\Leftrightarrow \\
\rho_{h,t}(m_t^{n,h} w_t)^{-\theta} &= \rho_{h,t+1}\left(m_{t+1}^{n,h}\left(w_t - m_t^{n,h} w_t\right)\right)^{-\theta} \\
&\Leftrightarrow \\
m_t^{n,h} &= \frac{1}{1+\left(\frac{\rho_{h,t+1}}{\rho_{h,t}}\right)^{\frac{1}{\theta}}\left(m_{t+1}^{n,h}\right)^{-1}}. \qquad (28)
\end{aligned}$$

Substituting

$$m_{t+1}^{n,h} = \frac{1}{1+\left(\frac{\rho_{h,t+2}}{\rho_{h,t+1}}\right)^{\frac{1}{\theta}}\left(m_{t+2}^{n,h}\right)^{-1}}$$

in (28) gives

$$m_t^{n,h} = \frac{1}{1+\left(\frac{\rho_{h,t+1}}{\rho_{h,t}}\right)^{\frac{1}{\theta}}+\left(\frac{\rho_{h,t+2}}{\rho_{h,t}}\right)^{\frac{1}{\theta}}\left(m_{t+2}^{n,h}\right)^{-1}}.$$

By repeating this argument until $m_T^{n,h} = 1$, we obtain the following closed form description of planned MPCs

$$m_t^{n,h} = \begin{cases} 1 & \text{for } t = T \\ \frac{1}{1+\sum_{k=t+1}^{T}\left(\frac{\rho_{h,k}}{\rho_{h,t}}\right)^{\frac{1}{\theta}}} & \text{for } t \leq T-1. \end{cases}$$

B.4 Proof of Proposition 5

Part (i) is obvious and part (iii) follows from (i) and (ii). It remains to prove part (ii). Suppose to the contrary that Condition SDC holds at age h but that there exists some $h' > h$ such that

$$\frac{\rho_{h',k}}{\rho_{h',t}} \neq \rho_{t,k} \qquad (29)$$

for some $t \in \{h'+1, T-1\}$ and some $k > t$. Note that (11) implies

$$\frac{\rho_{h,t}}{\rho_{h,h'}} = \rho_{h',t}\ ,\ \frac{\rho_{h,k}}{\rho_{h,h'}} = \rho_{h',k},\ \frac{\rho_{h,k}}{\rho_{h,t}} = \rho_{t,k}$$

and therefore

$$\frac{\rho_{h',k}}{\rho_{h',t}} = \frac{\frac{\rho_{h,k}}{\rho_{h,h'}}}{\frac{\rho_{h,t}}{\rho_{h,h'}}} = \frac{\rho_{h,k}}{\rho_{h,t}} = \rho_{t,k};$$

a contradiction to (29).

B.5 Proof of Proposition 6

For age $T-1$ we have trivially

$$m_T^{n,T-1} = m_T^n = m_T^s = 1 \tag{30}$$

so that by the if-part of Lemma 1(iii)

$$m_{T-1}^n = m_{T-1}^s. \tag{31}$$

The model is always dynamically consistent at ages T and $T-1$. Suppose now that the model is dynamically consistent at age $t = T-2$, we have, by definition,

$$m_t^{n,T-2} = m_t^n \text{ for } t \geq T-1.$$

This gives us by (30) and (31)

$$m_t^{n,T-2} = m_t^s \text{ for } t \geq T-1$$

so that by the if-part of Lemma 1(iii)

$$m_{T-2}^n = m_{T-2}^s. \tag{32}$$

By backward induction, we obtain the proposition for arbitrary $h \leq T-2$.

B.6 Proof of Proposition 7

Focus on $\theta < 1$ and suppose to the contrary that $m_t^n \leq m_t^s$ does not become strict for some $t \geq h$ but that

$$m_t^n = m_t^s \text{ for all } t \geq h. \tag{33}$$

By Lemma 1(iii), $m_h^n = m_h^s$ implies $m_t^{n,h} = m_t^s$ for all $t \geq h+1$. This gives us, by (33),

$$m_t^{n,h} = m_t^n \text{ for all } t \geq h+1,$$

which contradicts the assumption of dynamic inconsistency at age h.

B.7 Proof of Proposition 8

Focus on $\theta < 1$ and suppose to the contrary that $m_h^n = m_h^s$ instead of $m_h^n < m_h^s$. By Lemma 1(iii), we have that

$$m_t^{n,h} = m_t^s \text{ for all } t \geq h+1.$$

If the model is dynamically consistent at age h, we further have

$$m_t^{n,h} = m_t^n \text{ for all } t \geq h+1,$$

implying

$$m_t^n = m_t^s \text{ for all } t \geq h+1. \tag{34}$$

But if the model is dynamically inconsistent at age $h' \geq h+1$, we obtain, by Proposition 7(i), $m_t^n < m_t^s$ for some $t \geq h'$, a contradiction to (34).

B.8 Three-period model: Derivation of Equations (14) and (15)

Since $m_1 = \frac{1}{1+\rho_{1,2}^{\frac{1}{\theta}}}$ we obtain for the MPCs of the two agents

$$m_0^n = \frac{1}{1+\rho_{0,1}^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}\rho_{1,2}}\right)^{\frac{1}{\theta}}\rho_{1,2}^{\frac{1}{\theta}}\right)}$$

$$m_0^s = \frac{1}{1+\rho_{0,1}^{\frac{1}{\theta}}\left(\left(\frac{1}{1+\rho_{1,2}^{\frac{1}{\theta}}}\right)^{1-\theta}+\rho_{1,2}^{\frac{1}{\theta}}\left(\frac{1}{1+\rho_{1,2}^{\frac{1}{\theta}}}\right)^{1-\theta}\frac{\rho_{0,2}}{\rho_{0,1}\rho_{1,2}}\right)^{\frac{1}{\theta}}}.$$

Next, let $1+\epsilon = \frac{\rho_{0,2}}{\rho_{0,1}\rho_{1,2}}$ for $\epsilon \neq 0$ and define $d_0^i = \frac{1-m_1^i}{m_1^i}$ for $i \in \{n, s\}$ to get

$$\begin{aligned}
\frac{\partial d_0^n}{\partial \epsilon} &= \frac{1}{\theta}\rho_{0,1}^{\frac{1}{\theta}}\rho_{1,2}^{\frac{1}{\theta}}(1+\epsilon)^{\frac{1}{\theta}-1} > 0 \\
\frac{\partial d_0^s}{\partial \epsilon} &= \frac{1}{\theta}\rho_{0,1}^{\frac{1}{\theta}}\rho_{1,2}^{\frac{1}{\theta}}\left(\left(\frac{1}{1+\rho_{1,2}^{\frac{1}{\theta}}}\right)^{1-\theta} + \rho_{1,2}^{\frac{1}{\theta}}\left(\frac{1}{1+\rho_{1,2}^{\frac{1}{\theta}}}\right)^{1-\theta}(1+\epsilon)\right)^{\frac{1}{\theta}-1}\left(\frac{1}{1+\rho_{1,2}^{\frac{1}{\theta}}}\right)^{1-\theta} \\
&= \frac{1}{\theta}\rho_{0,1}^{\frac{1}{\theta}}\rho_{1,2}^{\frac{1}{\theta}}\left(\frac{1}{1+\rho_{1,2}^{\frac{1}{\theta}}} + \rho_{1,2}^{\frac{1}{\theta}}\frac{1}{1+\rho_{1,2}^{\frac{1}{\theta}}}(1+\epsilon)\right)^{\frac{1}{\theta}-1} \\
&= \frac{1}{\theta}\rho_{0,1}^{\frac{1}{\theta}}\rho_{1,2}^{\frac{1}{\theta}}\left(1+\epsilon\frac{\rho_{1,2}^{\frac{1}{\theta}}}{1+\rho_{1,2}^{\frac{1}{\theta}}}\right)^{\frac{1}{\theta}-1} \\
&= \frac{1}{\theta}\rho_{0,1}^{\frac{1}{\theta}}\rho_{1,2}^{\frac{1}{\theta}}\left(1+\epsilon(1-m_1)\right)^{\frac{1}{\theta}-1} > 0
\end{aligned}$$

and we thus get

$$y \equiv \frac{\frac{\partial d_0^s}{\partial \epsilon}}{\frac{\partial d_0^n}{\partial \epsilon}} = z^{\frac{1}{\theta}-1} \text{ where } z \equiv \frac{1+\epsilon(1-m_1)}{1+\epsilon}.$$

Furthermore, since

$$\frac{\partial \ln y}{\partial \theta} = \underbrace{-\left(\frac{1}{\theta}\right)^2 \ln(z)}_{>0} + \left(\frac{1}{\theta}-1\right)\underbrace{\frac{\partial \ln(z)}{\partial \theta}}_{>0}$$

we obtain

$$\frac{\partial y}{\partial \theta} > 0, \text{ for } \psi \equiv \frac{1}{\theta} > \frac{1}{\bar{\theta}} \in (0,1) \tag{35}$$

where for the case $\theta > 1$ the threshold value $\bar{\theta}$ is defined as the root of the non-linear equation

$$-\left(\frac{1}{\theta}\right)^2 \ln(z) + \left(\frac{1}{\theta}-1\right)\frac{\partial \ln(z)}{\partial \theta} = 0.$$

which by the intermediate value theorem exists and is unique.

The intuition for the upper bound parameter $\bar{\theta} \in (1,\infty)$—and thus corresponding lower bound IES of $\underline{\psi} \in (0,1)$ is that for the limiting case $\theta = \infty$ we have perfect complements of consumption over time so that consumption is constant, $c_h^s = c_h^n = \bar{c}$ for all h and therefore the sophisticated and the naive agent save and consume the same amount if $\theta = \infty$ and $\theta = 1$ and thus the derivative in (15) changes its sign for $\theta < \bar{\theta}$.

It remains to show that $\frac{\partial \ln(z)}{\partial \theta} > 0$. We have

$$\frac{\partial \ln(z)}{\partial \theta} = -\frac{1}{1+\epsilon(1-m_1)}\epsilon\frac{\partial m_1}{\partial \theta} > 0$$

because

$$\frac{\partial m_1}{\partial \theta} = -\frac{1}{(1+\rho_{1,2})^2}\frac{\partial \rho_{1,2}^{\frac{1}{\theta}}}{\partial \theta} < 0.$$

C Calibration

For the calibration of the quantitative illustration in Section 4.4 we assume that agents live for $T = 60$ years, are born with a wealth endowment of $w_0 = 100$, which, to generate a life-cycle asset profile, we distribute as endowment income of 2 until retirement and a pension income of 1 after retirement at age 40. We assume two types of agents with $\theta_i \in \{0.5, 2\}$ with equal measure of 50% and within each preference type we assume equal shares of naive and sophisticated hyperbolic time discounters with long-run discount factor δ and short-run discount factor β so that the economy is populated by 4 different agents with equal measure of 25%. To discipline discounting in our deterministic model with a zero interest rate, we calibrate the long-run discount factor δ so that the aggregate asset holdings in the model economy match smoothed data on assets taken from Groneck et al. (2016), whereby we hold constant the difference between the short and the long-run discount rates Δ as calibrated by Angeletos et al. (2001), i.e., $\Delta = 1/0.7 - 1/0.957 = 0.383$. Matching asset profiles gives $\delta = 1.035$ and $\beta = \delta - \Delta = 0.741$ for the long- and short-run discount factors, respectively.[18]

[18] Note that in our model with a zero interest rate a long-run discount factor above one is not surprising.

References

Ando, A. and F. Modigliani (1963). The 'Life-Cycle' Hypothesis of Saving: Aggregate Implications and Tests. American Economic Review 53(1), 55–84.

Andreoni, J. and C. Sprenger (2012). Risk Preferences Are Not Time Preferences. American Economic Review 102(7), 3357–3376.

Bansal, R., and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. The Journal of Finance, 59(4), 1481–1509.

Bleichrodt, H. and L. Eeckhoudt (2006). Survival Risks, Intertemporal Consumption, and Insurance: The Case of Distorted Probabilities. Insurance: Mathematics and Economics 38(2), 335–346.

Bommier, A., D. Harenberg, and F. Le Grand (2021). Recursive Preferences and the Value of Life: A Clarication, Working Paper.

Bommier, A., D. Harenberg, F. Le Grand, and C. O'Dea (2020). Recursive Preferences, the Value of Life, and Household Finance. Cowles Foundation Discussion Paper No. 2231.

Chakraborty, A. Y. Halevy, and K. Saito (2020). The Relation between Behavior under Risk and over Time. American Economic Review: Insights 2 (1): 1–16.

Cordoba, J. C. and M. Ripoll (2017). Risk Aversion and the Value of Life. Review of Economic Studies 84, 1472–1509.

Deaton, A. (1992). Understanding Consumption. Oxford: Clarendon Press.

Drouhin, N. (2015). A Rank-dependent Utility Model of Uncertain Lifetime. Journal of Economic Dynamics and Control 53, 208–224.

Eichberger, J., S. Grant, and D. Kelsey (2007). Updating Choquet Beliefs. Journal of Mathematical Economics 43(7), 888–899.

Eichberger, J., S. Grant, and D. Kelsey (2012). When is Ambiguity-Attitude Constant? Journal of Risk and Uncertainty 45(3), 239–263.

Epper, T., H. Fehr-Duda, and A. Bruhin (2011). Viewing the Future through a Warped Lens: Why Uncertainty Generates Hyperbolic Discounting. Journal of Risk and Uncertainty 43, 169–203.

Epstein, L. G. (2006). An Axiomatic Model of Non-Bayesian Updating. The Review of Economic Studies 73, 413–436.

Epstein, L. G. and M. Le Breton (1993). Dynamically Consistent Beliefs must be Bayesian. Journal of Economic Theory 61, 1–22.

Epstein, L. G. and M. Schneider (2003). Recursive Multiple-Priors. Journal of Economic Theory 113, 1–31.

Epstein, L. G., J. Noor, and A. Sandroni (2008). Non-Bayesian Updating: A Theoretical Framework. Theoretical Economics 3, 193–229.

Epstein, L. G. and S.E. Zin (1989). Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. Econometrica, 57(4), 937–969.

Epstein, L. G. and S.E. Zin (1991). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis. Journal of Political Economy, 99(2), 263–286.

Gennaioli, N. and A. Shleifer (2010). What Comes to Mind. The Quarterly Journal of Economics 125, 1399–1433.

Ghirardato, P. (2002). Revisiting Savage in a Conditional World. Economic Theory 20(1), 83–92.

Gilboa, I. (1987). Expected Utility with Purely Subjective Non-Additive Probabilities. Journal of Mathematical Economics 16, 65–88.

Gilboa, I. and D. Schmeidler (1993). Updating Ambiguous Beliefs. Journal of Economic Theory 59(1), 33–49.

Grevenbrock, N., M. Groneck, A. Ludwig, and A. Zimper (2020). Cognition, Optimism and the Formation of Age-Dependent Survival Beliefs. International Economic Review 62, 887–918.

Groneck, M., A. Ludwig, and A. Zimper (2016). A Life-Cycle Model with Ambiguous Survival Beliefs. Journal of Economic Theory 162, 137–180.

Groneck, M., A. Ludwig, and A. Zimper (2021). Who Saves More, the Naïve or the Sophisticated Agent? Working paper version, CEPR DP16412.

Halevy, Y. (2008). Strotz Meets Allais: Diminishing Impatience and the Certainty Effect. American Economic Review 98(3), 1145–62.

Harris, C. and D. Laibson (2001). Dynamic Choices of Hyperbolic Consumers. Econometrica 69(4), 935–957.

Hugonnier, J., F. Pelgrin, and P. St-Amour (2013). Health and (Other) Asset Holdings. Review of Economic Studies 80, 663–710.

Kimball, M., and P. Weil (2009). Precautionary saving and consumption smoothing across time and possibilities. Journal of Money, Credit and Banking, 41(2-3), 245–284.

Krebs, T. (2003). Human Capital Risk and Economic Growth. The Quarterly Journal of Economics 118(2), 709–744.

Krueger, D., Ludwig, A., and S. Villalvazo (2021). Optimal taxes on capital in the OLG model with uninsurable idiosyncratic income risk. Journal of Public Economics, 201, 104491.

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. Quarterly Journal of Economics 62(2), 443–477.

Laibson, D. (1998). Life-cycle Consumption and Hyperbolic Discount Functions. European Economic Review 42(3), 861–871.

Ludwig, A. and A. Zimper (2013). A Parsimonious Model of Subjective Life Expectancy. Theory and Decision 75, 519–542.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. The Review of Economics and Statistics, 247–257.

Modigliani, F. and R. Brumberg (1954). Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data. In K. K. Kurihara (ed), Post-Keynesian Economics. New Brunswick, NJ.: Rutgers University Press.

Mullainathan, S., J. Schwartzstein, and A. Shleifer (2008). Coarse Thinking and Persuasion. The Quarterly Journal of Economics 123, 577–619.

O'Donoghue, T. and M. Rabin (1999). Doing It Now or Later. American Economic Review 89(1), 103–124.

Ortoleva, P. (2012). Modeling the Change of Paradigm: Non-Bayesian Reactions to Unexpected News. The American Economic Review 102, 2410–2436.

Phelps, E. S. and R. Pollak (1968). On Second-Best National Saving and Game- Equilibrium Growth. Review of Economic Studies, 35, 185–199.

Pollak, R. A. (1968). Consistent Planning. Review of Economic Studies 35, 201–208.

Rabin, M. (2002). Inference by Believers in the Law of Small Numbers. The Quarterly Journal of Economics 117, 775–816.

Rabin, M. and J. L. Schrag (1999). First Impressions Matter: A Model of Confirmatory Bias. The Quarterly Journal of Economics 114, 37–82.

Saito, K. (2011). Strotz Meets Allais: Diminishing Impatience and the Certainty Effect: Comment. American Economic Review 101(5): 2271–75.

Samuelson, P. A. (1937). A Note on Measurement of Utility. The Review of Eco¡nomic Studies 4(2), 155–161.

Samuelson, P. A. (1969). Lifetime Portfolio Selection by Dynamic Stochastic Programming. The Review of Economics and Statistics 51(3), 239–246.

Schmeidler, D. (1989). Subjective Probability and Expected Utility Without Additivity. Econometrica 57(3), 571–587.

Siniscalchi, M. (2011). Dynamic Choice under Ambiguity. Theoretical Economics 6, 379–421.

Strotz, R. H. (1955). Myopia and Inconsistency in Dynamic Utility Maximization. The Review of Economic Studies 23(3), 165–180.

Tversky, A. and D. Kahneman (1992). Advances in Prospect Theory: Cumulative Representations of Uncertainty. Journal of Risk and Uncertainty 5(4), 297–323.

Wakker, P. P. (2010). Prospect Theory: For Risk and Ambiguity. Cambridge, UK: Cambridge University Press.

Wakker, P. P. and A. Tversky (1993). An Axiomatization of Cumulative Prospect Theory. Journal of Risk and Uncertainty 7, 147–176.

Weil, P. (1989). The Equity Premium Puzzle and the Risk-free Rate Puzzle. Journal of Monetary Economics, 24(3), 401–421.

Supplementary Appendix

Extension to EZW Preferences

(Not for Publication)

A Dynamic Budget Constraint

The budget constraint in terms of financial wealth e_t is

$$a_{t+1} = a_t R_t^p(\alpha_{t-1}) + e_t - c_t$$

for $a_0 = 0$ given. In terms of cash on hand $x_t = a_t R_t^p(\alpha_{t-1}) + e_t$ we can rewrite the budget constraint as

$$x_{t+1} = (x_t - c_t)\, R_{t+1}^p(\alpha_t) + e_{t+1}. \tag{36}$$

Since human capital as the discounted sum of future deterministic labor income obeys

$$h_{t+1} = h_t R^f - e_{t+1}. \tag{37}$$

Consolidating budget constraints (36) and (37) gives (17) and (18).

B Solution

The marginal propensities to consume and a characterization of the optimal portfolio choice resulting from the solution of the consumption savings and portfolio allocation problem of the naive and the sophisticated agents are given in the next proposition, which we formally prove below:

Proposition 9. *Consider the EZW life-cycle model with arbitrary discount factors. The marginal propensities to consume are given as follows:*

- *for the sophisticated agent:*

$$m_h^{s,h} = \begin{cases} 1 & \text{for } h = T \\ \frac{1}{1+\left(\rho_{h,h+1}\zeta_{h+1}^h \Theta\left(\hat{\alpha}_t, R^f, R_{h+1}, \pi\right)\right)^{\frac{1}{\theta}}} & \text{for } h < T, \end{cases} \tag{38}$$

where ζ_{h+1}^h *follows from the backward recursion in* $t = T-1, \ldots, h$

$$\zeta_t^h = m_t^{s^{1-\theta}} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \left(1 - m_t^s\right)^{1-\theta} \zeta_{t+1}^h \cdot \Theta\left(\hat{\alpha}_t^*, R^f, R_{t+1}, \pi\right) \tag{39}$$

for $\zeta_T^h = 1$*, where for all* $t = h, \ldots, T-1$

$$\Theta\left(\hat{\alpha}_t, R^f, R_{t+1}, \pi\right) = \max_{\hat{\alpha}_t} \left\{ \left(\int R_{t+1}^p(\hat{\alpha}_t)^{1-\sigma} d\pi \right)^{\frac{1-\theta}{1-\sigma}} \right\}. \tag{40}$$

- *for the naive agent:*

$$m_t^{n,h} = \begin{cases} 1 & \text{for } t = T \\ \dfrac{1}{1 + \left(\frac{\rho_{h,t+1}}{\rho_{h,t}} \Theta\left(\hat{\alpha}_t, R^f, R_{t+1}, \pi\right) \right)^{\frac{1}{\theta}} \left(m_{t+1}^{n,h}\right)^{-1}} & \text{for } t < T, \end{cases} \tag{41}$$

where $\Theta(\cdot)$ *is given by (40).*

- *for both agents the optimal portfolio choice* $\hat{\alpha}_t^s = \hat{\alpha}_t^n = \hat{\alpha}_t$ *is the solution to*

$$\int R_{t+1}^p(\hat{\alpha}_t)^{-\sigma} d\pi = 0 \tag{42}$$

We thus find that the separation between risk attitudes as measured by σ and inter-temporal attitudes as measured by θ inherent to EZW preferences is reflected in the solution of this model to the effect that both households choose the same optimal portfolio share $\hat{\alpha}_t$ as the solution to (42)—which due to the convexity of the function $R_{t+1}^p(\hat{\alpha}_t)^{-\sigma}$ in the portfolio share is decreasing in risk aversion σ—, whereas the relationship between the marginal propensities to consume out of total wealth across the two types of households is exclusively driven by inter-temporal attitudes as measured by θ. Specifically, as in our recursive proof in Subsection A we likewise find that

$$m_h^n \leq m^s \qquad \Leftrightarrow \qquad \left(m_{h+1}^{n,h}\right)^{\theta} \zeta_{h+1}^h \leq 1$$

and since

$$\frac{\rho_{h,t+1}}{\rho_{h,t}} \Theta\left(\hat{\alpha}_t, R^f, R_{t+1}, \pi\right) = \left(\frac{1 - m_t^{n,h}}{m_t^{n,h}} \right)^{\theta} m_{t+1}^{n,h}$$

we can use the above in equation (39) to obtain (21). An application of the analogous steps as in the backward recursive proof of Theorem 1 finally gives us Corollary 4.

C Implications for Portfolio Choice

Our finding on marginal propensities to consume in Theorem 1 combined with the finding of equal (across the two types) optimal portfolio shares $\hat{\alpha}_t$ leads us to the next observation regarding the portfolio shares as a fraction of financial wealth α_t^i for $i \in \{n, s\}$. Recall from the definition of $\hat{\alpha}_t^i$ in (18) that

$$\alpha_t^i = \hat{\alpha}_t \left(1 + \frac{h_t}{x_t^i - c_t^i}\right)$$

and since (the optimal) $\hat{\alpha}_t$ and h_t are the same for both types of households, differences in the optimal portfolio choice out of financial wealth, α_t^i, across the two types are solely due to differences in $x_t^i - c_t^i$. Specifically, we get

$$\alpha_t^s \leq \alpha_t^n \qquad \Leftrightarrow \qquad x_t^s - c_t^s \geq x_t^n - c_t^n \qquad \Leftrightarrow \qquad w_t^s(1 - m_t^s) \geq w_t^n(1 - m_t^n).$$

Next, assume that the return realizations R_t are the same for the naive and the sophisticated household (aggregate return risk). Then, since at all t wealth accumulation, or decumulation, obeys (17) and since $\hat{\alpha}_t^i = \hat{\alpha}_t$, for $i \in \{n, s\}$ we obtain

$$m_t^s \leq m_t^n \;\Leftrightarrow\; (1 - m_t^s)w_t^s \geq (1 - m_t^s)w_t^n \;\Rightarrow\; w_{t+1}^s \geq w_{t+1}^n.$$

where the last inequality follows from (17) because $\hat{\alpha}_t^i = \hat{\alpha}_t$ and by our assumption of aggregate return risk so that return realizations are the same for the naive and the sophisticated household. We thus arrive at Corollary 5.

D Proof of Proposition 9

Our proof of Proposition 9 is based on recursive methods.

Sophisticated Agent. Our proof is by backward induction.

Claims: The value function of the sophisticated agent in any period $t \geq h$ is given by

$$U_t^h(w_t) = \frac{1}{1-\theta} \zeta_t^h w_t^{1-\theta} \tag{43}$$

with associated policy function

$$c_h^s = m_h^s w_h. \tag{44}$$

Base case: In period T we have $c_T^s = w_T$ and thus $U_T^h = \frac{1}{1-\theta} w_T^{1-\theta}$ and $m_T^s = 1$.

Backward Induction Steps: Suppose the claims (43) and (44) have been shown for all periods $h+1, \ldots, T$. Then iterate backward for all $t = T-1, \ldots, h+1$ using (43) in (16) to get, also using resource constraint (17),

$$\begin{aligned} U_t^h &= u(c_t) + \frac{\rho_{h,t+1}}{\rho_{h,t}} \frac{1}{1-\theta} \left(\mathbb{E}\left[\left((1-\theta) U_{t+1}^h \right)^{\frac{1-\sigma}{1-\theta}} \right] \right)^{\frac{1-\theta}{1-\sigma}} \\ &= \frac{1}{1-\theta} \left(\left(c_t^s\right)^{1-\theta} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \zeta_{t+1}^h \left(\mathbb{E}\left[\left(w_{t+1}^{1-\theta} \right)^{\frac{1-\sigma}{1-\theta}} \right] \right)^{\frac{1-\theta}{1-\sigma}} \right) \\ &= \frac{1}{1-\theta} \left(\left(m_t^s\right)^{1-\theta} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \left(1 - m_t^s\right)^{1-\theta} \zeta_{t+1}^h \Theta\left(\hat{\alpha}_t, R^f, R_{t+1}, \pi \right) \right) w_t^{1-\theta} \\ &= \frac{1}{1-\theta} \zeta_t^h w_t^{1-\theta}, \end{aligned} \tag{45}$$

which defines (40) and establishes the backward recursion of ζ_t^h in (39).

Next, in period h use (43) in (16) to get

$$U_h^h = \frac{1}{1-\theta} \max_{c_h^s, w_{h+1}, \hat{\alpha}_h^s} \left\{ \left(c_h^s\right)^{1-\theta} + \rho_{h,h+1} \zeta_{h+1}^h \left(\mathbb{E}\left[\left(w_{h+1}^{1-\theta} \right)^{\frac{1-\sigma}{1-\theta}} \right] \right)^{\frac{1-\theta}{1-\sigma}} \right\}. \tag{46}$$

Use the resource constraint (17) in the above to obtain, by the separation between the optimal consumption and the optimal portfolio choice,

$$U_h^h = \frac{1}{1-\theta} \max_{c_h^s} \left\{ \left(c_h^s\right)^{1-\theta} + \rho_{h,h+1} \left(w_h - c_h^s\right)^{1-\theta} \right\} \zeta_{h+1}^h \underbrace{\max_{\hat{\alpha}_h} \left\{ \left(\mathbb{E}\left[\left(R_{h+1}^p (\hat{\alpha}_h)^{1-\theta} \right)^{\frac{1-\sigma}{1-\theta}} \right] \right)^{\frac{1-\theta}{1-\sigma}} \right\}}_{=\Theta\left(\hat{\alpha}_h, \zeta_{h+1}^h, R^f, R_{h+1}, \pi\right)}$$

with first-order condition for c_h^s

$$\left(c_h^s\right)^{-\theta} - \rho_{h,h+1} \left(w_h - c_h\right)^{-\theta} \zeta_{h+1}^h \Theta\left(\hat{\alpha}_h, R^f, R_{h+1}, \pi\right) = 0,$$

where $\hat{\alpha}_h^*$ is the optimal portfolio share further characterized below. We thus get

$$c_h^s = m_h^s w_h$$

where

$$m_h^s = \frac{1}{1 + \left[\rho_{h,h+1} \zeta_{h+1}^h \Theta\left(\hat{\alpha}_h, R^f, R_{h+1}, \pi\right) \right]^{\frac{1}{\theta}}}.$$

which is (46) and proves the claims.

Naive Agent. For the naive agent, we essentially follow the same steps with the following modifications:

- The maximization problem in (46) is solved for all $t = h, \dots, T-1$, thus

$$U_t^{n,h} = \frac{1}{1-\theta} \max_{c_t^{n,h}, w_{t+1}, \hat{\alpha}_t^{n,h}} \left\{ \left(c_t^{n,h}\right)^{1-\theta} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \zeta_{t+1}^h \left(\mathbb{E}\left[\left(w_{t+1}^{1-\theta}\right)^{\frac{1-\sigma}{1-\theta}}\right]\right)^{\frac{1-\theta}{1-\sigma}} \right\},$$

which, using the resource constraint and the separation between the optimal consumption and the portfolio choice, gives

$$m_t^{n,h} = \frac{1}{1 + \left[\frac{\rho_{h,t+1}}{\rho_{h,t}} \zeta_{t+1}^h \Theta\left(\hat{\alpha}_t, R^f, R_{t+1}, \pi\right)\right]^{\frac{1}{\theta}}}. \tag{47}$$

- Using the solution back in the value function as in (45) gives

$$\begin{aligned} U_t^{n,h} &= \frac{1}{1-\theta} \left(\left(m_t^{n,h}\right)^{1-\theta} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \zeta_{t+1}^{n,h} \left(1 - m_t^{n,h}\right)^{1-\theta} \Theta\left(\hat{\alpha}_t^{n,h}, R^f, R_{t+1}, \pi\right) \right) w_t^{1-\theta} \\ &= \frac{1}{1-\theta} \left(\left(m_t^{n,h}\right)^{1-\theta} + \left(1 - m_t^{n,h}\right)^{1-\theta} \left(\frac{1 - m_t^{n,h}}{m_t^{n,h}}\right)^{\theta} \right) w_t^{1-\theta} \\ &= \frac{1}{1-\theta} \left(m_t^{n,h}\right)^{-\theta} w_t^{1-\theta}. \end{aligned}$$

- We thus find $\zeta_t^h = m_t^{h^{-\theta}}$. Using this in (47) we finally obtain

$$m_t^h = \frac{1}{1 + \left(\frac{\rho_{h,t+1}}{\rho_{h,t}} \Theta\left(\hat{\alpha}_t^{n,h}, R^f, R_{t+1}, \pi\right)\right)^{\frac{1}{\theta}} \left(m_{t+1}^{n,h}\right)^{-1}}.$$

Optimal Portfolio Choice. Since $\Theta\left(\hat{\alpha}_t, R^f, R_{t+1}, \pi\right)$ is the same for both agents we obtain $\hat{\alpha}_t^s = \hat{\alpha}_t^{n,h} = \hat{\alpha}_t$, where from the first-order condition of the optimal portfolio allocation problem $\hat{\alpha}_t^*$ is the solution to

$$\mathbb{E}\left[R_{t+1}^p(\hat{\alpha}_t)^{-\sigma}\right] = \int R_{t+1}^p(\hat{\alpha}_t)^{-\sigma} d\pi = 0$$

and thus the optimal portfolio allocation problem at t is a static decision problem, which is parameterized by risk aversion σ. □□