The Impact of Biases in Survival Beliefs on Savings Behavior*

Max Groneck† Alexander Ludwig‡ Alexander Zimper§

March 24, 2017

Abstract

On average young people "undersave" whereas old people "oversave" with respect to the rational expectations model of life-cycle consumption and savings. According to numerous studies on subjective survival beliefs, young people also "underestimate" whereas old people "overestimate" their objective survival chances on average. We take a structural behavioral economics approach to jointly address both empirical phenomena by embedding subjective survival beliefs that are consistent with these biases into a rank-dependent utility (RDU) model over life-cycle consumption. The resulting consumption behavior is dynamically inconsistent. Considering both naive and sophisticated RDU agents we show that within this framework underestimation of young age and overestimation of old age survival probabilities may (but need not) give rise to the joint occurrence of undersaving and oversaving. In contrast to this RDU model, the familiar quasi-hyperbolic discounting (QHD), which is nested as a special case, cannot generate oversaving.

JEL Classification: D91, D83, E21.

Keywords: saving puzzles; subjective survival beliefs; behavioral economics; prospect theory; neo-additive probability weighting; dynamic inconsistency; sophisticated versus naive behavior; quasi-hyperbolic discounting

*We thank Menusch Khadjavi, Ulrich Schmidt, Peter Wakker and Mirko Wiederholt for valuable comments and suggestions. Alex Ludwig gratefully acknowledges financial support by Research Center SAFE, funded by the State of Hessen initiative for research LOEWE.

†Stockholm School of Economics; Netspar; P.O. Box 6501 SE-113 83 Stockholm, Sweden; E-mail: max.groneck@hhs.se

‡SAFE, Goethe University Frankfurt; CMR; MEA; Netspar; House of Finance; Theodor-W.-Adorno-Platz 3; 60629 Frankfurt am Main (Germany); E-mail: ludwig@safe.uni-frankfurt.de

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

1 Introduction

How households consume and save over the life-cycle and how time preferences as well as the time horizon affect these decisions are classical economic questions. The workhorse model to address this problem of inter-temporal allocation is the life-cycle model of Modigliani and Brumberg (1954) and Ando and Modigliani (1963). This standard model considers an expected utility maximizing agent with an additively separable per period utility function. The agent's future utility is discounted by an effective discount factor comprised of the pure rate of time-preference and her belief to survive into the future. Since the pioneering work of Samuelson (1937), pure time-preferences are typically described by an exponential discount function. Following Muth (1961) it has also become standard to express survival beliefs as objective survival probabilities. This standard *rational expectations (RE)* life-cycle model gives rise to three well established saving puzzles: compared to the RE model, households in the data save too little at young age (undersaving), cf., e.g., Laibson et al. (1998) and Bernheim and Rangel (2007), and hold on to their assets until too late in life (oversaving, respectively high old-age asset holdings), cf., e.g., De Nardi et al. (2010), Hurd and Rohwedder (2010) and Lockwood (2012). In addition, there is ample empirical evidence for dynamically inconsistent savings behavior, again see Laibson et al. (1998) and Bernheim and Rangel (2007).[1]

Our main objective is to develop a behavioral theory that can jointly accommodate these savings puzzles. The point of departure of our analysis is the robust finding from survey data on subjective survival beliefs that "young" respondents (younger than about 65) tend to underestimate whereas "old" respondents (older than about 70) tend to overestimate their survival chances (Hammermesh 1985; Manski 2004; Gan et al. 2005; Peracchi and Perotti 2010; Elder 2013; Ludwig and Zimper 2013). Intuitively, one would conjecture that such age-dependent biases between perceived and objective survival chances are an important driver of the empirically observed savings puzzles. After all, individuals who do not expect to live for long will consume in the presence rather than save for the future to the effect that underestimation of survival chances at young age should give rise to undersaving early in life. Conversely, overestimation of survival chances at an old age should lead to oversaving later in life

[1] Also see Barsky et al. (1997), Angeletos et al. (2001), Choi et al. (2006) and Lusardi and Mitchell (2011).

because we would expect that such overly optimistic individuals transfer more wealth to their future than their rational expectations counterparts.

To rigorously analyze this conjecture, we incorporate a model of subjective survival beliefs into a variant of a life-cycle model of consumption and savings. Specifically, we consider a simple transformation of objective survival beliefs, whereby a *likelihood-insensitivity* parameter controls the decision weight on objective survival information. The higher is the likelihood insensitivity, the less relevant are objective survival rates for economic decisions. In presence of such likelihood insensitivity a second *optimism* parameter governs the strength of underestimation versus overestimation of survival chances. We show that this two-parameter transformation of objective survival beliefs can easily replicate the age-dependent survival belief biases reported in survey data. Such a transformation is known as a *neo-additive* probability weighting function which is popular in the literature because it approximates in a parsimonious way the inverse S-shaped probability weights typically elicited in experimental *prospect theory* (Kahneman and Tversky 1979; Tversky and Kahneman 1992).[2]

Next, we assume that in presence of risky survival chances individuals prefer longer consumption streams (longer horizons) to shorter ones. This natural notion gives rise to a *rank dependent utility* (RDU) (cf. Quiggin 1981, 1982) life-cycle model defined over gains that arise from consumption streams.[3] Since RDU and prospect theory coincide on the domain of gains, our model stands for an application of prospect theory to life-cycle consumption with survival risk.[4] We further assume additive separability with exponential time-discounting and per-period utility functions that are of the power form featuring a constant inter-temporal elasticity of substitution (IES). We show that using neo-additive survival beliefs in this RDU model gives rise to dynamically inconsistent consumption behavior. That is, future consumption plans generally deviate from present plans for these future periods.

Our choice of a neo-additive probability weighting function results in an analyti-

[2]Cf., e.g., Abdellaoui et al. (2011). Wakker (2009, p. 208-2010) discusses in depth RDU models with neo-additive probability weighting functions (also see the references therein).
For an axiomatic foundation of neo-additive probability measures within Choquet expected utility theory (Gilboa 1987; Schmeidler 1989) see Chateauneuf et al. (2007).

[3]See Quiggin (1993) for a textbook treatment. As Machina (1994) observes in a review of this book the "publication history of the rank-dependent expected utility model attests to its role as the most natural and useful modification of the classical expected utility formula."

[4]In addition, our RDU model with a neo-additive probability weighting function can also be regarded as a formal special case of *security- and potential level models* which generalize expected utility under risk without any reference to probability weighting (Cohen 1992; Essid 1997).

cally very tractable model enabling us to characterize the entire solution of the optimization problem of naive and sophisticated RDU agents in a multi-period life-cycle model in closed form. This tractability is useful when we compare the decision problems of naive and sophisticated RDU agents to other nested models. The tractability pays off in particular when we employ a three-period variant of our model to compare under- versus overestimation of survival beliefs to conditions for under- versus oversaving in terms of the single optimism parameter.

In terms of related models, first, our RDU model falls under the class of models with general discount functions initiated by Strotz (1955) and Pollak (1968). In particular, we revisit an important equivalence result derived in Pollak (1968), namely that consumption behavior of naive and sophisticated agents coincide for a logarithmic per period utility function. We provide a new interpretation for this finding. For the three-period variant of our model we also establish that the sophisticated agent saves more (less) than the naive agent in the first period of life when the IES is less than one (above one). In our numerical analysis of that model variant we further show that these differences become stronger when likelihood insensitivity is large and the IES is low. In quantitative work, this insight might be useful to differentiate between naive and sophisticated agents, respectively to identify the degree of sophistication.

Second, we show that our RDU model nests the familiar quasi-hyperbolic time discounting (QHD) model made popular by Laibson (1997) as a special case. QHD models introduce a short-run discount factor between the present and the first future period which is lower than the discount factor between any other two future subsequent periods thereby giving rise to diminishing impatience in the form of short-run impatience.[5] Applied to a life-cycle model, this short-run impatient QHD decision maker becomes formally equivalent to a neo-additive RDU decision maker who is extremely pessimistic in the sense that her optimism parameter takes on the value zero. To the best of our knowledge, the closed form solutions and interpretations we provide for the multi-period model are also novel to the QHD literature. For the general parametrization with non-zero optimism parameter, one crucial difference to this nested QHD model emerges: the QHD model naturally gives rise to undersaving but not to oversaving, hence the QHD model cannot simultaneously address both savings

[5]Quasi-hyperbolic discounting is therefore in line with overwhelming experimental evidence that reports a conflict between decision makers' long-run desire to be patient and their short-run desire for instantaneous gratification. See, e.g., Ainslie (1992), Loewenstein and Thaler (1989), Laibson (1997), Laibson 1998.

puzzles. This aspect is therefore an important dimension along which the two models are not observationally equivalent. Also, Halevy (2015) presents experimental data according to which *time-variant* preferences are empirically more relevant than *time-invariant* preferences as presumed by the QHD model. Since the RDU preferences of our life-cycle model are time-variant if, and only if, they do not reduce to QHD (or RE) preferences, we regard Halevy's (2015) finding as independent empirical evidence in favour of a general RDU life-cycle model that cannot be reduced to a QHD model.

With respect to the implications of biased survival beliefs for savings behavior in our general RDU model, our closed form solutions establish two main results. We derive these results by employing the three-period variant of our model for which we present clear-cut analytical conditions. First, overestimation of old age survival chances is sufficient and necessary for oversaving at an old age. Second, in presence of old age overestimation, sufficiently strong underestimation of survival chances at young age results in undersaving at young age. The details of this latter result crucially depend on whether the agent is aware of her dynamically inconsistent RDU preferences (i.e., sophisticated) or not (i.e., naive). A combination of both results pins down parameter conditions for which underestimation at young combined with overestimation at old ages generates undersaving at young combined with oversaving at old ages.

This analytical characterization of savings behavior in the RDU model refers to undersaving and oversaving in terms of the flow of savings but not in terms of the stock of asset holdings. Yet, the previously cited quantitative life-cycle literature also tries to understand high old age asset holdings, which results from persistent oversaving relative to the RE benchmark model. In our context, where biases in subjective survival beliefs are the key driver of savings behavior, high old-age asset holdings emanate simultaneously with undersaving at young age if overestimation of survival beliefs is sufficiently strong so that oversaving in old age eventually dominates undersaving at young age. We show this through a simple numerical analysis, again for the three-period variant of our model, thereby establishing that the RDU model may contribute to explaining the empirical phenomena of undersaving, oversaving and high old-age asset holdings. Observe that there is an important tension regarding the biases in survival beliefs: On the one hand, to generate undersaving at young age in presence of underestimation of survival chances at young age and overestimation at old age, underestimation of survival probabilities must be sufficiently strong. It must dominate the effects of old age overestimation on young age savings behavior for forward looking agents who anticipate that they will overestimate survival probabilities eventually. On the other hand, to simultaneously generate high old age asset holdings, overestimation must be strong enough. Eventually, it must dominate the effects of underestimation on asset holdings. How these tensions play out are quantitative questions which we address in Groneck, Ludwig, and Zimper (2016).

Our analysis will provide useful guidance for quantitative work to distinguish between different, partially nested, behavioral models of life-cycle consumption, both on qualitative (with respect to our RDU model versus the standard QHD model) as well as on quantitative grounds (with respect to the degree of sophistication).

The remainder of our analysis proceeds as follows. Section 2 revisits the key stylized facts on biases in survival perceptions and introduces neo-additive survival beliefs. Section 3 constructs our RDU life-cycle model and Section 4 solves for the consumption behavior for a multi-period RDU life-cycle model in closed form. Section 5 employs a three-period variant of our model to revisit the aforementioned savings puzzles. Section 6 complements our analytical results by numerical analyses. Section 7 concludes with a discussion of our main results and an outlook on possible avenues for future research. All formal proofs are relegated to the Appendix and a Supplementary Appendix contains additional results.

2 Misperceptions of Survival Chances

This section revisits the key stylized facts of survival misperceptions by documenting biases measured in the Health and Retirement Study (HRS). We then proceed with our theoretical framework which applies neo-additive probability weighting functions to objective survival probabilities. We show that for suitable values of the decision maker's degree of *optimism* such neo-additively transformed survival probabilities naturally result in underestimation of survival chances at young and overestimation at old age, just as observed in the data.

2.1 Point of Departure: Stylized Facts

Figure 1 reports average differences between subjective survival rates as elicited in the Health and Retirement Study (HRS) and objective cohort survival rates by age.[6] Each data point represents the average bias in long-run survival chances: It measures the average distance between the subjective belief to survive from the age at interview—depicted on the abscissa—to some target age which is several years ahead. The data pattern mirrors findings in numerous empirical studies on subjective survival beliefs, cf. Groneck, Ludwig, and Zimper (2016) and references therein: Until the age of about 70, respondents underestimate whereas later in life they overestimate their chances to survive into the future.[7] These biases are relatively large. For example, on average 50 year old respondents in the sample underestimate their chances to survive until age 80 by about 15 percentage points whereas 85 year old respondents on average overestimate their chances to survive until age 95 by roughly 18 percentage points.

2.2 Theoretical Framework: Neo-additive Survival Beliefs

Consider an agent of age $h \geq 0$ and fix some $T \geq h$ with the interpretation that the agent possibly lives until the maximal age of T. We construct the additive probability space $(\Omega, \mathcal{F}, \psi)$ such that the state space is given as $\Omega = \{1, \dots, T\}$ and the σ-algebra $\mathcal{F}$ is given as the powerset of Ω. We interpret $D_t \equiv \{t\}$, $t \in \Omega$ as the event in $\mathcal{F}$ that the agent dies at the end of age t. Observe that $D_t \cup \cdots \cup D_T$ stands for the event in $\mathcal{F}$ that the agent of age $h < t$ survives until (at least) the beginning of age t.

We interpret ψ as the objective (unconditional) probability measure that comprehensively governs the agent's mortality risk. Denote by ψ^h the conditional probability measure of an agent who has reached age $h < T$; that is, for all $A \in \mathcal{F}$,

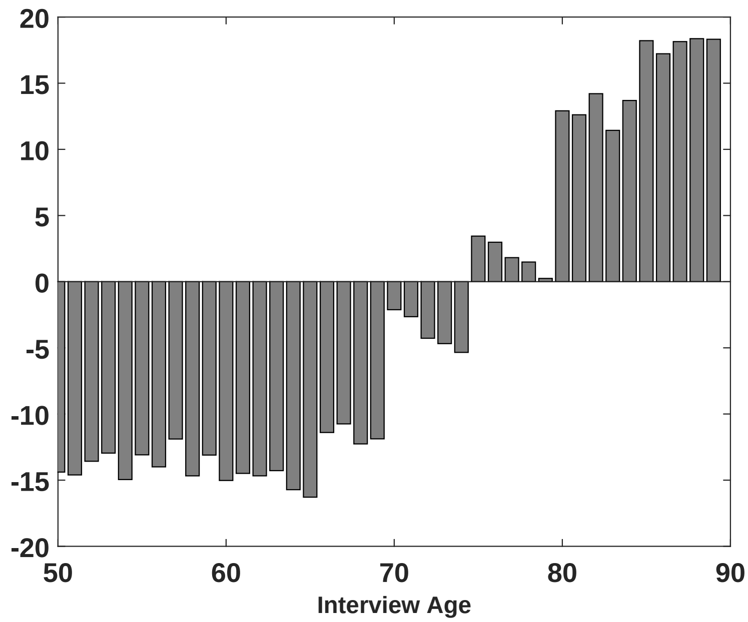
$$\psi^h(A) \equiv \frac{\psi(A \cap (D_h \cup \cdots \cup D_T))}{\psi(D_h \cup \cdots \cup D_T)}.$$

As a notational convention, we write for the objective probability that an agent of

[6] Trends in life-expectancy are taking into account. A detailed description of the data is contained in Ludwig and Zimper (2013), and Groneck, Ludwig, and Zimper (2016).

[7] This pattern is robust to more elaborate measures of objective survival probabilities. For example, in Grevenbrock et al. (2016) we compute objective survival rates at the individual level instead of using cohort life tables. Averaging differences between subjective beliefs and objective probabilities across individuals by age gives rise to similar differences as those shown in Figure 1.

Figure 1: Difference of Subjective Survival Beliefs and Cohort Data



Notes: Deviations in percentage points of subjective survival probabilities from objective data based on cohort life tables. Future objective data is predicted with the Lee-Carter procedure (Lee and Carter 1992). Each bar depicts the difference of unconditional probabilities to survive to a specific target age.
Source: Groneck, Ludwig, and Zimper (2016).

current age h survives until (at least) the beginning of age $t > h$

$$\psi_{h,t} \equiv \psi^h \left(D_t \cup \cdots \cup D_T\right) = \sum_{k=t}^{T} \psi^h \left(D_k\right). \tag{1}$$

Definition 1 *Fix parameters $\delta \in [0,1]$ and $\lambda \in [0,1]$. A neo-additive probability weighting function $\nu : [0,1] \to [0,1]$ satisfies, for all h and all $A \in \mathcal{F}$,*

$$\nu\left(\psi^h(A)\right) = \begin{cases} 0 & \textit{if } \psi^h(A) = 0 \\ \delta\lambda + (1-\delta)\psi^h(A) & \textit{if } \psi^h(A) \in (0,1) \\ 1 & \textit{if } \psi^h(A) = 1. \end{cases}$$

As notational convention we write $\nu_{h,t} \equiv \delta\lambda + (1-\delta)\psi_{h,t}$ for all $\psi_{h,t} \in (0,1)$.

The first parameter $\delta \in [0,1]$ measures the *deviation* of the neo-additive belief from the objective probability. One possible (cognitive) interpretation is that δ captures the empirical phenomenon of *likelihood-insensitivity* (cf. Wakker 2004). The second parameter $\lambda \in [0,1]$ determines in how far the decision maker over- vs. underestimates objective probabilities whenever $\delta > 0$. Since $\delta\lambda$ (resp. $\delta(1-\lambda)$) corresponds in our life-cycle model to the additional decision weight attached to the event in which the decision maker lives until a maximally possible age (resp. already dies at the end of the current period), we henceforth refer to λ as *optimism* parameter.

In the present paper, we choose neo-additive weighting functions because (i) they naturally give rise to the age-dependent survival biases depicted in Figure 1 whereby (ii) the mathematical formalism that describes such biases remains as parsimonious as possible. To see this, observe that underestimation of survival chances at a young age s is captured through a neo-additively transformed probability whenever we have that

$$\delta\lambda + (1-\delta)\,\psi_{s,s+1} < \psi_{s,s+1}. \tag{2}$$

Conversely, overestimation of survival chances at an old age $t > s$ corresponds to

$$\psi_{t,t+1} < \delta\lambda + (1-\delta)\,\psi_{t,t+1}. \tag{3}$$

Throughout this paper we consider the empirically relevant case of monotonically strictly decreasing conditional survival probabilities, such that $\psi_{t,t+1} < \psi_{s,s+1}$ whenever $t > s$. Neo-additively transformed survival probabilities thus generate underestimation at young age s combined with overestimation at old age t if, and only if, $\delta > 0$ whereby λ must satisfy

$$\psi_{t,t+1} < \lambda < \psi_{s,s+1}. \tag{4}$$

Based on the parsimonious characterization of empirically observed age-dependent survival belief biases (4), we investigate in the remainder of the paper the relationship between underestimation of survival chances and undersaving at young ages, on the one hand, and overestimation of survival chances and oversaving at old ages, on the other hand.

3 Rank Dependent Utility (RDU) Life-Cycle Model

3.1 RDU Defined over Risky Consumption Streams

Recall the general definition of rank dependent utility.[8] Denote by

$$\mathbf{x} \equiv (\psi_1 : x_1, \cdots, \psi_n : x_n)$$

a *risky prospect* where ψ_k stands for the objective probability of deterministic outcome x_k. Under the assumption that the decision maker has the following preference ranking over the deterministic outcomes

$$x_n \succeq \ldots \succeq x_1,$$

the RDU of prospect $\mathbf{x}$ is defined as the following *Choquet integral*

$$RDU(\mathbf{x}, \omega) \equiv \sum_{k=0}^{n-1} U(x_{n-k}) \left[\omega\left(\psi_n + \ldots + \psi_{n-k}\right) - \omega\left(\psi_n + \ldots + \psi_{n-k+1}\right)\right]$$

where $U(\cdot)$ is a continuous utility function and $\omega : [0,1] \to [0,1]$ is a probability weighting function that is increasing and satisfies $\omega(0) = 0$ and $\omega(1) = 1$.[9]

Turn now to our RDU life-cycle model for which risk exclusively arises from the decision maker's mortality risk. Hence, the decision maker's life-cycle consumption is risky because she does not know for how long she is going to live (i.e., to consume). For a given age h, fix the consumption stream

$$(c_h, c_{h+1}, \ldots, c_T) \in \mathbb{R}_+^{T-h+1} \tag{5}$$

whereby we assume that the c_k, $k \in \{0, \ldots, T\}$, of all consumption streams are bounded away from zero. That is, we restrict attention to consumption streams for which there exists an arbitrarily small but fixed $\varepsilon > 0$ such that $c_k \geq \varepsilon$ for all $k \in \{0, \ldots, T\}$.[10] Denote by $\mathbf{c}^t \equiv (c_h, c_{h+1}, \ldots, c_t)$ the truncation of (5) that contains

[8]For early formulations of RDU see Quiggin (1981), Quiggin (1982), Yaari (1987), and Chateauneuf (1999). Note that RDU is identical to *cumulative propect theory* (CPT) whenver (i) CPT is restricted to the domain of gains and (ii) the CPT decision maker faces objective probabilities (cf., Tversky and Kahneman 1992).

[9]The standard convention $\omega\left(\psi_n + \psi_{n+1}\right) = 0$ applies.

[10]By this restriction, we ensure that the values of our chosen CRRA period utility function remain finite and do not approach minus infinity for $c_k \to 0$. (See our discussion of the role of the preference

all consumption entries up to age $t \geq h$. According to our interpretation, $\mathbf{c}^t$ is the deterministic outcome that the agent receives if she consumes in accordance with (5) and dies at the end of age t. We call

$$\mathbf{c} \equiv \left(\psi^h\left(D_h\right) : \mathbf{c}^h, \ldots, \psi^h\left(D_T\right) : \mathbf{c}^T\right) \tag{6}$$

the h-old agent's *consumption prospect* if she consumes in accordance with (5). We further assume the following preference ranking over the deterministic outcomes of (6):

$$\mathbf{c}^T \succeq \cdots \succeq \mathbf{c}^h,$$

that is, we restrict attention to a decision maker who prefers for any given consumption stream (5) to live (i.e., to consume) longer.

Definition 2 *The RDU of consumption prospect* $\mathbf{c}$ *of an* h*-old agent with respect to the probability weighting function* ω *is given as follows:*

$$\begin{aligned} & RDU^h\left(\mathbf{c}, \omega\right) \qquad (7) \\ & \equiv \sum_{t=0}^{T-h} U\left(\mathbf{c}^{T-t}\right)\left[\omega\left(\psi^h\left(D_T\right) + \ldots + \psi^h\left(D_{T-t}\right)\right) - \omega\left(\psi^h\left(D_T\right) + \ldots + \psi^h\left(D_{T-t+1}\right)\right)\right] \end{aligned}$$

where $U\left(\cdot\right) \in \mathbb{R}_+$ *denotes a continuous utility function satisfying*

$$U\left(\mathbf{c}^T\right) \geq \ldots \geq U\left(\mathbf{c}^h\right). \tag{8}$$

Assumption 1 *The utility of any truncated consumption stream* $\mathbf{c}^t$ *is additively separable with exponential discount factor, i.e.,*

$$U\left(\mathbf{c}^t\right) = \sum_{s=h}^{t} \beta^{s-h} u\left(c_s\right) \tag{9}$$

where $\beta \in (0, 1]$ *is the pure time-preference discount factor and* $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ *is a strictly increasing and strictly concave per period utility function.*[11]

shifter in Assumption 4). This restriction will be without loss of generality when we determine the (interior) solution to the optimal consumption stream as the optimal consumption will be strictly greater zero in all periods.

[11]Note that we require of the per period utility function $u(c_s) \geq 0$ for all s to ensure that condition (8) is satisfied.

Using the notational convention (1), we can transform (7) under Assumption 1:

$$\begin{aligned} RDU^h(\mathbf{c},\omega) &= \sum_{t=0}^{T-h} U\left(\mathbf{c}^{T-t}\right)\left[\omega\left(\psi_{h,T-t}\right)-\omega\left(\psi_{h,T-t+1}\right)\right] \\ &= \sum_{t=0}^{T-h}\sum_{s=h}^{T-t} \beta^{s-h} u\left(c_s\right)\left[\omega\left(\psi_{h,T-t}\right)-\omega\left(\psi_{h,T-t+1}\right)\right] \\ &= u\left(c_h\right)+\sum_{t=h+1}^{T}\omega\left(\psi_{h,t}\right)\beta^{t-h}u\left(c_t\right). \end{aligned} \tag{10}$$

Assumption 2 *The probability weighting function $\omega(\cdot)$ is given as some neo-additive probability weighting function $\nu(\cdot)$.*

Substituting a neo-additive probability weighting function $\nu(\cdot)$ for $\omega(\cdot)$ in (10) immediately gives us the following characterization of the RDU defined over risky consumption streams.

Theorem 1 *Under Assumptions 1 and 2, the RDU (7) of consumption prospect c of an h-old agent becomes*

$$\begin{aligned} RDU^h(\mathbf{c},\nu) &= u(c_h)+\sum_{t=h+1}^{T}\nu_{h,t}\beta^{t-h}u\left(c_t\right) \\ &= u(c_h)+\sum_{t=h+1}^{T}\left(\delta\lambda+(1-\delta)\psi_{h,t}\right)\beta^{t-h}u\left(c_t\right). \end{aligned} \tag{11}$$

We henceforth refer to (11) for $h\in\{0,\ldots,T\}$ as *the RDU life-cycle* model without explicitly mentioning Assumptions 1 and 2. For the parametric special case $\delta=0$, (11) reduces to the standard RE life-cycle model

$$U^h(\mathbf{c},\varphi)=u(c_h)+\sum_{t=h+1}^{T}\psi_{h,t}\beta^{t-h}u\left(c_t\right), \tag{12}$$

which will serve as our natural reference model.

Remark 1 *Without Assumption 1 but under Assumption 2 the RDU (7) of consumption prospect c of an h-old agent with respect to the neo-additive probability weighting* *function ν would, more generally than (11), take on the following structural form (cf. the proof of Proposition 2 in Groneck, Ludwig, and Zimper (2016))*

$$\delta\left(\lambda U\left(\mathbf{c}^{T}\right)+(1-\lambda)\,U\left(\mathbf{c}^{h}\right)\right)+(1-\delta)\sum_{t=0}^{T-h}U\left(\mathbf{c}^{T-t}\right)\psi^{h}\left(D_{T-t}\right). \tag{13}$$

Note that the optimism parameter λ measures in how far any deviation $\delta>0$ from the rational expectations approach is resolved in a rather optimistic way: large values of λ put large decision weights on the best possible outcome $U\left(\mathbf{c}^{T}\right)$ according to which the agent will achieve her maximally possible age T. In contrast, small values of λ put large decision weights on the worst possible outcome $U\left(\mathbf{c}^{h}\right)$ according to which the agent will already die at the end of the current period.

Remark 2 *An application of Cohen (1992)'s security and potential level model to consumption prospects becomes under Assumption 2*

$$a\left(\mathbf{c}^{h},\mathbf{c}^{T}\right)\sum_{t=0}^{T-h}U\left(\mathbf{c}^{T-t}\right)\psi^{h}\left(D_{T-t}\right)+b\left(\mathbf{c}^{h},\mathbf{c}^{T}\right) \tag{14}$$

where $a\left(\mathbf{c}^{h},\mathbf{c}^{T}\right)$ and $b\left(\mathbf{c}^{h},\mathbf{c}^{T}\right)$ are real-valued functions in (i) the worst possible consumption stream $\mathbf{c}^{h}$ (=security level) and (ii) in the best possible consumption stream $\mathbf{c}^{T}$ (=potential level).

Setting $a\left(\mathbf{c}^{h},\mathbf{c}^{T}\right)\equiv(1-\delta)$ and $b\left(\mathbf{c}^{h},\mathbf{c}^{T}\right)\equiv\delta\left(\lambda U\left(\mathbf{c}^{T}\right)+(1-\lambda)\,U\left(\mathbf{c}^{h}\right)\right)$ shows that the RDU model with neo-additive probability weighting (13) can be regarded a as special case of Cohen's security and potential level model (which furthermore satisfies stochastic dominance, cf. Proposition 5 in Cohen (1992), as well as continuity).[12]

Remark 3 *Our RDU life-cycle model is related to similar models in Bleichrodt and Eeckhoudt (2006), Halevy (2008) and Drouhin (2015). The RDU representation (10) for a 0-old agent already appears in Bleichrodt and Eeckhoudt (2006) (cf. Equation (8)*

[12]By a similar argument, the RDU model (13) becomes for an extremely pessimistic RDU decision maker with $\lambda=0$ a formal special case of the *security level models* considered in Gilboa (1988) and Jaffray (1988) for which only the worst possible outcome of a risky prospect, here $\mathbf{c}^{h}$, matters in addition to the prospect's expected utility. See also the discussion in Section 3.2 which establishes the formal equivalence between a RDU life-cycle model with $\lambda=0$ and a life-cycle model with quasi-hyperbolic time discounting.

therein).[13] *More restrictively than these authors, Halevy (2008) derives the following RDU representation for a 0-old agent (Theorem 2 in Halevy 2008)*

$$RDU^{0}\left(\mathbf{c},\omega\right)=u(c_{0})+\sum_{t=1}^{T}\omega\left(\left(1-\kappa\right)^{t}\right)\beta^{t}u\left(c_{t}\right), \tag{15}$$

where $\kappa \in (0,1)$ denotes a constant hazard rate. Denote by $\bar{\psi}_{0,t}=(1-\kappa)^{t}$ the additive survival probability from age $h=0$ to $t\geq 1$ to see that Halevy (2008)'s representation is a special case of (10) under the assumption that objective survival probabilities exhibit a constant hazard rate, i.e.,

$$\psi_{h,h+1}=1-\kappa \tag{16}$$

for all ages h. The constant hazard rate assumption (16) is, however, not supported by the data. Also, Halevy (2008)'s representation (15) cannot express underestimation at a young age combined with overestimation at an old age; (to see this note that (4) holds under (16) with equality).[14]

3.2 Discussion: Quasi-Hyperbolic Time Discounting

Recall that *hyperbolic discounting* models (cf. Strotz (1955) and Pollak (1968)) describe agents who are more sensitive to a given time delay if it occurs closer to the present than if it occurs farther in the future. The weaker concept of *quasi-hyperbolic discounting (QHD)*, first proposed by Phelps and Pollak (1968) and made very popular in the behavioral economics literature by Laibson (1997), assumes that agents are sensitive to a time delay with respect to the present period only. More specifically,

[13] Bleichrodt and Eeckhoudt (2006) also solve the RDU life cycle model for the optimal consumption stream. As a drawback of their solution, however, they ignore the dynamic inconsistency of (10) for $h=0,\ldots,T$. Therefore, their Theorem 5 is based on the *planned* consumption stream of a naive RDU agent which does (typically) not coincide with the *actual* consumption behavior whenever the RDU agent does not reduce to a RE agent (cf. our detailed discussion in Section 4).

[14] Halevy writes: "Although the constant probability of stopping (hazard) assumption is maintained throughout this paper, it can be easily relaxed—resulting in a more general (but less tractable) representation." (Footnote 2 in Halevy (2008)). The advantage of our neo-additive RDU life-cycle model (11) is exactly its highly tractable representation without any need to impose restrictions on additive survival probabilities.

define the QHD discount factors

$$\left\{1, \gamma\beta, \gamma\beta^2, \gamma\beta^3, \ldots \gamma\beta^T\right\}, \tag{17}$$

where $0 < \gamma, \beta < 1$, where γ denotes the short term discount factor.

Consider at first a QHD decision maker whose survival beliefs coincide with objective survival probabilities. By Assumption 1, the utility of consumption plan $\mathbf{c}$ becomes for this QHD decision maker of age h

$$U^h_{QHD}(\mathbf{c}, \psi) = u(c_h) + \sum_{t=h+1}^{T} \psi_{h,t}\gamma\beta^{t-h} u(c_t). \tag{18}$$

Obviously, the QHD utility representation (18) is formally equivalent to the special case of our RDU representation (11) for $\lambda = 0$ and $1 - \delta = \gamma$, i.e.,

$$\nu_{h,t} = \gamma\psi_{h,t}. \tag{19}$$

The life-cycle model of a RDU decision maker who is extremely pessimistic, i.e., $\lambda = 0$, about her survival chances can thus not be distinguished from the life-cycle model of a QHD decision maker who holds correct survival beliefs but discounts the future in accordance with (17).

Next, consider the RDU life-cycle model (11) with parametrization $\delta = 1$ and $\lambda \in (0, 1)$ implying

$$\nu_{h,t} = \lambda. \tag{20}$$

Setting $\lambda = \gamma$ shows that this RDU life-cycle model is formally equivalent to a QHD model without mortality risk

$$U^h_{QHD}(\mathbf{c}) = u(c_h) + \sum_{t=h+1}^{T} \gamma\beta^{t-h} u(c_t). \tag{21}$$

That is, life-cycle models of RDU decision makers with extreme likelihood insensitivity, i.e., $\delta = 1$, cannot be distinguished from deterministic QHD life-cycle models and vice versa.

Finally, notice that the RDU life-cycle model (11) collapses for the degenerate parametrization

$$\delta = \lambda = 1 \tag{22}$$

to the standard deterministic life-cycle model with exponential discounting. Since this well-understood deterministic special case is of no interest to us, we exclude the parametrization (22) from the remainder of our analysis.

Remark 4 *There exists an interesting discussion in the theoretical and experimental literature about whether hyperbolic time discounting is conceptually nothing else than rank dependent decision making under an uncertain stopping time as, e.g., represented by mortality risks (cf., e.g., Halevy (2008) and Epper et al. (2011) as well as references therein). Although QHD life-cycle models can, by (18) and (21), be formally subsumed under our RDU life-cycle model, we do not claim that hyperbolic time discounting reduces in general to rank dependent decision making.*

Remark 5 *Interpreted in terms of Halevy (2015)'s properties of time preferences, RDU preferences are non-stationary as well as time-inconsistent whenever they do not reduce to expected utility preferences. Moreover, RDU preferences are time-invariant if, and only if, the neo-additive beliefs satisfy*

$$\begin{aligned} \frac{\nu_{0,3}}{\nu_{0,2}} &= \frac{\nu_{1,3}}{\nu_{1,2}} \\ &\Leftrightarrow \\ \frac{\lambda\delta + (1-\delta)\,\psi_{0,3}}{\lambda\delta + (1-\delta)\,\psi_{0,2}} &= \frac{\lambda\delta + (1-\delta)\,\psi_{1,3}}{\lambda\delta + (1-\delta)\,\psi_{1,2}}. \end{aligned} \tag{23}$$

But (23) only holds for $\delta = 0$, for $\lambda = 0$, or for $\delta = 1$. That is, our RDU-life cycle model becomes time-invariant if, and only if, it reduces to either the RE- or the QHD model. Halevy (2015) reports experimental data according to which time-variance is more common among decision makers with non-stationary and time-inconsistent preferences than time-invariance. We interpret Halevy's finding as empirical evidence in favour of a general time-variant RDU life-cycle model over the nested special case of a time-invariant QHD life-cycle model.

4 Solving the Model

4.1 Additional Assumptions

We assume that there exists an initial amount of total wealth that can be spend over the life-cycle. There is no borrowing constraint in the model so that total wealth is the sum of financial wealth, current period labor income and human capital wealth, which is the discounted value of current and future labor income.

Assumption 3 *The budget constraint is given by*

$$w_{t+1} = (w_t - c_t)\, R \quad \textit{for } t \in \{0, 1, \ldots, T-1\}\,. \tag{24}$$

for initial wealth $w_0 > 0$ *and market return* $R \geq 1$*. In addition the standard no-Ponzi condition applies, hence* $w_{T+1} \geq 0$*.*

For analytical convenience, we restrict attention to the family of power utility functions with a *constant inter-temporal elasticity of substitution* (IES), also known as *constant relative risk aversion* (CRRA) utility functions.[15]

Assumption 4 *The period-utility function for an IES of* $\frac{1}{\theta}$ *is given as*

$$u(c) = \chi + \begin{cases} \frac{c^{1-\theta}}{1-\theta} & \textit{for } \theta \neq 1, \theta > 0 \\ \ln(c) & \textit{for } \theta = 1. \end{cases} \qquad \chi \geq 0 \tag{25}$$

The additive preference shifter χ is irrelevant for the standard EU framework with additive beliefs. However, for our RDU framework a value of χ needs to be chosen such that the period utility function $u(c)$ is always non-negative. Else, condition (8) would break down to the effect that, contrary to our interpretation, the consumer might prefer dying earlier than living longer so that the RDU representation of Theorem 1 would no longer apply. For $\theta < 1$ the per period utility function is positive so that χ can be set to zero. For $\theta \geq 1$, χ must take on a sufficiently large value.[16]

[15] We prefer the interpretation of $\frac{1}{\theta}$ as the IES over the interpretation of θ as a measure of risk aversion (which is the coefficient of relative risk aversion in an atemporal context). While risk-aversion is comprehensively characterized by the concavity of the vNM utility function for an expected utility decision maker, this is no longer the case for a RDU decision maker because probability weighting additionally reflects risk attitudes, cf. Hong, Karni, and Safra (1987),Wu and Gonzalez (1996),Chateauneuf and Coen (2000).

[16] Given that our programming problem is convex (see below), the additive preference shifter does

4.2 Characterization of the Solution for Three Types

Under the above assumptions we derive analytical solutions for consumption Euler equations (consumption growth rates) and consumption policy functions for three agent types, indexed by $i \in \{r, n, s\}$, which we subsequently interpret in Section 4.3.[17] As our benchmark case, we consider RE agents (indexed by r) who base their decisions on objective survival beliefs ψ_t. In contrast to RE agents, (non-degenerate) RDU agents exhibit dynamically inconsistent preferences. That is, from an ex ante perspective a RDU agent would prefer a different choice of future consumption than her future self who actually makes this choice.

To model how RDU agents cope with this dynamic inconsistency, we distinguish between *naive* and *sophisticated* agents. Naive RDU agents (indexed by n) are, and remain, completely unaware of their dynamically inconsistent preferences. For them planned and realized consumption streams will therefore (in general) differ except for the current period. Sophisticated RDU agents (indexed by s) fully understand the diverging preferences of their future selves. Moreover, they correctly anticipate how their current consumption choice will impact on the consumption behavior of their future selves. In contrast to naive RDU agents, sophisticated RDU agents thus want to influence future consumption behavior in their favor through the impact of their current consumption choice on future budget constraints.

In what follows, we denote by c_t^i the actual period t consumption behavior of an agent of type $i \in \{r, n, s\}$. To keep track of the difference between planned and actual future consumption of the naive agent, we write $c_t^{n,h}$ for the planned period $t \geq h$ consumption of the naive agent from her current perspective at age h whereby $c_h^{n,h} = c_h^n$ is the actual consumption of the naive agent at age h.

Theorem 2 *The consumption policy functions of all agents of type $i \in \{r, n, s\}$ and age $h \in \{0, \ldots, T\}$ are linear in total wealth, i.e., $c_h^i = m_h^i w_h$, whereby the marginal propensity to consume (MPC) is given at the final age as $m_T^i = 1$. For all other ages*

not affect the solution. Upon characterizing the solution one can therefore determine ex-post from the optimal consumption stream some value for χ such that condition (8) holds. More precisely, for $\theta \geq 1$ we can pick any $\varepsilon > 0$ that is strictly smaller than all optimal period consumption levels and set $\chi = -\frac{\varepsilon^{1-\theta}}{1-\theta}$.

[17]In Groneck, Ludwig, and Zimper (2016) we characterize the solution for a stochastic model and a general concave utility function using a slightly different representation of survival beliefs. For convenience of the reader we restate this solution in Appendix B.1 for the RDU model of the present paper.

$h < T$ we obtain the following MPCs recursively:

1. *for the RE agent*
$$m_h^r = \frac{1}{1 + \frac{\left(\beta R^{1-\theta} \psi_{h,h+1}\right)^{\frac{1}{\theta}}}{m_{h+1}^r}}; \tag{26}$$

2. *for the planned*[18] *consumption of the naive RDU agent for all $t \geq h$*
$$m_t^{n,h} = \frac{1}{1 + \frac{\left(\beta R^{1-\theta} \frac{\nu_{h,t+1}}{\nu_{h,t}}\right)^{\frac{1}{\theta}}}{m_{t+1}^{n,h}}}; \tag{27}$$

3. *for the sophisticated RDU agent*
$$m_h^s = \frac{1}{1 + \frac{\left(\beta R^{1-\theta} \nu_{h,h+1} \left(\Theta_{h+1} + \xi_{h+1}\right)\right)^{\frac{1}{\theta}}}{m_{h+1}^s}} \tag{28}$$

   where
$$\Theta_{h+1} \equiv m_{h+1}^s + \frac{\nu_{h,h+2}}{\nu_{h,h+1} \cdot \nu_{h+1,h+2}} \left(1 - m_{h+1}^s\right) \geq 1, \tag{29}$$
$$\xi_{h+1} \equiv \beta R^{1-\theta} \frac{\nu_{h,h+2}}{\nu_{h,h+1}} \left(1 - m_{h+1}^s\right)^{1-\theta} m_{h+1}^{s\,\theta} \left(\zeta_{h+2}^h - \zeta_{h+2}^{h+1}\right) \geq 0 \tag{30}$$
$$\zeta_t^h = \begin{cases} 1 & \textit{for } t = T \\ m_t^{s^{1-\theta}} + \beta \frac{\nu_{h,t+1}}{\nu_{h,t}} (1 - m_t^s)^{1-\theta} \zeta_{t+1}^h & \textit{otherwise.} \end{cases} \tag{31}$$

The analytical characterization for RE and naive RDU agents directly follows from the analysis in Samuelson (1969) which is based on the insight that the combination of multiplicative dynamic budget constraints of the form in (24) with homothetic utility functions gives rise to linear policy functions of consumption. While the solution for the sophisticated RDU agent's problem is slightly more involved, the closed form expressions stated in Theorem 2 are also due to these properties.

Before we give a detailed interpretation of the consumption behavior of the different agent types, let us state the intertemporal Euler equations corresponding to the MPCs derived in Theorem 2.

[18]Recall that planned consumption of the naive type coincides with actual consumption for $t = h$.

Corollary 1 *The consumption growth rates are, for all $h < T$,*

1. *for the RE agent*
$$\frac{c^r_{h+1}}{c^r_h} = \left(\beta\psi_{h,h+1}R\right)^{\frac{1}{\theta}}; \tag{32}$$

2. *for the planned consumption of the naive RDU agent for all $t \geq h$*
$$\frac{c^{n,h}_{t+1}}{c^{n,h}_t} = \left(\beta\frac{\nu_{h,t+1}}{\nu_{h,t}}R\right)^{\frac{1}{\theta}}; \tag{33}$$

3. *for the sophisticated RDU agent*
$$\frac{c^s_{h+1}}{c^s_h} = \left(\beta\nu_{h,h+1}R\left(\Theta_{h+1} + \xi_{h+1}\right)\right)^{\frac{1}{\theta}}. \tag{34}$$

4.3 Interpretation

4.3.1 The Naive Agent

Whenever our RDU life-cycle model (11) does not reduce to the RE model (12), i.e., whenever $\delta > 0$, the life-cycle utility maximization problem becomes dynamically inconsistent. This can be seen from the discrepancy between the h-old naive agent's plan for future consumption and her actual future consumption path. The planned consumption path, represented by the intertemporal Euler equation (33), would maximize the h-old RDU agent's life-cycle utility regardless of whether she is naive or sophisticated. That the naive agent's actual future consumption path will be deviating from this plan is an expression of the fact that the RDU agent's future selves have different preferences about future consumption than the h-old agent.

To be more specific, observe that the first-order condition for period $h+1$ from the perspective of period h must, by equation (33), satisfy

$$\frac{c^{n,h}_{h+2}}{c^{n,h}_{h+1}} = \left(\beta R\frac{\nu_{h,h+2}}{\nu_{h,h+1}}\right)^{\frac{1}{\theta}}.$$

When the naive agent turns age $h+1$, the corresponding condition from the perspective of period $h+1$ becomes $\frac{c^{n,h+1}_{h+2}}{c^{n,h+1}_{h+1}} = (\beta R\nu_{h+1,h+2})^{\frac{1}{\theta}}$. Dynamic consistency requires that the marginal valuation of consumption in the two subsequent periods for agents of age h and $h+1$ must be the same, i.e.,

$$\frac{\nu_{h,h+2}}{\nu_{h,h+1}} = \nu_{h+1,h+2} \qquad \Leftrightarrow \qquad \frac{\delta\lambda + (1-\delta)\,\psi_{h,h+2}}{\delta\lambda + (1-\delta)\,\psi_{h,h+1}} = \delta\lambda + (1-\delta)\,\psi_{h+1,h+2}.$$

However, this is (generically) only the case for $\delta = 0$ because $\frac{\psi_{h,h+2}}{\psi_{h,h+1}} = \psi_{h+1,h+2}$.

Using the closed form solutions, we can characterize the discrepancy between the optimal plan for period $h+1$ consumption from the h-old agent's perspective and the actual consumption in period $h+1$ as follows:

Proposition 1 *The naive RDU agent consumes more in period $h+1$ than originally planned at age h, i.e.,*

$$m^{n,h}_{h+1} < m^{n,h+1}_{h+1} = m^{n}_{h+1} \qquad \Leftrightarrow \qquad c^{n,h}_{h+1} < c^{n,h+1}_{h+1} = c^{n}_{h+1}.$$

The proof of the proposition proceeds in two steps. First, we derive an important sufficient condition for which the naive RDU agent ends up saving less (consuming more) at age $h+1$ than she planned to save at age h. Since we relate back to this sufficient condition we state it explicitly:

$$\frac{\nu_{h,h+2}}{\nu_{h,h+1}\nu_{h+1,h+2}} > 1. \tag{35}$$

In a next step, we establish that condition (35) always holds in our non-degenerate (i.e., $\delta > 0$) RDU model.

Remark 6 *Importantly, condition (35) also holds in the QHD model. As we established in the discussion of Section 3.2, our RDU model embeds the QHD model as the special parameterizations $\delta \in (0,1), \lambda = 0$ as well as $\delta = 1, \lambda \in (0,1)$. In the first case, condition (35) is equal to $\frac{1}{1-\delta} > 1$, in the second it is equal to $\frac{1}{\lambda} > 1$. Therefore, in terms of the revisal of the naive RDU agent's plan, our model qualitatively shares the same dynamics as the QHD model.*

4.3.2 The Sophisticated Agent

Equation (28) is the solution to the sophisticated agent's problem who understands her dynamically inconsistent preferences. To interpret the solution for the sophisticated agent, again recall from the discussion in Section 3.2 that our RDU model becomes formally equivalent to variants of QHD models for the degenerate parameterizations $\lambda = 0, \delta \in (0,1)$ and $\lambda \in (0,1), \delta = 1$. In this case, $\xi_h = 0$ for all t and Θ_{h+1} is the adjustment term in the Euler equation. The latter is familiar from the QHD literature, cf. Harris and Laibson (2001), who label the adjusted Euler equation as *generalized Euler equation*. Formally, Θ_{h+1} shows up because the standard Envelope condition seizes to hold in a dynamic programming problem under dynamic inconsistency. Because of (35), $\Theta_{t+1} \geq 1$ as in the QHD model. The term reflects the sophisticated RDU agent's high marginal valuation of savings, $\left(1 - m^s_{h+1}\right)$, and self h correspondingly expresses higher patience than according to the pure short-run discount factor $\beta\nu_{h,h+1}$. We label this the *intertemporal smoothing motive* of the sophisticated agent to the effect that this motive implies higher savings compared relative to the naive RDU agent.

Yet, Θ_{h+1} is inversely related to next period's marginal propensity to consume (MPC), m^s_{h+1}. If the marginal propensity of ones own future self will increase, the sophisticated RDU agent decreases her consumption growth rate in the current period, i.e., decreases savings thereby leaving fewer resources to her own future self. We label this the *constrain-ones-future-self motive.* Notice that the two motives work in opposite directions.

Our general case with $\lambda \in (0,1), \delta \in (0,1)$ implies that an additional *adjustment factor* shows up in the sophisticated agent's first-order condition, cf. Appendix B.1. The reason for the appearance of this adjustment factor is that the continuation values from the perspectives of an agent's self h and her future self $h+1$ from periods $h+2$ onwards generally differ in our RDU model whereas they are identical in the QHD model.[19] This adjustment factor reflects the difference in the marginal value of wealth from period $t+2$ onwards between selves t and $t+1$. In Theorem 2 this difference in the marginal value from wealth between selves t and $t{+}1$ is reflected in term ξ_{+1} which in turn involves the distance $\zeta^h_{h+2} - \zeta^{h+1}_{h+2}$, cf. equation (30). Our proof of Theorem 2 establishes that this distance measures the difference in current self h and future self's $h+1$ marginal valuation of wealth in period $h+2$. Because $\zeta^h_{h+2} - \zeta^{h+1}_{h+2} \geq 0$, we have that self h's marginal valuation of wealth in period $h+2$ is higher such that

[19] Discounting in the QHD model is geometric after age $h{+}1$ for the agent of age h and after age $h{+}2$ for the agent of age $h+1$ so that they both apply the same discount function after age $h+2$. By only looking at the next period, *quasi* hyperbolic discounting models stand for a shortcut of *proper* hyperbolic discounting models. Our RDU model shares with such proper hyperbolic discounting models that discounting continues to be non-geometric for all future periods.

that self h values savings from $h+1$ to $h+2$ more than future self $h+1$. Hence, savings are increased already at age h for the sophisticated agents.

4.3.3 Naive versus Sophisticated Agent

To compare the sophisticated agent to the naive agent, recall that for $t = h+1$ we have $\frac{\nu_{h,t+1}}{\nu_{h,t}} = \nu_{h,h+1}$ so that the difference between planned consumption behavior of a naive and actual consumption behavior of a sophisticated agent is due to $\Theta_{h+1}+\xi_{h+1} \geq 1$. This implies the following result:

Proposition 2 *The sophisticated RDU agent's actual consumption growth rate exceeds the planned consumption growth rate of the naive RDU agent at age h, i.e.,*

$$\frac{c^s_{h+1}}{c^s_h} \geq \frac{c^{n,h}_{h+1}}{c^n_h}. \tag{36}$$

with the inequality being strict in all periods $h < T-2$ for which $\xi_{h+1} > 0$ if $\delta \in (0,1), \lambda \in (0,1)$.

On the one hand, Proposition 2 establishes that actual consumption growth of the sophisticated RDU agent generally exceeds planned consumption growth of the naive RDU agent. On the other hand, Proposition 1 establishes that actual future consumption of the naive RDU exceeds her planned future consumption, implying for actual consumption growth of the two agents

$$\frac{c^s_{h+1}}{c^s_h} \gtreqless \frac{c^{n}_{h+1}}{c^n_h}, \tag{37}$$

i.e., the consumption growth rate of the sophisticated agents can be smaller or larger than consumption growth of the naive agent—and will be equal for a special case discussed in the next section. The interplay between the opposing forces is best understood by analyzing the consumption behavior in the first and second to last periods, periods $T-1$ and $T-2$. For period $T-1$ observe that $\Theta_T = 1$ (as well as $\xi_{T+1} = 0$) as well as $m^s_T = m^n_T = 1$. Hence, the consumption behavior of naive and sophisticated agents coincide. In period $T-2$ we have $\Theta_{T-1} > 1$ (and $\xi_{T-1} = 0$). Due to the *inter-temporal smoothing motive* consumption at age $T-2$ of the sophisticated RDU agent is lower relative to the naive agent, i.e. the sophisticated agent saves more. Contrary, the sophisticated agent anticipates the (excessive) consumption behavior of her own future self whereas the naive RDU agent does not, implying that $m^s_{T-1} = m^{n,T-1}_{T-1} > m^{n,T-2}_{T-1}$ (cf. Proposition 1). This induces the *constrain ones future self motive* of the sophisticated agents according to which consumption at age $T-2$ of the sophisticated is higher relative to the naive RDU agent. The sophisticated agent anticipates the overconsumption of her own future self and accordingly shifts fewer resources to the future by consuming more in the present. Also notice that the two effects interact: An increase of the MPC at age $T-1$, $m^s_{T-1} = m^{n,T-1}_{T-1}$, decreases Θ_{T-1} thereby reducing the inter-temporal smoothing effect. As will be discussed in Section 5.3.2, these observations imply that sophistication might even lead to lower savings if the *constrain ones future self motive* dominates the *inter-temporal smoothing motive.*

Remark 7 *Also with respect to the comparison between the naive and the sophisticated agent, our model qualitatively shares the same dynamics as the QHD model.*

4.4 Characterizing MPC Through Discount Functions: When Sophisticates act Naive

We now present an alternative non-recursive characterization of these MPCs in terms of *discount functions.* Since this characterization is central to our subsequent analysis of young-age undersaving, we give it the status of a theorem.

Theorem 3 *Let $\rho^i(h,t)$ be the discount function for agent $i \in \{r,n,s\}$ where*

$$\begin{aligned} \rho^r(h,t) &= \beta^{t-h}\psi_{h,t}, \\ \rho^n(h,t) &= \rho^s(h,t) = \beta^{t-h}\nu_{h,t}. \end{aligned}$$

Then the MPCs of Theorem 2 can be equivalently expressed as follows:

1. *for the* RE *agent:*

$$m^r_h = \begin{cases} 1 & \textit{for } h = T \\ \frac{1}{1+\left(\sum_{t=h+1}^{T}\rho^r(h,t)\left(R^{t-h}m^r_t\prod_{j=h+1}^{t-1}\left(1-m^r_j\right)\right)^{1-\theta}\right)^{\frac{1}{\theta}}} & \textit{for } h < T \end{cases}$$

2. for the naive *RDU agent:*

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

3. for the sophisticated *RDU agent:*

$$m_h^s = \begin{cases} 1 & \textit{for } h = T \\ \frac{1}{1+\left(\sum_{t=h+1}^{T} \rho^s(h,t)\left(R^{t-h} m_t^{s} \prod_{j=h+1}^{t-1}\left(1-m_j^{s}\right)\right)^{1-\theta}\right)^{\frac{1}{\theta}}} & \textit{for } h < T \end{cases}$$

The formulation of the MPCs in Theorem 3 in terms of discount functions explicitly shows that our RDU life-cycle model falls under a class of ("n-point") discounting models already considered in Pollak (1968). Remarkably, Pollak (1968) proves for this class of discounting models that equilibrium consumption paths coincide for sophisticated and naive RDU decision makers with logarithmic utility function. This equivalence result for the special case of a logarithmic period utility function is trivially implied by Theorem 3. To see this observe at first that the discount functions of the naive and the sophisticated agents are identically given as $\beta^{t-h}\nu_{h,t}$. Next let $\theta = 1$ and observe that the naive and sophisticated types' MPCs become, for $h < T$,

$$m_h^n = m_h^s = \frac{1}{1 + \sum_{t=h+1}^{T} \beta^{t-h}\nu_{h,t}}.$$

Corollary 2 *The realized consumption paths of sophisticated and naive RDU agents coincides for the logarithmic utility function, i.e.,* $\theta = 1$*.*

This insight shares an interesting parallel to the familiar offsetting inter-temporal substitution and income effects from changes in the interest rate R. Recall that the inter-temporal substitution effect means that an increase of R—and thereby a decrease of the relative price of consumption in the next (and all future) periods, $q = 1/R$—leads to an inter-temporal shifting of consumption expenditures so that current period consumption falls. In contrast, the income effect means that a decrease of the relative price of consumption also decreases the absolute price level and so consumption increases in the current and future periods. For logarithmic utility the two effects just cancel.

The formal explanation for the opposing *inter-temporal smoothing* and *constrain ones future self* motives to cancel each other out is directly related. Our closed form solutions for the consumption policy functions rest on the combination of homothetic preferences with a multiplicative dynamic budget constraint (24). For logarithmic utility these multiplicative terms become additive in logs and thereby do not affect consumption behavior. Rewriting the dynamic budget constraint in all periods as $w_{t+1} = (1 - m_t)w_t R$ then readily establishes the formal analogy of changes in the interest rate R and changes in marginal propensities to consume m_t, respectively to save, $1 - m_t$. Beyond the mathematical equivalence result in Pollak (1968), our analysis thus offers an interpretation of this finding in terms of the two opposing motives, which happen to cancel each other out for logarithmic utility.

Remark 8 *Although the actual consumption behaviors of a naive and a sophisticated RDU decision maker, respectively, coincide for the special case of a logarithmic period-utility function, this does not mean that the RDU life-cycle model has suddenly become dynamically consistent for a logarithmic period-utility function. Proposition 2 illustrates that the dynamic inconsistency of our RDU model is exclusively driven by the neo-additive structure of the RDU discount function $\beta^{t-h}\nu_{h,t}$ compared to the dynamically consistent RE discount function $\beta^{t-h}\psi_{h,t}$.*

5 Savings Puzzles Revisited

5.1 A Three-Period Variant of the Model

To investigate the relationship between survival beliefs and savings puzzles we need to distinguish between young and old agents who are still uncertain about their future survival. The simplest specification of the RDU life-cycle model that can address the puzzles of under- and oversaving is therefore a three-period model ($T = 2$): at age $h = 0$ the agent is young, at age $h = 1$ the agent is already old but still uncertain about her future survival chances, at age $h = 2$ the agent knows for sure that she is going to die at the end of the period.[20]

[20] The restriction to a three period model has an additional formal advantage. Recall from Section 4.2 that the only difference between naive and sophisticated RDU agent's marginal propensities shows up in the first period of life ($h = 0$) and that $\xi_1 = 0$, which simplifies the analysis for the sophisticated agent.

Without loss of generality we normalize initial wealth to one, $w_0 = 1$. To further simplify the analysis, we also set the gross interest rate to one, $R = 1$. The next result then immediately follows from Theorem 2.

Proposition 3 *As actual consumption behavior in the 3-period RDU life-cycle model, we obtain from Theorem 2*

1. *for the* RE *agent:*

$$c_2^r = m_2^r w_2^r = 1 - c_0^r - c_1^r, \tag{38}$$
$$c_1^r = m_1^r w_1^r = \frac{1}{1 + \left(\beta\psi_{1,2}\right)^{\frac{1}{\theta}}} \left(1 - c_0^r\right), \tag{39}$$
$$c_0^r = m_0^r w_0 = \frac{1}{1 + (\beta\psi_{0,1})^{\frac{1}{\theta}} + (\beta^2\psi_{0,2})^{\frac{1}{\theta}}}, \tag{40}$$

2. *for the* naive *RDU agent:*

$$c_2^n = m_2^n w_2^n = 1 - c_0^n - c_1^n, \tag{41}$$
$$c_1^n = m_1^n w_1^n = \frac{1}{1 + (\beta\nu_{1,2})^{\frac{1}{\theta}}} \left(1 - c_0^n\right), \tag{42}$$
$$c_0^n = m_0^n w_0 = \frac{1}{1 + (\beta\nu_{0,1})^{\frac{1}{\theta}} + (\beta^2\nu_{0,2})^{\frac{1}{\theta}}}, \tag{43}$$

3. *for the* sophisticated *RDU agent:*

$$c_2^s = m_2^s w_2^s = 1 - c_0^s - c_1^s, \tag{44}$$
$$c_1^s = m_1^s w_1^s = \frac{1}{1 + (\beta\nu_{1,2})^{\frac{1}{\theta}}} \left(1 - c_0^s\right), \tag{45}$$
$$c_0^s = m_0^s w_0 = \frac{1}{1 + \left(1 + (\beta\nu_{1,2})^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}} \left(\beta\nu_{0,1} + \beta^2\nu_{0,2}(\beta\nu_{1,2})^{\frac{1}{\theta}-1}\right)^{\frac{1}{\theta}}}. \tag{46}$$

By an application of Theorem 3, we obtain an alternative characterization of the solution to the three-period model through the corresponding discount factor notation:

Proposition 4 *Expressed in terms of the discount functions* $\rho^r(h,t) = \beta^{t-h}\psi_{h,t}$ *and* $\rho^n(h,t) = \rho^s(h,t) = \beta^{t-h}\nu_{h,t}$ *the MPCs of the three-period RDU life-cycle model are given as follows:*

1. *for the* RE *agent:*

$$m_2^r = 1, \tag{47}$$

$$m_1^r = \frac{1}{1 + (\rho^r(1,2))^{\frac{1}{\theta}}}, \tag{48}$$

$$m_0^r = \frac{1}{1 + \left(\rho^r(0,2)\left(m_2^r\left(1-m_1^r\right)\right)^{1-\theta}\right)^{\frac{1}{\theta}}}, \tag{49}$$

2. *for the* naive *RDU agent:*

$$m_2^n = 1, \tag{50}$$

$$m_1^n = \frac{1}{1 + (\rho^n(1,2))^{\frac{1}{\theta}}}, \tag{51}$$

$$m_0^n = \frac{1}{1 + \left(\rho^n(0,2)\left(m_2^{n,0}\left(1-m_1^{n,0}\right)\right)^{1-\theta}\right)^{\frac{1}{\theta}}}, \tag{52}$$

where

$$m_2^{n,0} = 1 \text{ and } m_1^{n,0} = \frac{1}{1 + \left(\beta\frac{\nu_{0,2}}{\nu_{0,1}}\right)^{\frac{1}{\theta}}}.$$

3. *for the* sophisticated *RDU agent:*

$$m_2^s = 1, \tag{53}$$

$$m_1^s = \frac{1}{1 + (\rho^s(1,2))^{\frac{1}{\theta}}}, \tag{54}$$

$$m_0^s = \frac{1}{1 + \left(\rho^s(0,2)\left(m_2^s\left(1-m_1^s\right)\right)^{1-\theta}\right)^{\frac{1}{\theta}}}. \tag{55}$$

The 3-period variant of our model gives rise to the following straightforward definitions of age-dependent undersaving and oversaving for a RDU decision maker in terms of her MPCs. Whether these conditions hold depends on the values of the model parameters $\theta, \delta, \lambda, \psi_{0,1}, \psi_{1,2}, \beta$.

Definition 3 *We say that the RDU decision maker of type $i \in \{n, s\}$*

- *oversaves at an old age if, and only if, $m_1^i < m_1^r$,*
- *undersaves at a young age if, and only if, $m_0^i > m_0^r$.*

5.2 Oversaving at Old Age

We first turn to the characterization of oversaving at old age 1 in terms of biases in survival beliefs. By Proposition 4, the MPCs of the naive and sophisticated agent coincide at age 1, i.e., $m_1^n = m_1^s$, so that the condition equally applies to both agents.

Proposition 5 *Overestimation of survival chances at an old age is necessary and sufficient for oversaving at an old age, i.e., the RDU decision maker oversaves at an old age if, and only if, $\lambda > \psi_{1,2}$.*

Recall from the discussion in Section 3 that the QHD life-cycle model is nested in our more general RDU life-cycle model as a special case for $\lambda = 0$. Since Proposition (5) requires $\lambda > \psi_{1,2} > 0$, we immediately obtain the following result:

Corollary 3 *Oversaving is not possible in the QHD life-cycle model.*

5.3 Undersaving at Young Age

We next derive conditions for undersaving at young age 0 whereby we keep the same parameters $\theta, \delta, \lambda, \psi_{0,1}, \psi_{1,2}, \beta$ fixed for the naive and the sophisticated agent. At first, translate the condition in Definition 3 in terms of (derived) model parameters from Propositions 3 and 4, respectively. Undersaving at young age occurs if, and only if, we have for the naive RDU agent:

$$m_0^n > m_0^r \tag{56}$$

$$\Leftrightarrow$$

$$\nu_{0,1}^{\frac{1}{\theta}} + (\beta\nu_{0,2})^{\frac{1}{\theta}} < \psi_{0,1}^{\frac{1}{\theta}} + (\beta\psi_{0,2})^{\frac{1}{\theta}}; \tag{57}$$

and for the sophisticated RDU agent:

$$
\begin{aligned}
m_0^s &> m_0^r \qquad (58)\\
&\Leftrightarrow\\
\frac{\left(1+(\beta\nu_{1,2})^{\frac{1}{\theta}}\right)}{\left(1+(\beta\nu_{1,2})^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}}}\left(\nu_{0,1}+\frac{\nu_{0,2}}{\nu_{1,2}}(\beta\nu_{1,2})^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}} &< \psi_{0,1}^{\frac{1}{\theta}}+(\beta\psi_{0,2})^{\frac{1}{\theta}}. \qquad (59)
\end{aligned}
$$

5.3.1 Naive Agent

Focus at first on the naive agent and observe that the LHS in (57) is strictly increasing in the optimism parameter λ. This pins down the threshold level for the optimism parameter, λ^n, that is necessary for young age undersaving of the naive agent:

Proposition 6 *The naive RDU agent undersaves at a young age if, and only if, $\lambda < \lambda^n$ whereby λ^n satisfies*

$$\left(\delta\lambda^n+(1-\delta)\,\psi_{_{0,1}}\right)^{\frac{1}{\theta}}+\left(\beta\left(\delta\lambda^n+(1-\delta)\,\psi_{_{0,2}}\right)\right)^{\frac{1}{\theta}}=\psi_{0,1}^{\frac{1}{\theta}}+(\beta\psi_{0,2})^{\frac{1}{\theta}}. \qquad (60)$$

Corollary 4 *The threshold level of the naive RDU agent satisfies the qualitative relationship*

$$\lambda^n \leq \psi_{0,1}. \qquad (61)$$

Thus, while underestimation of survival chances at age 0 by the naive agent is a necessary condition for undersaving, such underestimation needs to be sufficiently pronounced to generate undersaving.

To see this, let us analyze condition (57) directly. Underestimation of survival chances at age 0, i.e., $\lambda < \psi_{0,1}$ implies that $\nu_{0,1} < \psi_{0,1}$. However, it does not imply $\nu_{0,2} < \psi_{0,2}$ which would otherwise be sufficient to generate undersaving. The reason is that $\nu_{0,2} < \psi_{0,2}$ requires $\lambda < \psi_{0,2} < \psi_{0,1}$. In other words, for a fixed λ the relative underestimation of survival chances is stronger for the probability to survive from age 0 to 1 than it is for the probability to survive from age 0 to age 2, i.e., $\psi_{0,1}-\lambda > \psi_{0,2}-\lambda$. In particular, for a λ such that $\psi_{0,2} < \lambda < \psi_{0,1}$ there are two forces working in opposite directions: (i) underestimation of survival from age 0 to age 1 ($\lambda < \psi_{0,1}$) leads to undersaving, (ii) overestimation of survival from age 0 to age 2 ($\lambda > \psi_{0,2}$) leads to oversaving at age 0. Also notice that the second force becomes stronger when β is increased.

5.3.2 Naive versus Sophisticated Agents

To prepare the discussion of the characterization of the threshold level for the optimism parameter λ^s, which determines undersaving of the sophisticated RDU agent, we compare savings behavior of the naive and the sophisticated RDU agent, which is also of interest per se. Assume, for the moment, that the young sophisticated agent has a smaller MPC than the young naive agent. By Proposition 4, we have that

$$m_0^s < m_0^n \tag{62}$$

$$\Leftrightarrow$$

$$\frac{1}{1+\left(\rho^s(0,2)\,(1-m_1^s)^{1-\theta}\right)^{\frac{1}{\theta}}} < \frac{1}{1+\left(\rho^n(0,2)\,\left(1-m_1^{n,0}\right)^{1-\theta}\right)^{\frac{1}{\theta}}} \tag{63}$$

$$\Leftrightarrow$$

$$\left(1-m_1^{n,0}\right)^{1-\theta} < (1-m_1^s)^{1-\theta} \tag{64}$$

whereby the last step follows from $\rho^s(0,2) = \rho^n(0,2)$. For $\theta < 1$, (64) is equivalent to $m_1^{n,0} > m_1^s$ whereas, for $\theta > 1$, (64) is equivalent to $m_1^{n,0} < m_1^s$. Recall that, by Proposition 1, $m_1^{n,0} < m_1^{n,1}$, whereas, by Proposition 4, $m_1^{n,1} = m_1^s$, so that we always have $m_1^{n,0} < m_1^s$. Consequently, the relationship (62) is always satisfied for $\theta > 1$ but always violated for $\theta < 1$.

Proposition 7 *The following relationships hold for the MPCs of sophisticated and naive agents as functions of θ:*

$$\begin{aligned}
m_0^s &> m_0^n \text{ if } \theta \in (0,1),\\
m_0^s &= m_0^n \text{ if } \theta = 1,\\
m_0^s &< m_0^n \text{ if } \theta \in (1,\infty).
\end{aligned}$$

In words: Sophisticated agents save less at young age than naive agents for $\theta < 1$ (high IES) whereas the converse statement is true for $\theta > 1$ (low IES). In terms of the terminology introduced in Section 4.3.3 this finding means that the *inter-temporal smoothing motive* dominates the *constrain ones future self motive* for $\theta > 1$ (low IES) and vice versa for $\theta < 1$ (high IES).

From Proposition 7 we immediately get the following corollary on undersaving of naive and sophisticated RDU agents:

Corollary 5 *By Proposition 7,*

1. *for $\theta < 1$ (high IES) undersaving of the naive RDU agent implies undersaving of the sophisticated RDU agent and*
2. *for $\theta > 1$ (low IES) undersaving of the sophisticated RDU agent implies undersaving of the naive RDU agent.*

5.3.3 Sophisticated Agent

In terms of threshold levels of the optimism parameter, Corollary 5 has the following implication for undersaving of the sophisticated agent:

Corollary 6 *By Corollary 5,*

1. *for $\theta < 1$ (high IES) $\lambda < \lambda^n$ is a sufficient condition and*
2. *for $\theta > 1$ (low IES) $\lambda < \lambda^n$ is a necessary condition*

for undersaving of the sophisticated agent.

This characterization is not very sharp. We therefore now turn to proving existence of the threshold level λ^s such that for all $\lambda < \lambda^s$ there is undersaving of the sophisticated agent. From Corollary 2 we know that naive and sophisticated agents behave identically for $\theta = 1$ so that $\lambda^s = \lambda^n$ when $\theta = 1$. To characterize λ^s for $\theta \neq 1$ it is sufficient to show that the LHS of (59) is strictly increasing in the optimism parameter λ. Analytically, we could only show this for the case $\theta > 1$ but not for $\theta < 1$. Hence, we treat both cases separately.

Low IES ($\theta > 1$). We first prove existence of a λ^s for $\theta > 1$ so that for all $\lambda < \lambda^s$ there is undersaving of the sophisticated agent:

Proposition 8 *Let $\theta > 1$. Then the sophisticated RDU agent undersaves at a young age if, and only if, $\lambda < \lambda^s$, whereby λ^s satisfies*

$$\begin{aligned}
&\left(1+\left(\beta\left(\delta\lambda^s+(1-\delta)\,\psi_{_{1,2}}\right)\right)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}} && (65)\\
&\cdot\left(\left(\delta\lambda^s+(1-\delta)\,\psi_{_{0,1}}\right)+\beta\left(\delta\lambda^s+(1-\delta)\,\psi_{_{0,2}}\right)\left(\beta\left(\delta\lambda^s+(1-\delta)\,\psi_{_{1,2}}\right)\right)^{\frac{1}{\theta}-1}\right)^{\frac{1}{\theta}}\\
=\;\; & \psi_{0,1}^{\frac{1}{\theta}}+(\beta\psi_{0,2})^{\frac{1}{\theta}}.
\end{aligned}$$

From Propositions 7 and 8 we therefore have (in analogy to Corollary 5):

Corollary 7 *For $\theta > 1$, we have*

$$\lambda^s \leq \lambda^n. \tag{66}$$

Hence, for $\theta > 1$, underestimation of survival beliefs must be stronger for sophisticated agents than for naive agents to generate undersaving.

Low IES ($\theta < 1$). Our formal proof of Proposition 8 (relegated to the Appendix) shows that $\left(\nu_{0,1}+\frac{\nu_{0,2}}{\nu_{1,2}}(\beta\nu_{1,2})^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}}$ is strictly increasing in λ for all values of θ whereas $\frac{\left(1+(\beta\nu_{1,2})^{\frac{1}{\theta}}\right)}{\left(1+(\beta\nu_{1,2})^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}}}$ is increasing in λ if, and only if, $\theta \geq 1$. Consequently, we cannot unambiguously determine the overall effect of an increase in λ on the LHS of (59) for the case of $\theta < 1$ (high IES). We therefore address the case $\theta < 1$ numerically (with details relegated to the Appendix), establishing the following conjecture:

Conjecture 1 *For $\theta < 1$ (high IES) we have*

1. *that the sophisticated RDU agent undersaves at a young age if, and only if, $\lambda < \lambda^s$ with λ^s implicitly given by (65) and*

2. *that underestimation of young-age survival chances, $\lambda < \psi_{0,1}$, is a necessary condition for young-age undersaving.*

We can therefore summarize our discussion on undersaving as follows: independent of the specific value of $\theta > 0$, underestimation must be sufficiently strong to generate undersaving for both the naive and the sophisticated RDU agent.

5.3.4 The Special Case of Logarithmic Utility

For general values of $\theta \neq 1$ we cannot derive an analytical characterization of undersaving at a young age that goes beyond the implicit characterizations in Propositions 6 and 8. For the special case of a logarithmic period utility function, i.e., $\theta = 1$, however, we can present a simple analytical characterization of undersaving at a young age that applies to both types of agents because of Corollary 2:

Proposition 9 *Let $\theta = 1$. Then the RDU decision maker undersaves at a young age if, and only if, λ falls into the interval*

$$[0, \lambda^*) \subset \left[0, \psi_{0,1}\right] \tag{67}$$

where

$$\lambda^s = \lambda^n = \lambda^* = \frac{1 + \beta\psi_{1,2}}{1 + \beta} \cdot \psi_{0,1}. \tag{68}$$

The inequalities

$$\frac{1 + \beta\psi_{1,2}}{1 + \beta} < 1 \qquad \Rightarrow \qquad \lambda^* < \psi_{0,1}$$

establish in closed form our earlier general insights for the naive agent according to which (i) underestimation of survival chances has to be sufficiently strong and (ii) such underestimation has to be stronger when the discount factor increases.

Remark 9 *Note that the likelihood-insensitivity parameter δ does not appear in (68) so that its specific value (besides being non-zero) has no impact on the undersaving threshold value λ^*. That is, for a logarithmic utility function only the optimism but not the likelihood-insensitivity parameter δ determines whether the RDU agent will undersave at a young age or not. An inspection of the inequalities in Proposition 6 shows that the situation will be different for non-logarithmic utility functions where the specific value of δ will impact on the question whether there is undersaving or not for fixed values of λ. Moreover, this impact will differ between naive and sophisticated agents. While an analytical investigation of the role of δ for undersaving with non-logarithmic utility functions turns out to be non-tractable, we will address this role in our numerical analysis of Section 6.*

5.4 Empirically Relevant Survival Belief Biases

In line with the HRS data, cf. Section 2, we now restrict attention to the empirically relevant case of biases in survival beliefs in which agents underestimate their survival chances at young but overestimate them at old ages, i.e., we assume that $\lambda \in (\psi_{1,2}, \psi_{0,1})$. To obtain a closed form solution for the co-occurrence of undersaving at young and oversaving at old ages, we again restrict attention to a logarithmic utility function so that our results in this section apply to both naive and sophisticated RDU agents (cf. Corollary 2).

While Proposition 9 establishes that $\lambda < \lambda^{\star}$ is required to generate undersaving, $\lambda > \psi_{1,2}$ is, by Proposition 5, required to generate oversaving. Combining both propositions thus implies undersaving at young and oversaving at old ages co-occur if, and only if, $\lambda \in (\psi_{1,2}, \lambda^{\star})$ with $\lambda^{\star}$ given by (68). For large values of $\psi_{1,2}$, however, this interval might be empty. The following proposition provides a clear characterization.

Proposition 10 *Let $\theta = 1$ and consider a RDU decision maker who underestimates her survival chances at a young and overestimates them at an old age, i.e., $\lambda \in (\psi^1_{1,2}, \psi^0_{0,1})$. The decision maker undersaves at a young and oversaves at an old age if, and only if, we have for the optimism parameter*

$$\lambda \in \left(\psi_{1,2}, \frac{1+\beta\psi_{1,2}}{1+\beta} \cdot \psi_{0,1}\right). \tag{69}$$

The interval is empty if $\frac{\psi_{0,1}}{1+\beta(1-\psi_{0,1})} < \psi_{1,2} < \psi_{0,1} < 1$.

Our RDU life-cycle model with neo-additively transformed survival probabilities can thus indeed generate undersaving at young combined with oversaving at old ages for decision makers who underestimate their survival chances at young and overestimate them at old ages. In this specific sense our model establishes the possibility that the observed savings puzzles can be explained through the age-dependent biases in survival beliefs as reported in the HRS data.

6 Numerical Illustrations

Although our analytical approach yields very useful insights, there are (at least) two main limitations even for the three-period variant of the RDU life-cycle model. First, it is impossible to establish closed form conditions for which oversaving dominates undersaving over the life-cycle. Under such conditions old-age asset holdings in the RDU model would exceed those of the RE model thereby addressing the additional stylized fact of too high old-age asset holdings. Second, while Proposition 7 establishes qualitative differences in the savings behavior of naive versus sophisticated agents, it would be interesting to see how the magnitude of these differences varies with the likelihood-insensitivity parameter δ. We would also like to see how these parameters impact on the threshold levels λ^n and λ^s that determine young-age undersaving for the naive agent and her sophisticated counterpart, respectively.

6.1 High Old-Age Asset Holdings

We say that the three-period RDU model generates high old-age asset holdings if, and only if, the period 2 asset holdings (i.e., the period 2 consumption) is higher for the RDU than for the RE agent. More precisely, our model generates high old-age asset holdings for the naive RDU agent if, and only if (for $\nu_{ij} = \delta\lambda + (1-\delta)\psi_{ij}$),

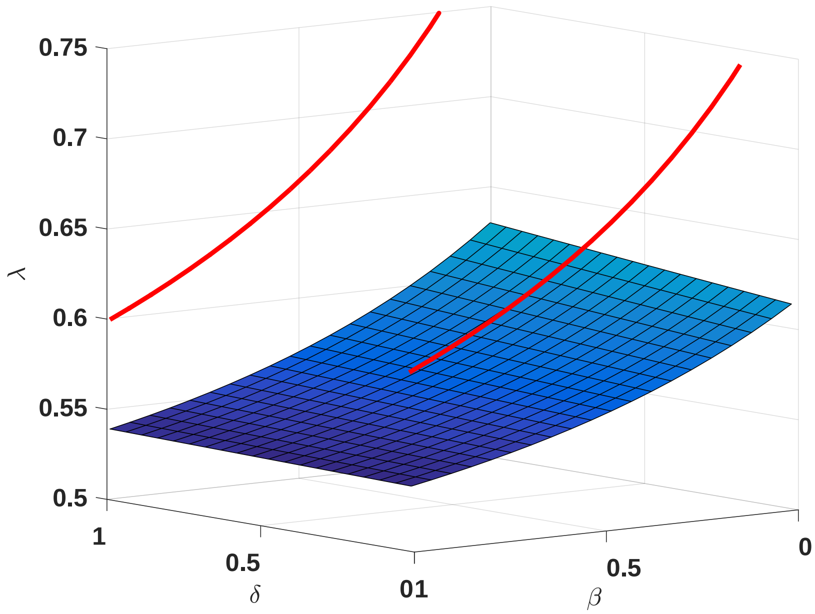
$$\begin{aligned} & (1-m_0^n)\cdot(1-m_1^n) \geq (1-m_0^r)\cdot(1-m_1^r) \\ \Leftrightarrow \quad & \frac{(\beta\nu_{0,1})^{\frac{1}{\theta}} + (\beta^2\nu_{0,2})^{\frac{1}{\theta}}}{1+(\beta\nu_{0,1})^{\frac{1}{\theta}} + (\beta^2\nu_{0,2})^{\frac{1}{\theta}}} \frac{(\beta\nu_{1,2})^{\frac{1}{\theta}}}{1+(\beta\nu_{1,2})^{\frac{1}{\theta}}} \geq \frac{(\beta\psi_{0,1})^{\frac{1}{\theta}} + (\beta^2\psi_{0,2})^{\frac{1}{\theta}}}{1+(\beta\psi_{0,1})^{\frac{1}{\theta}} + (\beta^2\psi_{0,2})^{\frac{1}{\theta}}} \frac{\left(\beta\psi_{1,2}\right)^{\frac{1}{\theta}}}{1+\left(\beta\psi_{1,2}\right)^{\frac{1}{\theta}}}. \end{aligned} \tag{70}$$

Our objective is to choose a parametrization that generates high old-age asset holdings. To this purpose we set $\theta = 1$ (so that our findings equally apply to naive and sophisticated RDU agents), $\psi_{0,1} = 0.99$, $\psi_{1,2} = 0.5$ and consider $\beta \in (0,1]$. For this parametrization of the model we first compute λ to characterize the interval (69). Recall that this is independent of δ.[21] This upper threshold of the interval (69) is characterized by the red solid curves in Figure 2. The lower threshold is $\psi_{1,2} = 0.5$. Hence, for any λ above 0.5 and below the red solid curves of Figure 2, undersaving at young and oversaving at old age co-occur.

Next, we additionally consider $\delta \in (0,1]$ and compute for all parameter constellations λ such that condition (70) holds. This is illustrated by the surface in Figure 2 with the interpretation that for all λ above that surface asset holdings of the RDU

[21]For our choice of parameters, this interval is never empty. We could achieve emptiness for some low values of β if we were to decrease $\psi_{0,1}$ towards $\psi_{1,2}$, cf. equation (69).

Figure 2: Threshold Values of λ for RDU Agents



Notes: Notes: Threshold values for λ to illustrate the regions of undersaving, oversaving and high old-age asset holdings, see the main text, for $\theta = 1, \psi_{0,1} = 0.99, \psi_{1,2} = 0.5$.

agent exceed those of the RE agent. Observe that this is of course stricter than the requirement of oversaving.

To summarize, this figure is a useful illustration of four cases:

1. For $\lambda < \psi_{1,2} = 0.5$ there is undersaving but no oversaving.

2. For $\lambda > \psi_{1,2}$ but less than the surface of Figure 2 there is undersaving and oversaving but period 2 asset holdings of the RE agent still exceed those of the RDU agent. Hence, overestimation of old-age survival beliefs (and thereby oversaving) is not strong enough.

3. For λ above the surface but below the red lines, all three phenomena (undersaving, oversaving and high old-age asset holdings) occur jointly.

4. For λ above the red lines, there is oversaving and high old-age asset holdings but overestimation of old-age survival beliefs (and thereby oversaving) is so strong that there no longer undersaving.

In light of the empirical and quantitative literature on savings puzzles over the life-cycle summarized in Section 1, the third case is the most interesting parameter constellation because it simultaneously generates all three saving puzzles, undersaving, oversaving and high old-age asset holdings, as a result of under- and overestimation of survival beliefs. However, these results also illustrate—thereby summarizing our main findings—that it is not an obvious conclusion that such biases in survival beliefs lead to a resolution of all savings puzzles.

6.2 Differences between Naive and Sophisticated Agents

We now illustrate the differences in the MPCs of the naive agent and her sophisticated counterpart as well as the differences in their respective threshold levels of the optimism parameter that induce young-age undersaving.

6.2.1 Differences in MPCs

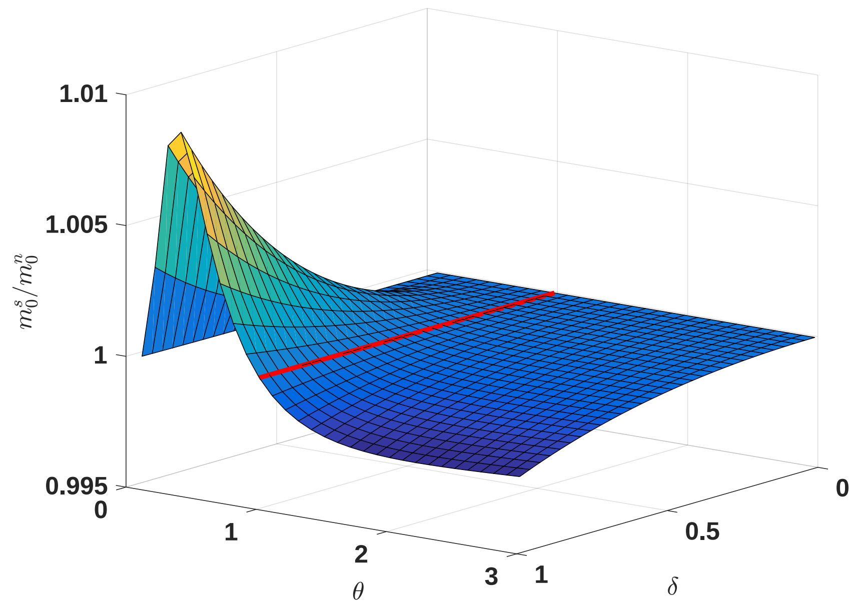
Let us first look at the magnitude of the difference between the savings behavior of the sophisticated versus the naive RDU agent. To this purpose, we now set $\beta = 0.99^{30} \approx 0.75$, $\psi_{0,1} = 0.99$, $\psi_{1,2} = 0.5$, $\lambda = 0.7$. We additionally vary $\theta \in \{0.01, \ldots, 1, \ldots, 5\}$ and the likelihood insensitivity parameter by considering $\delta \in \{0.01, \ldots, 1\}$. Results are shown in Figure 3. Recall from Proposition 7 that for $\theta = 1$ the MPCs of both types of agents coincide whereas the sophisticated agent consumes more, respectively less, than the naive agent for $\theta < 1$, respectively $\theta > 1$. As the figure shows, the magnitude of this difference in MPCs increases in likelihood-insensitivity, δ. For values of δ close to one, these differences are no longer negligible if the agents' per period utility features a high IES. For a low IES, differences continue to be moderate even with strong likelihood insensitivity.[22]

6.2.2 Differences in Optimism Parameter Thresholds

Turn now to the thresholds levels λ^n and λ^s as pinned down through the equations (60) and (65), respectively. By definition, the naive agent undersaves iff $\lambda < \lambda^n$ whereas the sophisticated agent undersaves iff $\lambda < \lambda^s$. The ratio $\frac{\lambda^s}{\lambda^n}$ thus measures the difference in underestimation of survival chances between both types of agents which

[22]It is therefore not surprising that some authors observe small differences between naive and sophisticated agents, cf., e.g., Angeletos et al. (2001, p. 52) who choose an IES of 0.5.

Figure 3: Ratio of Marginal Propensities to Consume, m_0^s/m_0^n



Notes: Ratio of marginal propensities to consume between the sophisticated and the naive RDU agent in period 0, m_0^s/m_0^n, for $\beta = 0.99^{30}, \psi_{0,1} = 0.99, \psi_{1,2} = 0.5, \lambda = 0.7$.

is (just) required to generate undersaving for the respective type. To characterize this ratio we consider the same parametrization as in the previous exercise of Figure 3. Results are shown in Figure 4. To interpret those, first recall from Proposition (10), that for $\theta = 1$

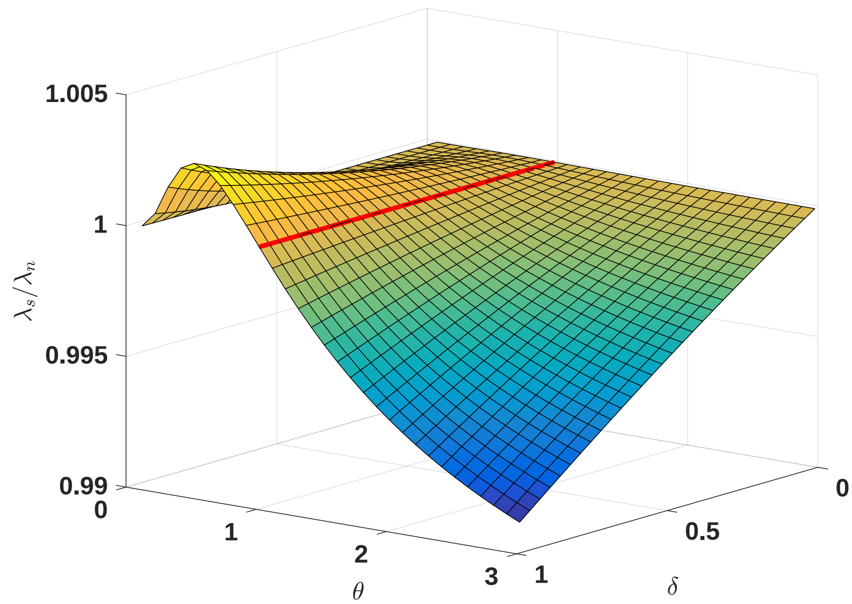
$$\lambda^n = \lambda^s = \lambda^* = \frac{1 + \beta\psi_{1,2}}{1 + \beta} \cdot \psi_{0,1} \simeq 0.0778.$$

and from Corollary 7 that $\lambda^n < \lambda^s$ for $\theta < 1$ and $\lambda^n > \lambda^s$ for $\theta > 1$. Turning to the results in the figure, notice that as with the ratio of the MPCs shown in Figure 3, the ratio $\frac{\lambda^s}{\lambda^n}$ increases in the value of the likelihood-insensitivity parameter δ. Again, for δ close to one, these differences are no longer negligible for a high IES.

7 Concluding Remarks

Our main research question in this paper is to investigate whether the RDU model can give rise to important savings puzzles discussed in the life-cycle literature, namely that it generates too little saving relative to the RE benchmark model when house-

Figure 4: Ratio of the Optimism Parameter Thresholds, λ^s/λ^n



Notes: Ratio of the optimism parameter thresholds between the sophisticated and the naive RDU agent, λ^s/λ^n, for $\beta = 0.99^{30}, \psi_{0,1} = 0.99, \psi_{1,2} = 0.5$.

holds are young and too much saving (or even too high old-age asset holdings) when households are old. As our main answer to this question we establish that only a subset of parameter constellations can give rise to the joint occurrence of oversaving and undersaving; in particular, the familiar quasi-hyperbolic time discounting (QHD) model—which is nested as a special case—cannot generate oversaving.

The basic intuition for our main finding is as follows: whenever households overestimate their survival chances when old, then they will oversave; if households also underestimate their survival chances when young, then they will have a tendency to undersave; however, for undersaving to indeed occur, this underestimation must be sufficiently strong in order to dominate the overestimation of survival to old age. These results refer to the flow of savings. We further establish (numerically) that to also generate high old-age asset holdings, overestimation has to be sufficiently strong. Our analysis therefore shows that underestimation of survival chances when young combined with overestimation when old are in no way sufficient conditions for generating the well-known savings puzzles.

The reason why our RDU model can lead to oversaving whereas the QHD model cannot is an additional degree of freedom in the parametrization. The QHD model only features a short term discount factor, giving rise to stronger impatience and therefore to undersaving relative to the RE benchmark. Our model features an optimism and a likelihood-insensitivity parameter for the description of subjective survival beliefs. This additional degree of freedom implies that one has to profoundly discipline the calibration. We adopt such a disciplined approach in our quantitative work (Groneck, Ludwig, and Zimper 2016) which is based on a model of Bayesian learning of neo-additive survival beliefs. There we estimate the two parameters—the optimism and the likelihood insensitivity parameter—outside the life-cycle model by using data on subjective survival beliefs from the Health and Retirement Study (HRS). The only parameter we identify by use of the structural model is the households' discount factor which we pin down by matching average asset holdings but not their shape. Our finding there is that this calibration goes a long way in explaining the observed savings puzzles: the quantitative model generates undersaving, oversaving and high old-age asset holdings and matches the average life-cycle asset profile in the data remarkably well. Our main point here is to show that these quantitative findings are in no way obvious, although intuition may suggest so at first glance.

Although our structural behavioral economics approach bridges the empirical literature on subjective survival beliefs and the decision theoretic literature on RDU in order to provide a clean characterization of conditions leading to oversaving and undersaving in a RDU life-cycle model, several aspects remain unaddressed. This gives rise to a number of research avenues out of which we find two particularly interesting. First, our RDU life-cycle model only considers survival risks but neither health risks nor does it feature a bequest motive. Both have been identified in the RE literature as crucial elements to generate high old-age asset holdings, cf. De Nardi, French, Jones, DeNardi, and Nardi (2010, 2016) and Lockwood (2012). Our results suggest that the estimates in these papers of bequest motives and of precautionary savings motives in light of health risks are upward biased because the underlying RE models do not feature overestimation of old-age objective survival risks. It is certainly a challenging endeavor—especially with respect to the identification of all the relevant model parameters and the modeling of subjective health risk—to extend our present theoretical analysis and our quantitative work in Groneck, Ludwig, and Zimper (2016) by these features. With respect to a bequest motive our results suggest that identification of preference parameters for such a motive must come from consumption rather than asset data. The reason is that overestimation of survival probabilities as well as a bequest motive lead to slow old-age asset decumulation. In the first case, assets are used to finance own consumption, in the second case they are used to finance the consumption of ones offspring. Second, we investigate in an ongoing empirical project (Grevenbrock et al. 2016) whether our cognitive and psychological interpretation (i.e., likelihood insensitivity and optimism) of biases of survival probabilities is valid by exploring newly available direct cognitive and psychological measures in the HRS. In this work, we identify age-specific inverse S-shaped survival probability weighting functions that can reasonably well be approximated by linear neo-additive probability weighting functions as in the present paper. Our results in this ongoing empirical project show that the shapes of these functions are influenced by the direct cognitive and psychological measures, which is consistent with our theory.

References

Abdellaoui, M., A. Baillon, L. Placido, and P. P. Wakker (2011). The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation. *American Economic Review 101*(2), 695–723.

Ainslie, G. (1992). *Picoeconomics*. Cambridge University Press.

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

Angeletos, G.-M., D. Laibson, A. Repetto, J. Tobacman, and S. Weinberg (2001). The Hyperbolic Consumption Model: Calibration, Simulation and Empirical Evaluation. *Journal of Economic Perspectives 15 (3)*, 47–68.

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro (1997). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *Quarterly Journal of Economics 112*(2), 537–579.

Bernheim, D. and A. Rangel (2007). Behavioral Public Economics: Welfare and Policy Analysis with Nonstandard Decision-Makers. In P. Diamond and H. Vartianen (Eds.), *Behavioral Economics and Its Applications*. Princeton University Press.

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

Chateauneuf, A. (1999). Comonotonicity Axioms and RDEU Theory for Arbitrary Consequences. *Journal of Mathematical Economics 32*(1), 21–45.

Chateauneuf, A. and M. Coen (2000). Choquet Expected Utility Model: A New approach to Individual Behavior under Uncertainty and Social Welfare. In M. Grabisch, T. Murofushi, and M. Sugeno (Eds.), *Fuzzy Measures and Integrals*, pp. 289–313. Physica-Verlag.

Chateauneuf, A., J. Eichberger, and S. Grant (2007). Choice under Uncertainty with the Best and Worst in Mind: Neo-Additive Capacities. *Journal of Economic Theory 127*, 538–567.

Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2006). Saving for Retirement on the Path of Least Resistance. In E. J. McCaffrey and J. Slemrod (Eds.), *Behavioral Public Finance: Toward a New Agenda*, pp. 304–351. New York: Russell Sage Foundation.

Cohen, M. (1992). Security Level, Potential Level, Expected Utility: A Three-Criteria Decision Model under Risk. *Theory and Decision 33*, 101–134.

De Nardi, M., E. French, and J. B. Jones (2016). Medicaid Insurance in Old Age. *The American Economic Review 106*(11), 3480–3520.

De Nardi, M., E. French, J. B. Jones, M. DeNardi, and M. D. Nardi (2010). Why Do the Elderly Save? The Role of Medical Expenses. *Journal of Political Economy 118*(1), 39–75.

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

Elder, T. E. (2013). The Predictive Validity of Subjective Mortality Expectations: Evidence From the Health and Retirement Study. *Demography 50*(2), 569–589.

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*(3), 169–203.

Essid, S. (1997). Choice under Risk with Certainty and Potential Effects: A General Axiomatic Model. *Mathematical Social Sciences 34*, 223–247.

Gan, L., M. D. Hurd, and D. L. McFadden (2005). Individual Subjective Survival Curves. In *Analyses in the Economics of Aging*, pp. 377–412. University of Chicago Press.

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

Gilboa, I. (1988). A Combination of Expected Utility and Maxmin Decision Criteria. *Journal of Mathematical Psychology 32*, 405–420.

Grevenbrock, N., M. Groneck, A. Ludwig, and L. Zimper (2016). Subjective Survival Beliefs and Ambiguity: The Role of Psychological and Cognitive Factors. Mimeo.

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

Halevy, Y. (2008). Strotz Meets Allais: Diminishing Impatience and the Certainty Effect. *American Economic Review 98*(3), 1145–1162.

Halevy, Y. (2015). Time Consistency: Stationarity and Time Invariance. *Econometrica 83*(1), 335–352.

Hammermesh, D. S. (1985). Expectations, Life Expectancy, and Economic Behavior. *Quarterly Journal of Economics 100(2)*, 389–408.

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

Hong, C. S., E. Karni, and Z. Safra (1987). Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities. *Journal of Economic Theory 42*(2), 370–381.

Hurd, M. and S. Rohwedder (2010). Wealth Dynamics and Active Saving at Older Ages: Do They Add Up. Mimeo.

Jaffray, J. Y. (1988). Choice under Risk and the Security Factor: An Axiomatic Model. *Theory and Decision 24*, 169–200.

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica 47*, 263–291.

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.

Laibson, D. I., A. Repetto, J. Tobacman, R. E. Hall, W. G. Gale, and G. A. Akerlof (1998). Self-Control and Saving for Retirement. *Brookings Papers on Economic Activity 1998*(1), 91–196.

Lockwood, L. M. (2012). Incidental Bequests: Bequest Motives and the Choice to Self-Insure Late-Life Risks. Working Paper.

Loewenstein, G. and R. H. Thaler (1989). Anomalies: Intertemporal Choice. *The Journal of Economic Perspectives 3*(4), 181–193.

Ludwig, A. and A. Zimper (2013). A Parsimonious Model of Subjective Life Expectancy.

Lusardi, A. and O. S. Mitchell (2011). Financial Literacy and Planning: Implication for Retirement Wellbeing. In A. Lusardi and O. S. Mitchell (Eds.), *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*, pp. 37. Oxford University Press.

Machina, M. (1994). Book Review of: John Quiggin (1993): "Generalized Expected Utility Theory - The Rank-Dependent Model". *Journal of Economic Literature 32*, 1237–1238.

Manski, C. F. (2004). Measuring Expectations. *Econometrica 72(5)*, 1329–1376.

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.

Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica 29*, 315–335.

Peracchi, F. and V. Perotti (2010). Subjective Survival Probabilities and Life Tables: Evidence from Europe. Working Paper.

Phelps, E. S. and R. A. Pollak (1968). On Second-Best National Saving and Game-Equilibrium Growth. *The Review of Economic Studies 35*(2), 185–199.

Pollak, R. A. (1968). Consistent Planning. *The Review of Economic Studies 35*(2), 201–208.

Quiggin, J. (1982). A Theory of Anticipated Utility. *Journal of Economic Behavior and Organization 3*, 323–343.

Quiggin, J. (1993). *Generalized Expected Utility Theory: The Rank-Dependent Model*. Springer.

Quiggin, J. P. (1981). Risk Perception and Risk Aversion among Australian Farmers. *Australian Journal of Agricultural Economics 25*, 160–169.

Samuelson, P. A. (1937). A Note on Measurement of Utility. *The Review of Economic 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*, 571–587.

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*, 297–323.

Wakker, P. P. (2009). *Prospect Theory for Risk and Ambiguity*. Cambridge, UK: Cambridge University Press.

Wu, G. and R. Gonzalez (1996). Curvature of the Probability Weighting Function. *Management Science 42*(12), 1676–1690.

Yaari, M. E. (1987). The Dual Theory of Choice under Risk. *Econometrica 55*(1), 95–115.

A Formal Proofs

Proof of Theorem 2. For all agents the objective is concave and the constraint is linear so that the programming problem is convex. Hence, first-order conditions together with the transversality condition are necessary and sufficient. Our proof is by backward induction iterating on the first-order conditions. We simplify notation and generally write $m_t^{i,h}, c_t^{i,h}$ for all $i \in \{r, s, n\}$, whereby $m_t^{i,h} = m_t^i, c_t^{i,h} = c_t^i$ for $i \in \{r, s\}$. The claim is that $c_t^{i,h} = m_t^{i,h} w_t$ for all i, t. The base case is that $m_T^{i,h} = 1$ for all i which is implied by the no-Ponzi and transversality condition. Both conditions, together with the lower Inada condition $\lim_{c \to 0} u_c = \infty$ also imply that $m_t^{i,h} \in (0, 1)$ for all $i, t < T$. The remainder of the proof shows the induction steps for the three agent types.

1. RE: from the first-order condition (76) we have

$$\begin{aligned} u_c(c_t^r) &= \beta \psi_{t,t+1} R u_c(c_{t+1}^r) \\ \Leftrightarrow \qquad {c_t^r}^{-\theta} &= \beta \psi_{t,t+1} R \left(m_{t+1}^r \left(w_t - c_t^r \right) R \right)^{-\theta} \end{aligned}$$

   from which (26) immediately follows.

2. Naive RDU: from the first-order condition (77) we get

$$\begin{aligned} u_c(c_t^{n,h}) &= \beta \frac{\nu_{h,t+1}}{\nu_{h,t}} R u_c(c_{t+1}^{n,h}) \\ \Leftrightarrow \qquad {c_t^{n,h}}^{-\theta} &= \beta \frac{\nu_{h,t+1}}{\nu_{h,t}^h} R \left(m_{t+1}^{n,h} \left(w_t - c_t^{n,h} \right) R \right)^{-\theta} \end{aligned}$$

   from which (27) immediately follows.

3. Sophisticated RDU: Start from equation (78), cf. the Proof of Proposition 11 in Appendix B.1.

   (a) First, derive the closed form solution for the derivative of the value function: From the proof of Proposition 11 we have that under the induction claim the derivative of the value function for any period t, $h \leq t < T$, cf. equation (82), writes as

$$\frac{\partial V_t^h(\cdot)}{\partial w_t} = {c_t^s}^{-\theta} m_t^s + \beta \frac{\nu_{h,t+1}}{\nu_{h,t}} R \left(1 - m_t^s \right) \frac{\partial V_{t+1}^h(\cdot)}{\partial w_{t+1}} \tag{71}$$

We now derive an expression for the derivative of the value function in terms of period t wealth, again by backward induction.

i. Claim: In any period $t \leq T$ the derivative of the value function writes as
$$\frac{\partial V_t^h(\cdot)}{\partial w_t} = \zeta_t^h w_t^{-\theta} \tag{72}$$
for some $\zeta_t^h > 0$.

ii. Base case: In period T we have $V_T^h = \frac{1}{1-\theta} {c_T^s}^{1-\theta} = \frac{1}{1-\theta} w_T^{1-\theta}$ so that $\frac{\partial V_T^h(\cdot)}{\partial w_T} = w_T^{-\theta}$ and $\zeta_T^h = 1$.

iii. Induction step: Use (72) in (71) to get
$$\frac{\partial V_t^h(\cdot)}{\partial w_t} = c_t^{s-\theta} m_t^s + \beta \frac{\nu_{h,t+1}}{\nu_{h,t}} R \left(1 - m_t^s\right) \zeta_{t+1}^h w_{t+1}^{-\theta}.$$
Using $c_t^s = m_t^s w_t$ and $w_{t+1} = (w_t - c_t^s)R = (1 - m_t^s) w_t R$ in the above we get
$$\frac{\partial V_t^h(\cdot)}{\partial w_t} = \left({m_t^s}^{1-\theta} + \beta \frac{\nu_{h,t+1}}{\nu_{h,t}} R^{1-\theta} \left(1 - m_t^s\right)^{1-\theta} \zeta_{t+1}^h \right) w_t^{-\theta} = \zeta_t^h w_t^{-\theta}$$
so that we can generally write
$$\zeta_t^h = \begin{cases} 1 & \text{for } t = T \\ {m_t^s}^{1-\theta} + \beta \frac{\nu_{h,t+1}}{\nu_{h,t}} R^{1-\theta} \left(1 - m_t^s\right)^{1-\theta} \zeta_{t+1}^h & \text{otherwise.} \end{cases}$$

(b) Next, use (72) to rewrite the distance of derivatives of value functions showing up in term Ξ_{h+1} in equation (78) as
$$\Delta V_{h+2}^{h,h+1} = \frac{\partial V_{h+2}^h(\cdot)}{\partial w_{h+2}} - \frac{\partial V_{h+2}^{h+1}(\cdot)}{\partial w_{h+2}} = \left(\zeta_{h+2}^h - \zeta_{h+2}^{h+1}\right) w_{h+2}^{-\theta}.$$

Use the above to rewrite Ξ_{h+1} to get

$$\begin{aligned}
\Xi_{h+1} &= \beta \frac{\nu_{h,h+2}}{\nu_{h,h+1}} R\left(1-m^s_{h+1}\right)\left(\frac{\partial V^h_{h+2}}{\partial w_{h+2}} - \frac{\partial V^{h+1}_{h+2}}{\partial w_{h+2}}\right) \\
&= \beta \frac{\nu_{h,h+2}}{\nu_{h,h+1}} R\left(1-m^s_{h+1}\right)\left(\zeta^h_{h+2} - \zeta^{h+1}_{h+2}\right)\left((1-m^s_{h+1})w_{h+1}R\right)^{-\theta} \\
&= \beta R^{1-\theta} \frac{\nu_{h,h+2}}{\nu_{h,h+1}} \left(1-m^s_{h+1}\right)^{1-\theta} m^{s^\theta}_{h+1}\left(\zeta^h_{h+2} - \zeta^{h+1}_{h+2}\right) m^{s^{-\theta}}_{h+1} w^{-\theta}_{h+1} \\
&= \xi_{h+1} m^{s^{-\theta}}_{h+1} w^{-\theta}_{h+1}.
\end{aligned}$$

which proves equation (30).

(c) Induction step: Use the above in the first-order condition (78), cf. Proposition 11 in Appendix B.1, to get

$$\begin{aligned}
& c^{s^{-\theta}}_h = \beta R \nu_{h,h+1}\left(\Theta_{h+1} c^{s^{-\theta}}_{h+1} + \xi_{h+1} m^{s^{-\theta}}_{h+1} w^{-\theta}_{h+1}\right) \\
\Leftrightarrow \qquad & c^s_h = \frac{1}{1 + \frac{\left(\beta R^{1-\theta}\nu_{h,h+1}\left(\Theta_{h+1}+\xi_{h+1}\right)\right)^{\frac{1}{\theta}}}{m^s_{h+1}}} w_h.
\end{aligned}$$

which proves equation (28).

(d) Finally, to see that $\zeta^h_{h+2} - \zeta^{h+1}_{h+2} \geq 0$ for all $h \leq T-2$, first, notice that $\zeta^h_T = 1$ for all h so that the distance is zero for $h = T-2$. Next, observe from (31) that for any t such that $h+2 \leq t < T$ the difference can be rewritten as

$$\begin{aligned}
& \zeta^h_t - \zeta^{h+1}_t = \beta R^{1-\theta}\left(1-m^s_t\right)^{1-\theta}\left(\frac{\nu_{h,t+1}}{\nu_{h,t}}\zeta^h_{t+1} - \frac{\nu_{h+1,t+1}}{\nu_{h+1,t}}\zeta^{h+1}_{t+1}\right) \geq 0 \\
\Leftrightarrow \qquad & \frac{\nu_{h,t+1}}{\nu_{h,t}}\zeta^h_{t+1} \geq \frac{\nu_{h+1,t+1}}{\nu_{h+1,t}}\zeta^{h+1}_{t+1}
\end{aligned}$$

for which a sufficient condition is

$$\frac{\nu_{h,t+1}}{\nu_{h,t}} \geq \frac{\nu_{h+1,t+1}}{\nu_{h+1,t}} \qquad \wedge \qquad \zeta^h_{t+1} \geq \zeta^{h+1}_{t+1}.$$

For the second part of this sufficient condition, we accordingly require that

$$\frac{\nu_{h,t+2}}{\nu_{h,t+1}}\zeta^h_{t+2} \geq \frac{\nu_{h+1,t+2}}{\nu_{h+1,t+1}}\zeta^{h+1}_{t+2}$$

for which a sufficient condition is

$$\frac{\nu_{h,t+2}}{\nu_{h,t+1}} \geq \frac{\nu_{h+1,t+2}}{\nu_{h+1,t+1}} \qquad \wedge \qquad \zeta_{t+2}^{h} \geq \zeta_{t+2}^{h+1}.$$

Since $\zeta_T^h = \zeta_T^{h+1} = 1$ we therefore have as sufficient condition for $\zeta_t^h \geq \zeta_{t+1}^h$ for any t such that $h+2 \leq t < T$ that

$$\frac{\nu_{h,j}}{\nu_{h,j-1}} \geq \frac{\nu_{h+1,j}}{\nu_{h+1,j-1}}, \text{ for all } j = t+1, \ldots, T.$$

Now, look at any $j \in \{t+1, \ldots, T\}$. We have

$$\begin{aligned}
& \frac{\nu_{h,j}}{\nu_{h,j-1}} \geq \frac{\nu_{h+1,j}}{\nu_{h+1,j-1}} \\
\Leftrightarrow \qquad & \frac{\delta\lambda + (1-\delta)\psi_{h,j}}{\delta\lambda + (1-\delta)\psi_{h,j-1}} \geq \frac{\delta\lambda + (1-\delta)\psi_{h+1,j}}{\delta\lambda + (1-\delta)\psi_{h+1,j-1}} \\
\Leftrightarrow \qquad & \frac{\delta\lambda + (1-\delta)\psi_{h,h+1}\psi_{h+1,j}}{\delta\lambda + (1-\delta)\psi_{h,h+1}\psi_{h+1,j-1}} \geq \frac{\delta\lambda + (1-\delta)\psi_{h+1,j}}{\delta\lambda + (1-\delta)\psi_{h+1,j-1}}
\end{aligned}$$

and the last inequality is strict for $\delta \in (0,1), \lambda \in (0,1)$ because $\psi_{h,h+1} \in (0,1)$ so that the LHS of the above expression is closer to one than the RHS. Hence, for $\delta \in (0,1), \lambda \in (0,1)$ we have $\zeta_t^h - \zeta_t^{h+1} > 0$ for any t such that $h+2 \leq t < T$.

Also notice that the above weak inequality holds with equality for (i) $\delta = 0$ (RE model), (ii) $\delta \in (0,1), \lambda = 0$ (QHD model, cf. Section 3.2) and for $\delta = 1, \lambda \in (0,1)$ (QHD model w/o mortality risk, cf. Section 3.2) so that in these cases $\zeta_t^h = \zeta_t^{h+1}$ for any t such that $h+2 \leq t < T$.

□

Proof of Proposition 1. The proof proceeds in three steps. First, we establish a necessary condition for $m_{h+1}^{n,h} \leq m_{h+1}^{n,h+1}$. Next, we establish a sufficient condition for the necessary condition to hold. Finally we show that this sufficient condition always holds in our model.

1. According to equation (27) the decision in period h for the MPC in period $h+1$

by the naive agent is

$$m_{h+1}^{n,h} = \frac{1}{1 + \frac{\left(\beta \frac{\nu_{h,h+2}}{\nu_{h,h+1}} R^{1-\theta}\right)}{m_{h+2}^{n,h}}} = \frac{1}{\sum_{t=h}^{T-1} \left(\beta R^{1-\theta}\right)^{t-h} \frac{\nu_{h,t+1}}{\nu_{h,h+1}}},$$

where the second term follows from backward substitution. However, when the agent turns $h+1$ his MPC is

$$m_{h+1}^{n,h+1} = \frac{1}{1 + \frac{\left(\beta \nu_{h+1,h+2} R^{1-\theta}\right)}{m_{h+2}^{n,h+1}}} = \frac{1}{\sum_{t=h}^{T-1} \left(\beta R^{1-\theta}\right)^{t-h} \nu_{h+1,t+1}}.$$

From this we observe that the naive RDU agent consumes more in period $h+1$ than planned in period h, i.e.,

$$c_{h+1}^{n} = c_{h+1}^{n,h+1} \geq c_{h+1}^{n,h}$$

if and only if

$$\sum_{t=h+1}^{T-1} \left(\beta R^{1-\theta}\right)^{t-h} \left(\nu_{h,t+1} - \nu_{h,h+1}\nu_{h+1,t+1}\right) \geq 0. \tag{73}$$

A sufficient condition for (73) is obviously that condition (35) holds, i.e.,

$$\nu_{h,t+1} \geq \nu_{h,h+1}\nu_{h+1,t+1} \quad \text{for all } t = h+1, \ldots, T-1.$$

2. We now show that our maintained assumption of decreasing survival chances

implies condition (35). To see this, notice that

$$
\begin{aligned}
& \nu_{h,t+1} \geq \nu_{h,h+1}\nu_{h+1,t+1} \\
\Leftrightarrow \quad & \frac{\nu_{h,t+1}}{\nu_{h+1,t+1}} \geq \nu_{h,h+1} \\
\Leftrightarrow \quad & \frac{\delta\lambda + (1-\delta)\psi_{h,t+1}}{\delta\lambda + (1-\delta)\psi_{h+1,t+1}} \geq \delta\lambda + (1-\delta)\psi_{h,h+1} \\
\Leftrightarrow \quad & \frac{\delta\lambda + (1-\delta)\psi_{h,h+1}\psi_{h+1,t+1}}{\delta\lambda + (1-\delta)\psi_{h+1,t+1}} \geq \delta\lambda + (1-\delta)\psi_{h,h+1} \\
\Leftrightarrow \quad & \frac{\delta\lambda + (1-\delta)\psi_{h,h+1}\psi_{h+1,t+1}}{\delta\lambda + (1-\delta)\psi_{h+1,t+1}} \geq \frac{\delta\lambda + (1-\delta)\psi_{h,h+1}\psi_{h+1,h+2}}{\delta\lambda + (1-\delta)\psi_{h+1,h+2}} \geq \delta\lambda + (1-\delta)\psi_{h,h+1},
\end{aligned}
$$

where the last line follows because $\psi_{h+1,h+2} \geq \psi_{h+1,t+1}$ for ages $t > h$ and because $\psi_{h,h+1} < 1$. The inequality is generally weak. It holds with equality for $\delta = 0$ (RE agent) and with strict inequality for $\delta \in (0,1), \lambda \in [0,1]$ (hence also for the QHD model, cf. Section 3.2) and for $\delta = 1, \lambda \in (0,1)$ (QHD model w/o mortality risk, cf. Section 3.2).

3. Finally, we establish that the sufficient condition (35) always holds in our model. To see this define the distance function

$$
D(\lambda) \equiv \nu_{h,h+2} - \nu_{h,h+1}\nu_{h+1,h+2}
$$

We next show that $D(\lambda = 0) > 0$ and $D(\lambda = 1) > 0$ whereby the relative magnitudes depend on the exact parametrization of survival risk. We then show that $D(\lambda)$ is an inverse-u shaped so that $D(\lambda) = 0$ is not possible for $\lambda \in (0,1)$. Observe that $D(\lambda)$ writes as

$$
\begin{aligned}
D(\lambda) &= \delta\lambda + (1-\delta)\psi_{h,h+2} - \left(\delta\lambda + (1-\delta)\psi_{h,h+1}\right)\left(\delta\lambda + (1-\delta)\psi_{h+1,h+2}\right) \\
&= \delta\lambda + (1-\delta)\psi_{h,h+2} - \left((\delta\lambda)^2 + \delta(1-\delta)\lambda(\psi_{h,h+1} + \psi_{h+1,h+2}) + (1-\delta)^2\psi_{h,h+2}\right).
\end{aligned}
$$

Next, evaluate $D(\lambda)$ at $\lambda = 0$ to get

$$
D(\lambda = 0) = \delta(1-\delta)\psi_{h,h+2} > 0,
$$

and at $\lambda = 1$:

$$\begin{aligned} D(\lambda = 1) &= \delta + (1-\delta)\psi_{h,h+2} - \left(\delta^2 + \delta(1-\delta)\left(\psi_{h,h+1} + \psi_{h+1,h+2}\right) + (1-\delta)^2\psi_{h,h+2}\right) \\ &= \delta(1-\delta) + \delta(1-\delta)\psi_{h,h+2} - \delta(1-\delta)\left(\psi_{h,h+1} + \psi_{h+1,h+2}\right) \\ &= \delta(1-\delta)\left(1 - \psi_{h,h+1}\right) > 0. \end{aligned}$$

Also notice that

$$\begin{aligned} D_\lambda &= \delta - 2\delta^2\lambda - \delta(1-\delta)\left(\psi_{h,h+1} + \psi_{h,h+2}\right) \\ &= \delta\left(1 - 2\delta\lambda - (1-\delta)\left(\psi_{h,h+1} + \psi_{h,h+2}\right)\right) \end{aligned}$$

so that

$$D_\lambda > 0 \qquad \Leftrightarrow \qquad \lambda < \frac{1 - (1-\delta)\left(\psi_{h,h+1} + \psi_{h,h+2}\right)}{2\delta}.$$

Furthermore, observe that

$$D_{\lambda\lambda} = -2\delta^2 < 0.$$

Hence, $D(\lambda) \neq 0$ for $\delta > 0, \lambda \in [0,1]$.

□

Proof of Theorem 3. We again simplify notation and generally write $m_t^{i,h}, c_t^{i,h}$ for all $i \in \{r,s,n\}$, whereby $m_t^{i,h} = m_t^i, c_t^{i,h} = c_t^i$ for $i \in \{r,s\}$. The utility function in period h writes as

$$U_h = u(c_h^{i,h}) + \sum_{t=h+1}^{T} \rho^i(h,t)u(c_t^{i,h}). \tag{74}$$

The consumption policy in all future periods $t > h$ is given by

$$c_t^{i,h} = m_t^{i,h} w_t$$

where

$$w_t = \left(w_{t-1} - c_{t-1}^{i,h}\right) R = \left(w_h - c_h^{i,h}\right) R^{t-h} \prod_{j=h+1}^{t-1} (1 - m_j^{i,h}),$$

where $\prod_{j=h+1}^{h}(1-m_j^{i,h})=1$. Using this in (74) we get

$$U_h = u(c_h) + \sum_{t=h+1}^{T} \rho^i(h,t) u \left(m_t^{i,h} \left(w_h - c_h^{i,h} \right) R^{t-h} \prod_{j=h+1}^{t-1} (1-m_j^{i,h}) \right).$$

Assuming CRRA the above rewrites as

$$U_h = \frac{1}{1-\theta} \left(c_h^{i,h^{1-\theta}} + \left(w_h - c_h^{i,h} \right)^{1-\theta} \sum_{t=h+1}^{T} \rho^i(h,t) \left(R^{t-h} m_t^{i,h} \prod_{j=h+1}^{t-1} (1-m_j^{i,h}) \right)^{1-\theta} \right)$$

Taking first-order conditions we get

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

hence

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

□

Proof of Proposition 8. Step 1. Obviously, $\nu_{0,1}$ and $(\beta\nu_{1,2})^{\frac{1}{\theta}}$ are strictly increasing in λ for all θ. To prove that the LHS of (59) is strictly increasing in λ for $\theta \geq 1$ it therefore suffices to show that

$$\frac{\left(1+(\beta\nu_{1,2})^{\frac{1}{\theta}}\right)}{\left(1+(\beta\nu_{1,2})^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}}} \tag{75}$$

as well as $\frac{\nu_{0,2}}{\nu_{1,2}}$ are increasing in λ for $\theta \geq 1$.

Step 2. To see that (75) is increasing in λ for $\theta \geq 1$, observe that

$$\begin{aligned}
\frac{d}{dx}\frac{\left(1+x^{\frac{1}{\theta}}\right)}{\left(1+x^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}}} &\geq 0 \\
&\Leftrightarrow \\
\frac{1}{\theta}x^{\frac{1}{\theta}-1}\left(1+x^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}} &\geq \frac{1}{\theta}\left(1+x^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}-1}\frac{1}{\theta}x^{\frac{1}{\theta}-1}\left(1+x^{\frac{1}{\theta}}\right) \\
&\Leftrightarrow \\
1 &\geq \frac{1}{\theta},
\end{aligned}$$

which gives the desired result.

Step 3. To see that $\frac{\nu_{0,2}}{\nu_{1,2}}$ is increasing in λ, observe that

$$\begin{aligned}
\frac{d}{d\lambda}\frac{\nu_{0,2}}{\nu_{1,2}} &\geq 0 \\
&\Leftrightarrow \\
\delta\left(\delta\lambda+(1-\delta)\,\psi_{_{1,2}}\right) &\geq \delta\left(\delta\lambda+(1-\delta)\,\psi_{_{0,1}}\psi_{_{1,2}}\right) \\
&\Leftrightarrow \\
1 &\geq \psi_{_{0,1}},
\end{aligned}$$

which is satisfied for all θ.$\square$

Numerical Analysis of Conjecture 1.

1. To numerically establish the first part of the conjecture, we first compute λ_s from equation (65) for a multidimensional grid of parameter values for $\mathcal{G} = \theta \in (0,1) \otimes \psi_{0,1} \in (0,1) \otimes \psi_{1,2} \in (0,1) < \psi_{0,1} \otimes \beta \in (0,1) \otimes \delta \in (0,1)$. Next, for each point on this multidimensional grid and the corresponding root $\lambda_s^{\star}(\theta, \psi_{0,1}, \psi_{1,2}, \beta, \delta)$ we compute on a grid of $\lambda \in (0, \lambda_s^{\star}(\cdot))$ the LHS of (59). There is no λ on this grid for which the MPC of the sophisticated RDU agent is lower than the MPC of the RE agent.

2. In terms of underestimation, recall from Corollary 6 that $\lambda < \lambda^n$ is only a sufficient condition for undersaving of the sophisticated agent. There may therefore be a parameter constellation such that $\lambda^s > \psi_{0,1}$. In our numerical analysis leading to Conjecture 1 we could, however, not find any parameter combinations

such that $\lambda^s > \psi_{0,1}$.

Proof of Proposition 9. From (57) we need

$$\begin{aligned}
& \nu_{0,1} + \beta\nu_{0,2} < \psi_{0,1} + \beta\psi_{0,2} \\
\Leftrightarrow \quad & \delta\lambda + (1-\delta)\psi_{0,1} + \beta\left(\delta\lambda + (1-\delta)\psi_{0,2}\right) < \psi_{0,1} + \beta\psi_{0,2} \\
\Leftrightarrow \quad & \lambda < \frac{\psi_{0,1} + \beta\psi_{0,2}}{1+\beta}
\end{aligned}$$

from which (68) immediately follows. $\square$