ECONSTOR

WWW.ECONSTOR.EU

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationszentrum Wirtschaft
The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics

Groneck, Max; Ludwig, Alexander; Zimper, Alexander

Conference Paper

Ambiguous Survival Beliefs and Hyperbolic Discounting in a Life-Cycle Model

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Belief Formation, No. C15-V2

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Groneck, Max; Ludwig, Alexander; Zimper, Alexander (2013) : Ambiguous Survival Beliefs and Hyperbolic Discounting in a Life-Cycle Model, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Belief Formation, No. C15-V2

This Version is available at:
http://hdl.handle.net/10419/79878

Nutzungsbedingungen:
Die ZBW räumt Ihnen als Nutzerin/Nutzer das unentgeltliche, räumlich unbeschränkte und zeitlich auf die Dauer des Schutzrechts beschränkte einfache Recht ein, das ausgewählte Werk im Rahmen der unter
→ http://www.econstor.eu/dspace/Nutzungsbedingungen
nachzulesenden vollständigen Nutzungsbedingungen zu vervielfältigen, mit denen die Nutzerin/der Nutzer sich durch die erste Nutzung einverstanden erklärt.

Terms of use:
The ZBW grants you, the user, the non-exclusive right to use the selected work free of charge, territorially unrestricted and within the time limit of the term of the property rights according to the terms specified at
→ http://www.econstor.eu/dspace/Nutzungsbedingungen
By the first use of the selected work the user agrees and declares to comply with these terms of use.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der Leibniz-Gemeinschaft

Ambiguous Survival Beliefs and Hyperbolic Discounting in a Life-Cycle Model

Max Groneck* Alexander Ludwig†
Alexander Zimper‡

Preliminary and Incomplete.

June 26, 2013

Abstract

On average, young people underestimate whereas old people overestimate their chances to survive into the future. We employ a subjective survival belief model proposed by Ludwig and Zimper (2013) which can replicate these patterns. The model is compared with hyperbolic discounting within a standard life-cycle setting of consumption and savings. We show theoretically that the first order conditions of the ambiguous survival belief model closely resemble the generalized Euler-equation from the hyperbolic discounting model. In the numerical section it is shown that the subjective survival belief model simultaneously leads to undersaving at younger ages and high asset holdings and little dissaving of the elderly. The model can thus replicate two important empirical facts of the life-cycle literature at once which is not possible with the standard hyperbolic discounting model.

JEL Classification: D91, D83, E21.
Keywords: Life-cycle hypothesis, hyperbolic discounting, subjective survival rates

*Max Groneck, CMR, University of Cologne, Netspar (groneck@wiso.uni-koeln.de)
†Alexander Ludwig, CMR, University of Cologne, MEA, Netspar (ludwig@wiso.uni-koeln.de)
‡Alexander Zimper, University of Pretoria (alexander.zimper@up.ac.za)

1 Introduction

In standard life-cycle consumption and saving models with uncertain lifetime, intertemporal decision making is inter alia governed by the discounting factor and the probability to survive into the future both of which form the so called effective discount factor. Since the pioneering work of Samuelson (1937) it has become common in the literature to assume time-invariant discounting, i.e. exponential discounting functions. In addition, when testing these models it is common to assume that individual's beliefs are given as objective probability distributions.[1] This well established modeling choice has been criticized with the apperance of empirical and experimental studies showing first, that people seem to exhibit time-inconsistent behavior and second, that on average people's subjective survival perceptions systematically deviate from objective life-table data.

In this paper we compare two different approaches that modify established models to account for these insights. We adopt a decision theoretic model of subjective survival belief formation proposed by Ludwig and Zimper (2013) which takes individuals' ambiguity attitudes into account. The model is based on Choquet expected utility (CEU) theory. We merge this subjective survival belief model into a standard stochastic life-cycle setting. The CEU model is compared to the well-known (quasi-)hyperbolic discounting model which modifies the functional form of the discount function but still takes survival beliefs as objectively given.[2] We compare the outcomes of two models with respect to saving and asset accumulation behavior.

Experimental studies have detected different forms of time-inconsistency, in particular a conflict between the long-run desire to be patient and the short-run desire for instantaneous gratification which is at odds with commonly used effective discount functions.[3] This has led to alternative discounting rules of which the (quasi-)hyperbolic or geometric discounting model has become very popular and which was propagated by the work of David Laibson.[4] Quasi-hyperbolic discounting (referred to as the HD model) introduces a short-run discount factor between the present and the first future period which is distinct—usually lower—from that between the future period and its successor, implying consumers as being short-run impatient. It has been shown that hyperbolic discounting can accommodate relevant empirical findings on household saving behavior which proved to be puzzling for rational expectations life-cycle models, in particular, time-inconsistent

[1] See e.g. Hurd (1989) Hubbard, Skinner, and Zeldes (1995), Imrohoroglu, Imrohoroglu, and Joines (1995), or De Nardi (2004) among others.

[2] In our companion paper Groneck, Ludwig, and Zimper (2013) we highlight the importance of subjective survival beliefs in a life cycle setting compared to a rational expectations model using exponential discounting and objective survival rates from the life-tables.

[3] See Ainslie (1992) and Loewenstein and Thaler (1989)

[4] See, e.g., Laibson (1997), Laibson (1998).

Another famous alternative that has been proposed to address the conflict between actions and intentions are the so called self-control preferences proposed by Gul and Pesendorfer (2001) and Gul and Pesendorfer (2004).

savings behavior and undersaving for retirement.[5]

Much less attention has been devoted to the survival probability as the second part of the effective discount factor, where standard quantitative studies make use of objective survival beliefs in. Only recently, researchers have focused on *subjective* survival probabilities and their deviations from life tables. In the Health and Retirement Study (HRS), respondents are explicitly asked for their belief to survive to some future age.[6] The data shows a systematic bias: on average younger people strongly underestimate their (relatively high) probability to survive to some target age. At the same time older people strongly overestimate their lower survival probability. In the CEU-model that we analyze, we follow Ludwig and Zimper (2013) and describe ambiguous survival beliefs as neo-additive capacities (Chateauneuf, Eichberger, and Grant 2007). Neo-additive capacities are non-additive probability measures based on Choquet decision theory (Schmeidler 1986, 1989, Gilboa 1987).[7] Biases of subjective beliefs from some additive probability measure are induced by two parameters only. One (ambiguity) parameter measures the degree of confidence a decision maker has in some additive probability measure. A second (optimism) parameter measures relative over/under-estimation through which this ambiguity is resolved. This approach is in line with empirical evidence in the decision theoretic literature of subjective probability weighting due to likelihood insensitivity, for example (cf. Wakker 2004, 2009). We estimate parameters of ambiguous survival beliefs and demonstrate that these calibrated beliefs can account for the empirical facts as elicited in the HRS.

Our main results can be summarized as follows: we show that (i) the CEU model can be interpreted as a microfounding of the hyperbolic discounting model and (ii) for the young and middle aged the CEU model has similar implications for consumption and saving behavior but (iii) due to optimism at older ages, the CEU model can account for the important stylized fact on old-age saving, namely that the elderly hold on to their assets and dissave less than prescribed by a standard life-cycle model (old-age-dissaving puzzle).

In Sections 2 and 3 we present the CEU and the HD model in more detail while we incorporate the two models into the life-cycle setting in section 4. We present results for both the sophisticated and naive agent type where the former is aware of her dynamic inconsistency while the latter is not. We show that our CEU model of subjective

[5]See, among others, Laibson, Repetto, and Tobacman (1998), Harris and Laibson (2001)Angeletos et al. (2001). and Krusell and Smith (2003).

[6]The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. In this survey people are asked about their subjective probability assessment to survive from some interview age up to a specific target age. Target age is mostly 10 to 15 years in advance, see, e.g., Ludwig and Zimper (2013) for details.

[7]Thus, the agent with ambiguous survival beliefs is labeled as the Choquet Expected Utility (CEU) maximizer.

survival beliefs closely resembles the hyperbolic discounting model in some respects but also differs in important aspects. Generally, both the CEU and the HD model share the similarity to effectively discount stronger for two subsequent events in the near future than subsequent events in the far future which are discounted less strongly. We theoretically derive the Euler-equations of different types of dynamically inconsistent agents.[8] The Euler-equation of the sophisticated CEU agent closely resembles the "generalized Euler-equation" derived by Harris and Laibson (2001), but it is extended by an additional adjustment factor. An important difference between the CEU and the HD model comes from the fact that the short term discount factor in the HD model is constant for all ages, while the ambiguity attitude in the CEU model is age dependent.

In a simplified version of our model in Section 5 which provides guidance for our quantitative analysis we show that we can expect naive CEU households with subjective survival beliefs (i) to save less than originally planned at young age if they are "moderately optimistic" about their future survival in old age, (ii) to save less than under rational expectations if they are "sufficiently pessimistic" at young age and (iii) to have higher asset holdings than their rational expectations counterparts if they are optimistic in old age for sufficiently many periods. On the contrary, we show that HD households can only explain time-inconsistent behavior and undersaving. For sophisticated agents theoretical results are not clear cut.

To gather quantitative results we return to the full model and calibrate it to simulate household decisions in Section 6. We show that the misperception of lifespan risk indeed simultaneously adds to existing explanations for three stylized facts in the data on savings behavior, if we assume agents to be naive: On average, naive CEU agents with subjective survival beliefs at working age have a saving rate of 21.5% compared to a rational expectations model with an average saving rate of 23.0%. In addition, the realized saving rate is 2.8 percentage points lower than what the CEU agent at age 20 actually planned to save. The results correspond to ample empirical evidence on dynamically inconsistent behavior. Survey evidence indicates that people save less for retirement than actually planned and they save less than they think they should.[9] Simultaneously, the CEU model is able to account for higher asset holdings than prescribed by a standard life-cycle framework with rational consumers.[10] The CEU model predicts average asset holdings at age 85 (95) of 46.4% (23.5%) of the assets at age 65 while the respective

[8] Time inconsistency refers to decision makers who have time-dependent preferences, i.e. the preferences of the presence contradict their own preferences at a later date.

[9] Bernheim and Rangel (2007), and Laibson et al. (1998) quote numerous studies indicating self-reported mistakes in terms of private saving decisions for retirement. See also Bernheim (1998) and Choi et al. (2006) as well as Munnell et al. (2010) for studies on undersaving.

[10] See Hurd and Rohwedder (2010) and De Nardi, French, and Jones (2010) for evidence.

There are numerous extensions of the standard life-cycle model which aim at explaining this feature of the data. The two main explanations for large assets holdings late in life are bequest motives (Hurd 1989, Lockwood 2012), and precautionary saving due to possibly large health expenditures (De Nardi, French, and Jones 2010).

values for rational agents are substantially lower. In contrast, the naive HD agent does also exhibit undersaving with a savings rate of $19,7\%$. On the contrary, the naive HD agent exhibit much lower asset holdings at age 85 (95) of 26.2% (3.9%) of the assets at age 65 which are even lower than those for the RE agents.

If one assumes sophisticated agents, undersaving is for CEU and HD agents is even more pronounced. But the aforementioned difference of old-age asset holdings of the CEU and the HD agents relative to RE agents disappears quantitatively. Still, assets of the sophisticated agent are slightly higher than for the RE agent at very old ages. In a concluding Section 8 we discuss further room for research.

2 Ambiguous Survival Beliefs

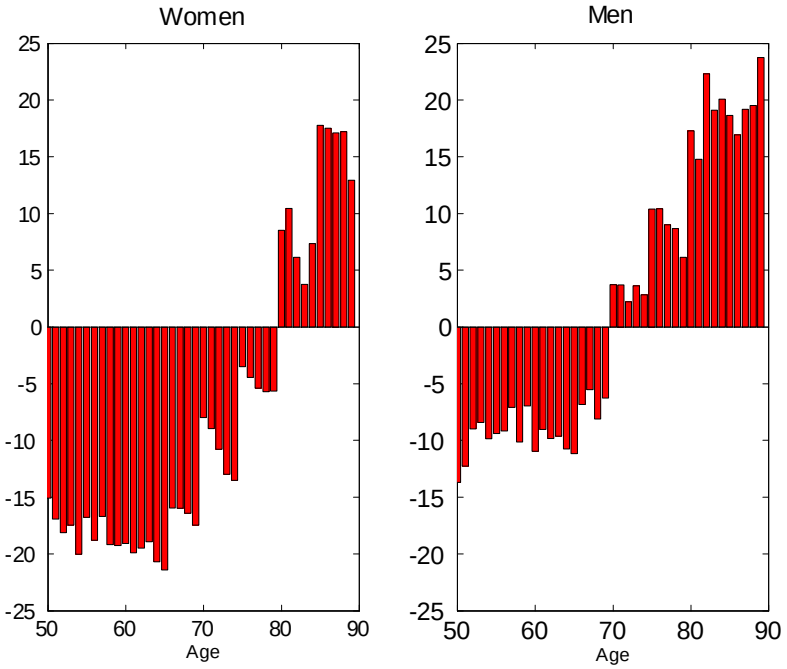
The model of ambiguous survival belief formation is developed in Ludwig and Zimper (2013) to accommodate the HRS data depicted in Figure 1. The figure shows average age-specific biases in survival beliefs—the difference between the respective average subjective belief and the average estimated objective data—for three waves of the HRS between 2000 and 2004. We observe that relatively "young"—younger than age 65-70—respondents underestimate whereas relatively "old"—above age 70—respondents overestimate their chances to survive into the future. Pessimistic biases of young respondents are stronger for women and optimistic biases for old respondents are stronger for men. For example, a 65 year old women underestimates her objective probability to become 80 years by about 20 percentage points. Respondents between ages 85 and 89 in the sample exhibit an average overestimation by about 15 to 20 percentage points. In addition, optimism further increases the older the agents get. Similar differences between subjective beliefs and objective data can be observed in most European countries according to data from the Survey of Health, Ageing and Retirement in Europe (SHARE), cf. Peracchi and Perotti (2010), see also Ludwig and Zimper (2013) for a discussion.

To explain the patterns observed in the HRS we follow Ludwig and Zimper (2013) who develop a closed-form model of Bayesian learning under ambiguity which gives rise to a parsimonious notion of *ambiguous survival beliefs*. The approach constitutes conditional neo-additive capacity in the sense of Chateauneuf et al. (2007) which is updated in accordance with the Generalized Bayesian update rule.[11] In contrast to standard additive probability measures, the neo-additive survival probability can replicate the patterns of Figure 1 because they do not necessarily converge through Bayesian updating to the objective survival probabilities $\psi_{k,t}$.

In the parsimonious version of the model in Ludwig and Zimper (2013), ambiguous survival beliefs are a weighted average of a standard additive probability measure and

[11] For the formal definitions of these decision theoretic terms see Appendix A and references therein.

Figure 1: Relative difference of subjective survival probabilities and cohort data



Notes: This graph shows deviations in percentage points of subjective survival probabilities from objective data. Objective survival rates are based on cohort life table data. Future objective data is predicted with the Lee and Carter (1992) procedure.
Source: Own calculations based on HRS, Human Mortality Database and Social Security Administration data.

the degree of optimism, respectively pessimism, by which the decision maker resolves her ambiguity towards objective information. The respective weight represents the degree of ambiguity—or the lack of confidence—that the decision maker has in the additive probability measure. Hence, there are two psychological biases, relative optimism and ambiguity. The model is information driven so that ambiguity—and thus the weight on relative optimism—increases in the amount of information. It is assumed that information rises with age to the effect that ambiguity increases in the agent's actual age (i.e., when the objective survival probability decreases). For the reader's benefit in Appendix A we review the decision theoretic preliminaries and the parsimonious learning model of Ludwig and Zimper (2013) as well as how the model is merged into a life-cycle setting as shown in Groneck, Ludwig, and Zimper (2013). In what follows, we only restate the model's parsimonious characterization of ambiguous survival beliefs.

Fix some $T \geq h \geq 0$ with the interpretation that the agent perceives it as possible to live until the end of period T whereas she perceives it as impossible to live longer than T. Denote by $\delta \in [0,1]$ an initial degree of ambiguity (or degree of likelihood insensitivity). $\lambda \in [0,1]$ denotes a psychological bias parameter which measures whether the agent resolves her ambiguity through over- or rather through under-estimation of the

true probability. In this framework, Ludwig and Zimper (2013) derive the following result which we build on:

Observation 1 *(Ludwig and Zimper (2013)). The h-old agent's age-dependent ambiguous belief to survive from age k with $h \leq k < T$ to target age t with $k < t \leq T$ is given by*

$$\nu_{k,t}^{h} = \delta_h \cdot \lambda + (1 - \delta_h) \cdot \psi_{k,t} \tag{1}$$

with

$$\delta_h = \frac{\delta}{\delta + (1-\delta) \cdot \frac{1}{1+\sqrt{h}}} \tag{2}$$

for age-independent parameters $\delta, \lambda \in [0, 1]$.

The h-old agent's belief to survive to some target age t is thus formally described as an age-dependent weighted average of the objective survival probability $\psi_{h,t}$ with weight $1 - \delta_h$ and the psychological bias parameter λ with weight δ_h. For $\delta = 0$ we have for all h that $\nu_{k,t}^{h} = \psi_{k,t}$ so that all ambiguous survival beliefs reduce to objective survival probabilities and the standard rational expectations approach is nested as a special case.

For any $\delta > 0$, the dynamics of the model imply that agents exhibit more pronounced ambiguity attitudes with increasing age. This feature captures the intuitive notion that, as the objective risk of survival becomes less likely, agents attach less and less certainty to this objective probability. According to our estimates of $\{\delta, \lambda\}$ presented in Section 6, objective survival probabilities $\psi_{k,t}$ decrease with age to values lower than λ. The model's convergence property hence implies that survival rates are overestimated eventually even when the initial degree of likelihood insensitivity, δ, is low.[12]

The model of neo-additive probabilities imply linear probability weighting functions which can be interpreted as an approximation of the inversely S-shaped probability weighting arising in cumulative prospect theory (CPT), cf. Tversky and Kahneman (1992). In cumulative prospect theory, it is standard to assume a single-parameter functional form for the probability weighting function. Applied to (age dependent) survival beliefs such a functional form, as, e.g. used by Wu and Gonzalez (1996), is given by

$$\varpi_h\left(\psi_{h,t}, \xi_h\right) = \frac{\left(\psi_{h,t}\right)^{\xi_h}}{\left[\left(\psi_{h,t}\right)^{\xi_h} + \left(1 - \psi_{h,t}\right)^{\xi_h}\right]^{\frac{1}{\xi_h}}} \tag{3}$$

For $\xi_h = 1$, we have $\varpi_h\left(\psi_{h,t}, \xi_h\right) = \psi_{h,t}$. Decreasing ξ_h means increasing curvature of the probability weighting function so that the inverse-S becomes more pronounced. The neo-

[12] If we were to assume that households do not have any memory to the effect that $\delta_h = \delta$ for all $h = 0, \ldots, T$, we would get qualitatively similar results. Quantitatively this would however imply that the degree of ambiguity would be substantially higher for many ages than with increasing δ_h. We will come back to this aspect in the interpretation of our main results below.

additive capacity (cf. Appendix A.1) used as our subjective survival belief for the CEU model can be seen as a linear approximation to the probability weighting function (3), $\tilde{\varpi}_h = \beta_0 + \beta_1 \psi_{k,t}$ with $\beta_0 = \lambda\delta_h$ and $\beta_1 = (1 - \delta_h)$. In Groneck, Ludwig, and Zimper (2013) we show that the estimates of our CEU model are well in line with conventional estimates of the probability weighting function.

3 (Quasi-)Hyperbolic Discounting

The idea of hyperbolic discounting dates back to Strotz (1956) and Pollak (1968). In our setting, the hyperbolic discounting (HD) model takes the survival rates as objectively given and introduces a time-varying discount function. The modification of the standard exponential discounting function aims to capture observed behavioral anomalies. More precisely, the expected utility model cannot account for the fact that people are more sensitive to a given time delay if it occurs closer to the present than if it occurs farther in the future, see Loewenstein and Prelec (1992). The continuous hyperbolic discount function introduced by Loewenstein and Prelec (1992) is defined as

$$\vartheta_h(\alpha, \varrho) = (1 + \alpha \cdot h)^{-\frac{\varrho}{\alpha}}, \tag{4}$$

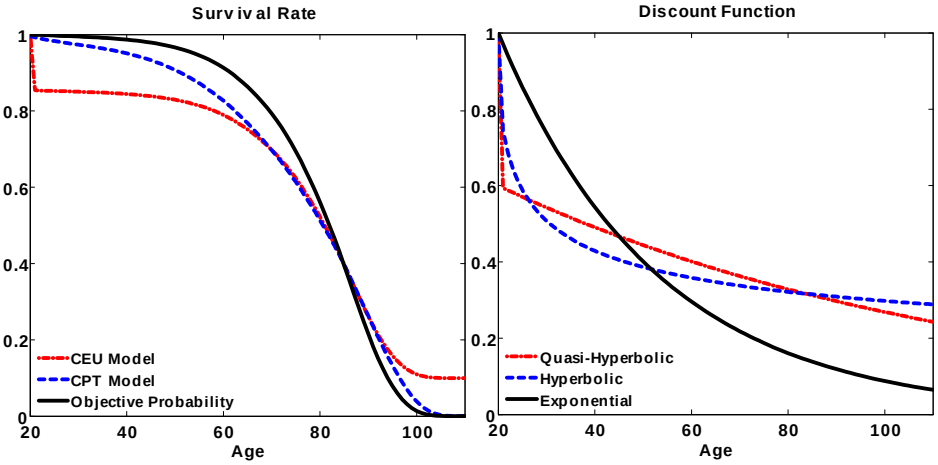
where α and ϱ are parameters and h is age.

The quasi-hyperbolic discounting model first proposed by Phelps and Pollak (1968) is an approximation of the hyperbolic function (4) and takes values $\left\{1, \gamma\beta, \gamma\beta^2, \delta\beta^3, ...\gamma\beta^T\right\}$ where $0 < \gamma, \beta < 1$, and γ is the short term discount factor. In essence, this form of discounting puts additional weight on the actual period, implying a downward jump of the discount function from period h to $h+1$. As will be shown, this feature is also generated in our ambiguous survival model.

Both the CEU and the HD model modify the effective discount factor in the life-cycle model in a similar fashion. The CEU model modifies the assumption on survival beliefs to better match the above mentioned empirical findings while we stay in line with the literature by using the time-invariant discount factor β. In contrast, the hyperbolic discounting (HD) model takes the survival rates as objectively given and introduces a time varying discount function accounting for the fact that people have an additional present bias. The similarity between the quasi-hyperbolic discounting model and the CEU model of subjective survival belief formation is highlighted in Figure 2.

On the right panel the hyperbolic and the quasi-hyperbolic discount function is plotted against the exponential discount functions while on the left panel the subjective beliefs according to the CEU and the CPT model are plotted against the estimated objective unconditional survival probabilities. Both the subjective survival rate and the quasi-

Figure 2: Survival Rates and Discount Functions



Notes: The first period in both panels is assumed to be the present, thus the survival rate and the discount factor is 1. We use the following values: $\gamma = 0.6$ and $\beta = 0.99$ is set for the quasi-hyperbolic, $\gamma = 1$ and $\beta = 0.97$ for exponential discounting. Both the values for the probability function and for the hyperbolic function are chosen to minimize the squared distance to the subjective belief function and the quasi-hyperbolic function, respectively. The resulting parameters are $\alpha = 0.95$ and $\varrho = 0.26$ for the hyperbolic discounting function (4) and $\xi = 0.74$ for the probability weighting function (3).

hyperbolic function exhibit an initial downward jump and they are flatter generally than their "objective" counterpart. In a life-cycle context with expected lifetime utility, both models put less weight on utility in the near future relative to rational expectations. But at the same time, they put relatively more weight on future utility if the event is further away.[13] Also, both the CEU model and the quasi-hyperbolic model are approximations of other models: The CEU is a linear approximation of the CPT model and the quasi-hyperbolic model is a discrete time version of the continuous hyperbolic model which is much smoother.

The close similarities of the two models with respect to their modifications to the effective discount function is the motivation for the model comparison in this paper. Despite the close similarities suggested by Figure 2 the two models have quite different consumption and saving behavior, as will be shown in Section 7.

4 The Model

We incorporate two different decision makers into an otherwise standard stochastic life-cycle framework. The basic model setup borrows from our companion paper Groneck, Ludwig, and Zimper (2013).

[13] It is important to note that this result only holds if we assume a higher long-term discount factor β for the quasi-hyperbolic model relative to exponential discounting, see further below for details.

4.1 Demographics

We consider a large number of ex-ante identical agents (=households). Households become economically active at age (or period) 0 and live at most until age T. The number of households of age t is denoted by N_t. Population is stationary and we normalize total population to unity, i.e., $\sum_{t=0}^{T} N_t = 1$. Households work full time during periods $1, \ldots, t_r - 1$ and are retired thereafter. The working population is $\sum_{t=0}^{t_r-1} N_t$ and the retired population is $\sum_{t=t_r}^{T} N_t$.

We refer to age $h \leq t$ as the planning age of the household, i.e., the age when households make their consumption and saving plans for the future. At ages $h = 1, \ldots, T$, households face objective risk to survive to some future period t. We denote corresponding objective survival probabilities for all in-between periods k, $h \leq k < t$, by $\psi_{k,t}$ where $\psi_{k,t} \leq 1$ for all $t \leq T$ and $\psi_{k,t} = 0$ for $t = T + 1$. Population dynamics are correspondingly given by $N_{t+1} = \psi_{t,t+1} N_t$, for N_0 given.

4.2 Endowments

There are discrete shocks to labor productivity in every period $t = 0, 1, ..., , t_r - 1$ denoted by $\eta_t \in E$, E finite, which are i.i.d. across households of the same age. The reason for stochastic labor productivity in our model is to impose discipline on calibration. For sake of comparability, our fully rational model should feature standard elements as used in numerous structural empirical studies, as, e.g., by Laibson et al. (1998), Gourinchas and Parker (2002) and several others. By $\eta^t = (\eta_1, \ldots, \eta_t)$ we denote a history of shocks and $\eta^t \mid \eta^h$ with $h \leq t$ is the history $(\eta_1, \ldots, \eta_h, ..., \eta_t)$. Let E be the powerset of the finite set E and E^∞ be the σ-algebra generated by $\mathsf{E}, \mathsf{E}, ...$ We assume that there is an objective probability space $(\times_{t=0}^{\infty} E, \mathsf{E}^\infty, \pi)$ such that $\pi_t(\eta^t \mid \eta^h)$ denotes the probability of η^t conditional on η^h.

We follow Carroll (1992) and assume that one element in E^w is zero (zero income). Accordingly, $\pi_t(\eta^t \mid \eta^h)$ reflects a (small) probability to receive zero income in period t. This feature gives rise to a self-imposed borrowing constraint and thereby to continuous policy functions. Continuity is convenient when we model a sophisticated agent. By thereby avoiding technicalities as addressed in Harris and Laibson (2001) we keep our analysis focused. As the zero income probability is small, results are not affected much by this assumption. In addition, we consider a fixed age-specific labor productivity profile, ϕ_t. w denotes age-independent gross wages.

After retirement at age t_r households receive a lump-sum pension income, b. Retirement income is modeled in order to achieve a realistic calibration. Without a pension system, the old-age saving motive would lead to unrealistic saving behavior. For the same technical reason mentioned above—continuity of policy functions—we assume that household's assume a small probability of government default on pension payments each

period. Again this assumption is made for sake of convenience only.[14] Pension contributions are levied at contribution rate τ.

Collecting elements, income of a household of age t is given by

$$y_t = \begin{cases} \eta_t \phi_t w (1-\tau) & \text{for } t < t_r \\ \eta_t b & \text{for } t \geq t_r. \end{cases}$$

During retirement, $\eta_t \in E^r = [1, 0]$ where the default probability of the government on its pension obligations is assumed to be low.

The interest rate is fixed at r. There are no annuity markets, an assumption which can be justified by the observed small size of private annuity markets.[15] We assume a fixed zero borrowing constraint. We define cash-on-hand as $x_t \equiv a_t (1+r) + y_t$ so the budget constraint writes as

$$x_{t+1} = (x_t - c_t)(1+r) + y_{t+1} \geq 0 \tag{5}$$

Define total income as $y_t^{tot} = y_t + r a_t$, saving as $s_t = y_t^{tot} - c_t$ and gross savings as assets tomorrow, a_{t+1}.

4.3 Government

We assume a pure PAYG public social security system. We denote by χ the net pension benefit level, i.e., the ratio of pensions to net wages. The government budget is assumed to be balanced each period and is given by

$$\tau w \sum_{t=0}^{t_r-1} \phi_t N_t = b \sum_{t=t_r}^{T} N_t = \chi (1-\tau) w \sum_{t=t_r}^{T} N_t. \tag{6}$$

Accidental bequests—arising because of missing annuity markets—are taxed away at a confiscatory rate of 100%.[16]

4.4 Preferences

Denote by $u(c_t)$ the agent's strictly increasing utility from consumption at age t whereby we assume that the agent is strictly risk-averse, i.e., $u'(c_t) > 0$ and $u''(c_t) < 0$. More specifically, given the productivity shock history η^h, denote by $\mathbf{c} \equiv (c_h, c_{h+1}, c_{h+2}...)$ a shock-contingent consumption plan such that the functions c_t, for $t = h, h+1, ...$, assign

[14] While households assume this small probability when determining their optimal behavior, default never happens in the forward simulation of our model.

[15] See Friedman and Warshawsky (1990). Observe that pessimistic survival beliefs extenuate the annuity puzzle.

[16] Revenue from this source is used for government consumption which is otherwise neutral.

to every history of shocks $\eta^t|\eta^h$ some amount of period t consumption.

Expected utility of an h-old agent from consumption in period $t > h$ contingent on the observed history of productivity shocks η^h is given as

$$E_t\left[u\left(c_t\right), \pi\left(\eta^t|\eta^h\right)\right] = \sum_{\eta^t|\eta^h} u\left(c_t\right) \pi\left(\eta^t|\eta^h\right).$$

As a shortcut we denote the expectation operator with respect to the productivity shock η^t in period t by $E_t\left[\bullet\right]$.

The agent's expected utility from consumption plan c is given by

$$E\left[U\left(\mathbf{c}\right), \nu^h\right] = u(c_h) + \sum_{t=h+1}^{T} \nu_{h,t}^h \cdot \varphi_t^h \cdot E_h\left[u\left(c_t\right)\right] \tag{7}$$

where $\nu_{h,t}^h$ is the (subjective) probability to survive from period h to period t and φ_t^h is the discount function. The following sections will specify preferences for the CEU model and the HD model.

Choquet Expected Utility (CEU)

While we model income risk in the standard objective EU way, we model uncertainty about life-expectancy in terms of a CEU agent who holds neo-additive survival beliefs as stated in Observation 1. In Groneck, Ludwig, and Zimper (2013) we formalize utility maximization over life-time consumption with respect to neo-additive probability measures, here we restate the main result of Choquet expected utility.[17]

Observation 2 *(Groneck, Ludwig, and Zimper 2013) The Choquet Expected Utility (CEU) model with exponential discounting, i.e.* $\varphi_t^h = \beta^{t-h}$*, is given by*

$$E\left[U\left(\mathbf{c}\right), \nu^h\right] = u(c_h) + \sum_{t=h+1}^{T} \nu_{h,t}^h \cdot \beta^{t-h} \cdot E_h\left[u\left(c_t\right)\right] \tag{8}$$

where $\mathbf{c}$ *is the consumption plan and* $\nu_{h,t}^h$ *is the subjective belief to survive from age* h *to* $t \geq h$ *given by equation (1).*

To analyze the behavior of CEU agents in the life-cycle framework we make the following definition which will be used further below

Definition 1 (Moderate optimism) *A CEU household is* moderately optimistic *if* $\nu_{t,t+1}^t < \frac{\nu_{h,t+1}^h}{\nu_{h,t}^h}$ *for* $t \geq h$*, where* $\nu_{h,h}^h = 1$.

[17] See Appendix A.3.

Expected Utility with Quasi-Hyperbolic Discounting (HD)

In the quasi-hyperbolic model, the raw time discount factor β is replaced by the quasi-hyperbolic discount function given by

$$\varphi_t^h = \gamma\beta^{t-h}, \quad \forall\ t > h \tag{9}$$

Observation 3 *Expected utility with Quasi-Hyperbolic Discounting (HD) with objective probability measures, i.e.* $\nu_{h,t}^h = \psi_{h,t}$ *is given by*

$$E\left[U\left(\mathbf{c}\right)\right] = u(c_h) + \sum_{t=h+1}^{T} \psi_{h,t} \cdot \gamma \cdot \beta^{t-h} \cdot E_h\left[u\left(c_t\right)\right] \tag{10}$$

where $\mathbf{c}$ *is the consumption plan and* γ *is the short-term discount factor.*

In the numerical simulation in Section 7 we compare the life-cycle patterns with the standard rational expectations (RE) model. Here, the agent discounts exponentially and uses objective survival rates, i.e. $\varphi_t^h = \beta^{t-h}$ and $\nu_{h,t}^h = \psi_{h,t}$. In contrast to the standard RE model, which is dynamically consistent, both the CEU and the HD might lead to dynamically inconsistent decision making.

4.5 Naivety and Sequential Sophistication

In models with time inconsistency one has to take a stance on what information the agent has concerning the future. The literature dealing with time inconsistency distinguishes between naive and sophisticated agents, cf. Strotz (1955) or inter alia O'Donoghue and Rabin (1999) for procrastination models, both inducing time inconsistency.

Naifs are not aware of their time inconsistency and have the optimistic belief that their future "selves" will be acting in their interest. Naive agents construct consumption and saving plans that maximize lifetime utility at age h. Self h then implements the first action of that sequence expecting future selves to implement the remaining plan. But coming to the next period, self $h+1$ conducts her own maximization problem and implements actions that do not necessarily coincide with the plan of self h.

In contrast, sophisticates are fully aware of their time inconsistent behavior, cf., e.g., Angeletos et al. (2001). Sophisticates correctly predict that their own future selves will not be acting according to the preference of the current self. Thus, they take actions that seek to constrain future selves behavior (commitment devices). The CEU framework differs from the aforementioned models in being information rather than preference driven. We can still adopt the concept of naifs and sophisticates to our model.

In order to characterize optimal behavior, it is convenient to work with the recursive representation of the planning problem. We therefore next assume that income risk is

first-order Markov such that $\pi(\eta^t \mid \eta^{t-1}) = \pi(\eta^t \mid \eta_t)$. It is then straightforward to set up the recursive formulation of model (8) and (10) for the naive agents.

The dynamic programming problem of an agent at planning age h is defined as follows. For the CEU agent we have for the value function V_t^h of age $t \geq h$ viewed from planning age h

$$(CEU) \qquad V_t^h(x_t, \eta_t) = \max_{c_t, x_{t+1}} \left\{ u(c_t) + \beta \frac{\nu_{h,t+1}^h}{\nu_{h,t}^h} E_t \left[V_{t+1}^h \left(x_{t+1}, \eta_{t+1} \right) \right] \right\} \qquad (11)$$
$$s.t. \text{ eq. } (5)$$

where the subjective survival belief $\nu_{h,t}^h$ is given by equation (1) and $\nu_{h,h}^h = 1$.

For the hyperbolic agent we have

$$(HD) \qquad V_t^h(a_t, \eta_t) = \max_{c_t^h, a_{t+1}^h} \left\{ u\left(c_t^h\right) + \varphi_t^h \psi_{t,t+1} E_t \left[V_{t+1}^h \left(x_{t+1}, \eta_{t+1} \right) \right] \right\} \qquad (12)$$
$$s.t. \text{ eq. } (5),$$

where the discount function is given by $\varphi_h^h = \gamma\beta$ for $t = h$ and $\varphi_h^h = \beta$ for $t > h$.

Naive Agents

Naive agents do not correctly predict their future selves behavior. Let's first look at the CEU agents. Self h knows that $\nu_{h,h}^h = 1$ and thus uses $\beta\nu_{h,t+1}^h$ to effectively discount the value function in $t+1$. For the next period $t+2$, self h assumes a discount function of $\beta\nu_{h,t+2}^h/\nu_{h,t+1}^h$ and does not anticipate that self $h+1$ will actually use $\beta\nu_{h+1,t+1}^{h+1}$. According to Definition 1 CEU households are *moderately optimistic* if $\nu_{h,t+2}^h/\nu_{h,t+1}^h > \nu_{h+1,t+1}^{h+1}$ which implies that the naive CEU agent assumes a higher effective discount factor for future selves than they actually apply.

Similarly, naive HD agents use $\gamma\beta$ as the discount factor from period $h = t$ to $t+1$ and expect future selve $h+1$ to use the discount factor β for $h+1$ to $h+2$. In fact, self $h+1$ is again present biased and uses $\gamma\beta$.

The naive agent's first order conditions are given by the standard Euler equations.

Observation 4 *The consumption plan* $\mathbf{c} = (c_h, c_{h+1}, ...)$ *of naive agents must satisfy, for all* $t \geq h$

- *CEU agents*

$$(CEU) \qquad \frac{du}{dc_t} \geq \beta(1+r) \cdot E_t \left[\frac{du}{dc_{t+1}} \right] \cdot \begin{cases} \nu_{h,t+1}^h, & for\ t = h \\ \frac{\nu_{h,t+1}^h}{\nu_{h,t}^h}, & for\ t > h \end{cases} \qquad (13)$$

- *HD agents*

$$(HD) \qquad \frac{du}{dc_t} \geq \beta\,(1+r)\,\psi_{t,t+1} \cdot E_t\left[\frac{du}{dc_{t+1}}\right] \cdot \begin{cases} \gamma, & \text{for } t = h \\ 1, & \text{for } t > h \end{cases} \tag{14}$$

Proof 1 *Relegated to Appendix B.1 and B.2.*

Observe the similarities between the two models. In period $t = h$ the CEU agent uses the subjective survival belief which is lower than the objective probability, i.e. $\nu^h_{h,t+1} < \psi_{t,t+1}$. The HD agents have a lower short-term discount factor for periods h to $h+1$, i.e. they discount with $\gamma\beta$. For subsequent periods $t > h$ the CEU agent planning from age h uses the ratio of beliefs which is higher than the conditional survival belief ($\nu^h_{h,t+1}/\nu^h_{h,t} > \nu^h_{t,t+1}$) under moderate optimism, cf. Definition 1. For the the HD agent we have $\gamma = 1$ for $t > h$.

That this plan implies time inconsistent behavior follows from inspection of the marginal rates of substitution (MRS) between any two subsequent periods from the perspective of different planning periods. Under time consistency, these objects would be identical. We have for any planning period h and periods $h < k < t$ that

$$MRS^h_{c_k,c_t} = \left[\left(\varphi^h_t\right)^{t-k}\right]^{-1} \frac{v^h_{h,k}}{v^h_{h,t}} \frac{E_h\left[\frac{du}{dc_k}\right]}{E_h\left[\frac{du}{dc_t}\right]}$$

Comparing the MRS of age h with the MRS of any age $h+i < k$ we find that for the CEU agent the decisive term is the ratio of subjective beliefs which obeys the relationship

$$\frac{\nu^h_{h,k}}{\nu^h_{h,t}} = \frac{\delta_h \cdot \lambda + (1-\delta_h)\cdot\psi_{h,k}}{\delta_h \cdot \lambda + (1-\delta_h)\cdot\psi_{h,t}} \neq \frac{\delta_{h+i} \cdot \lambda + (1-\delta_{h+i})\cdot\psi_{h,k}}{\delta_{h+i} \cdot \lambda + (1-\delta_{h+i})\cdot\psi_{h,t}} = \frac{\nu^{h+i}_{h+i,k}}{\nu^{h+i}_{h+i,t}}.$$

Therefore, $MRS^h_{c_k,c_t} \neq MRS^{h+i}_{c_k,c_t}$.

Similarly, the Hyperbolic agent's discount function is time dependent. For the HD agent the MRS between two subsequent periods $t > h+i$ and $k = t+1$ planned in period h and in $h+i = t$

is given by:

$$MRS^h_{c_{t+1},c_t} = \beta^{-1}\psi_{t,t+1}\frac{E_h\left[\frac{du}{dc_{t+1}}\right]}{E_h\left[\frac{du}{dc_t}\right]} \neq (\gamma\beta)^{-1}\,\psi_{t,t+1}\frac{E_{h+i}\left[\frac{du}{dc_{t+1}}\right]}{E_{h+i}\left[\frac{du}{dc_t}\right]} = MRS^{h+i}_{c_{t+1},c_t}.$$

Sophisticated Agents

As an alternative concept, we define a sophisticated agent as an agent who is fully aware of her dynamic inconsistency. Sophisticated agents have a desire for commitment devices that may force upon their future selves an ex ante optimal plan of actions. In absence of viable commitment mechanisms, these agents' optimization problems are modeled in terms of "subgame perfect" behavior whereby the agent correctly anticipates utility maximizing behavior of her future selves.

The maximization problems for self h of the CEU and the HD agents are the same than for the naive agents and are given by (11) and (12) for all $t = h, \ldots$ where V_t^h is the value function of age $t \geq h$ viewed from planning age h. Unlike the naive agents, sophisticated agents anticipate the correct value function for all future selves as additional constraints, i.e. they anticipate that their future selves will not be acting in their interest. Thus, self h evaluates present and future consumption allocations in periods $t = h, \ldots$ using her value function V_t^h but understands that future consumption allocations will be chosen by different selves $h+1, \ldots$. In absence of commitment devices, the only way to influence future selves behavior is via the savings decision of current self h.

Sophisticated CEU A way to influence future selves behavior via the choice of x_{t+1} is reflected in the conditions of optimality of self h which is for the CEU agent given by

$$\begin{aligned} \frac{du}{dc_h} &= \beta\left(1+r\right)\nu_{h,h+1}^h E_h\left[\frac{\partial V_{h+1}^h}{\partial x_{h+1}}\right] \\ \frac{\partial V_{h+1}^h}{\partial x_{h+1}} &= \frac{du}{dc_{h+1}}\cdot\frac{\partial c_{h+1}}{\partial x_{h+1}} + \beta\left(1+r\right)\frac{\nu_{h,h+2}^h}{\nu_{h,h+1}^h}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)E_{h+1}\left[\frac{\partial V_{h+2}^h}{\partial x_{h+2}}\right] \end{aligned} \tag{15}$$

Envelope conditions which are standard in rational expectations problems no longer apply to the effect that the partial derivatives of the consumption policy $\frac{\partial c_{h+1}}{\partial x_{h+1}} \neq 0$, in general. This reflects how self h can influence future self's $h+1$ choices with her choice of savings, x_{h+1}. Combining the above expressions with the condition of optimality for self $h+1$

$$\frac{du}{dc_{h+1}} = \beta\left(1+r\right)\nu_{h+1,h+2}^{h+1}E_{h+1}\left[\frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}^{h+1}}\right]$$

results in a "generalized Euler-equation with adjustment factor".

Proposition 1 *The generalized Euler-equation with adjustment factor for the sophisticated CEU agent is given by*

$$\frac{du}{dc_h} = \left(1+r\right)\beta\nu_{h,h+1}^h\cdot E_h\left[\left(\frac{\partial c_{h+1}}{\partial x_{h+1}} + \frac{\nu_{h,h+2}^h}{\nu_{h,h+1}^h\cdot\nu_{h+1,h+2}^{h+1}}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\right)\frac{du}{dc_{h+1}}\right] + \Lambda_t \tag{16}$$

where

$$\Lambda_t = \left[\beta\left(1+r\right)\right]^2 \nu^h_{h,h+2} E_h \left(\left(1 - \frac{\partial c_{h+1}}{\partial x_{h+1}}\right) E_{h+1} \left[\frac{\partial V^h_{h+2}}{\partial x_{h+2}} - \frac{\partial V^{h+1}_{h+2}}{\partial x_{h+2}} \right] \right) \tag{17}$$

and where $\nu^h_{h,t}$ is given by equation (1).

Proof 2 *The proof is relegated to Appendix B.3.*

Sophisticated HD The first order conditions and the derivative of the value function for self h is the same than for the naive agent and given by (14). As in the CEU case, cf. equation (15), the envelope theorem does not hold when taking the derivative of the value function for the next period, i.e. the consumption policy function in $h+1$ is not optimal from perspective of self h. The sophisticated HD agent knows that futures selves will also have a strong present bias and face the same problem (while the naive agent thinks that future selves will behave rational). Thus self h knows that self $h+1$ will have the same FOC for the subsequent periods.

Note that the sophisticated HD agent discounts any future period after the successive period with β, i.e. all selves discount future periods identical, leading to identical derivatives of the value functions for future periods, which is why the HD agent does not have an adjustment factor as in the CEU case.

Observation 5 *The generalized Euler-equation of the sophisticated HD agent derived, e.g. by Harris and Laibson (2001) is given by*

$$\frac{du}{dc_h} = \beta\gamma\left(1+r\right)\psi_{t,t+1} \cdot E_h \left[\left(\frac{\partial c_{h+1}}{\partial x_{h+1}} + \frac{1}{\gamma}\left(1 - \frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\right) \frac{du}{dc_{h+1}} \right] \tag{18}$$

Compared to the generalized Euler-equation of the CEU agent (16) the agent uses the objective probability $\psi_{t,t+1}$ to discount future consumption. In addition, the adjustment factor Λ_t is not present due to the fact that the HD agent's problem for future selves $h+1, \ldots$ does not change.

The Euler-equations of the CEU and the HD agents both have a stchastic and endoge-

nous adjustment to the effective discount factor" given by

$$(CEU) \qquad \left(\frac{\partial c_{h+1}}{\partial x_{h+1}} + \frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1} \cdot \nu^{h+1}_{h+1,h+2}} \left(1 - \frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\right), \tag{19}$$

$$(HD) \qquad \left(\frac{\partial c_{h+1}}{\partial x_{h+1}} + \frac{1}{\gamma}\left(1 - \frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\right). \tag{20}$$

which is a linear combination of the marginal propensity to consume and the marginal propensity to save out of cash-on-hand, $\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)$, where the latter is attached a certain weight. Since $\partial c_{h+1}/\partial x_{h+1} \leq 1$ this adjustment is larger one for the HD agent since $\gamma < 1$ and larger one for the CEU agent under *moderate optimism*, cf. Definition 1.

Intuitively, self h is aware of the fact that it values (marginal) savings in $h+1$ higher than her future self. Thus, the sophisticated agent at age h attaches more weight on the marginal utility tomorrow, if future self $h+1$ has a higher marginal propensity to save. The adjustment to the effective discount factor" hence just reflects how self h influences self $h+1$ by scaling up self $h+1$'s valuation of savings. An important difference between the two models is that the weight on the marginal propensity to save is $1/\gamma$ and thus constant for the HD model while it is age dependent for the CEU model. We will show in the quantitative Section 7 that the ratio of subjective beliefs $\nu^h_{h,h+2}/\nu^h_{h,h+1} \cdot \nu^{h+1}_{h+1,h+2}$ is increasing with age as agents are becoming increasingly optimistic. Thus, the adjustment factor rises with age implying higher savings for the CEU agent compared to the HD agent.

The Euler-equation of the sophisticated CEU agent contains an adjustment factor $\Lambda_t \neq 0$, cf. equation (17), which is not present in the HD model. Whereas in the latter, effective discount functions are identical for selves h and $h+1$ from period $h+2$ onwards, we have that the value functions of selves h and $h+1$ in periods $h+2$ are age-dependent and given by

$$V^h_{h+2} = u\left(c_{h+2}\right) + \beta \frac{\nu^h_{h,h+3}}{\nu^h_{h,h+2}} E_{h+2}\left[V^h_{h+3}\left(x_{h+3}, \eta_{h+3}\right)\right]$$
$$V^{h+1}_{h+2} = u\left(c_{h+2}\right) + \beta \frac{\nu^{h+1}_{h+1,h+3}}{\nu^{h+1}_{h+1,h+2}} E_{h+2}\left[V^{h+1}_{h+3}\left(x_{h+3}, \eta_{h+3}\right)\right],$$

where in general $\frac{\nu^h_{h,h+3}}{\nu^h_{h,h+2}} \neq \frac{\nu^{h+1}_{h+1,h+3}}{\nu^{h+1}_{h+1,h+2}}$. A positive difference $\frac{\partial V^h_{h+2}}{\partial x_{h+2}} - \frac{\partial V^{h+1}_{h+2}}{\partial x_{h+2}}$ implies that self h values savings from $h+1$ to $h+2$ higher than self $h+1$. The sign of Λ_t is again a quantitative question studied in Section 7.

4.6 Aggregation over Households

Recall that, when aggregating the economy, we assume constant pension income, i.e., we consider a stationary economy in absence of aggregate risk, hence $b_t = b$ for all t. Wealth dispersion within each age bin is only driven by productivity shocks. We denote the cross-sectional measure of agents with characteristics (a_t, η_t) by $\Phi_t(a_t, \eta_t)$.

Denote by $\mathcal{A} = [0, \infty]$ the set of all possible asset holdings and let $\mathcal{E}$ be the set of all possible income realizations. Define by $\mathcal{P}(\mathcal{E})$ the power set of $\mathcal{E}$ and by $\mathcal{B}(\mathcal{A})$ the Borel σ-algebra of $\mathcal{A}$. Let $\mathcal{Y}$ be the Cartesian product $\mathcal{Y} = \mathcal{A} \times \mathcal{E}$ and $\mathcal{M} = (\mathcal{B}(\mathcal{A}))$. The

measures $\Phi_t(\cdot)$ are elements of $\mathcal{M}$. We denote the Markov transition function—telling us how people with characteristics (t, a_t, η_t) move to period $t+1$ with characteristics $t+1, a_{t+1}, \eta_{t+1}$—by $Q_t(a_t, \eta_t)$. The cross-sectional measure evolves according to

$$\Phi_{t+1}(\mathcal{A} \times \mathcal{E}) = \int Q_t((a_t, \eta_t), \mathcal{A} \times \mathcal{E}) \cdot \Phi_t(da_t \times d\eta_t)$$

and for newborns

$$\Phi_1(\mathcal{A} \times \mathcal{E}) = N_1 \cdot \begin{cases} \Pi(\mathcal{E}) & \text{if } 0 \in \mathcal{A} \\ 0 & \text{else.} \end{cases}$$

The Markov transition function $Q_t(\cdot)$ is given by

$$Q_t((a_t, \eta_t), \mathcal{A} \times \mathcal{E}) = \begin{cases} \sum_{\eta_{t+1} \in \mathcal{E}} \pi(\eta_{t+1}|\eta_t) \cdot \psi_{t,t+1} & \text{if } a_{t+1}(a_t, \eta_t) \in \mathcal{A} \\ 0 & \text{else.} \end{cases}$$

for all $(a_t, \eta_t) \in Y$ and all $(\mathcal{A} \times \mathcal{E}) \in \mathcal{Y}$. Observe that the transition from t to $t+1$ is governed by the objective survival probabilities $\psi_{t,t+1}$.

Aggregation gives average (or aggregate)

consumption:	$\bar{c}_t = \int c_t(a_t, \eta_t)\Phi_t(da_t \times d\eta_t)$,
assets:	$\bar{a}_t = \int a_t(a_t, \eta_t)\Phi_t(da_t \times d\eta_t)$,
income:	$\bar{y}_t = (1-\tau)w\left(\sum_{t=0}^{t_r-1} \phi_t N_t + \chi \sum_{t=t_r}^{T} N_t\right)$,
total income:	$\bar{y}_t^{tot} = \bar{y}_t + r\bar{a}_t$,
savings:	$\bar{s}_t = \bar{y}_t^{tot} - \bar{c}_t$.

5 Simple Model

We now develop a simple three-period model ($T = 2$) without productivity risk ($\eta_t = 1$ for all t) in order to provide insights for the quantitative analysis below. We compare the results of the CEU and the HD agent as well as naive versus sophisticated agents. In addition, the rational expectations (RE) agent will serve as a benchmark case for all other agent types.

The CEU agent discounts exponentially thus lifetime utility can be rewritten as

$$U_{CEU}^0 = u(c_0) + \beta\nu_{0,1}^0\left(u(c_1) + \beta\frac{\nu_{0,2}^0}{\nu_{0,1}^0}u(c_2)\right) \tag{21}$$

In light of the data on subjective beliefs displayed in Figure 1 we interpret period 0 of the simple model as the period when survival beliefs express relative pessimism with respect to survival, i.e., up to actual age of about $65-70$. Period 1 then reflects the period when there is relative optimism in the data. Correspondingly, we make the following assumption:

Assumption 1 *We assume for some $\delta > 0$ that*

$$\psi_{0,1} > \nu^0_{0,1} = \delta_0\lambda + (1-\delta_0)\psi_{0,1} \tag{22}$$

i.e., that $\lambda < \psi_{0,1}$ (pessimistic beliefs), as well as

$$\psi_{1,2} < \nu^1_{1,2} = \delta_1\lambda + (1-\delta_1)\psi_{1,2} \tag{23}$$

i.e., that $\lambda > \psi_{1,2}$ (optimistic beliefs).[18]

As will be shown further below for the CEU agent we need certain conditions in order to generate the empirical facts of saving behavior we want to address. For this we make the following definitions:

Definition 2 (Moderate optimism) *A period 1 naive CEU household is* moderately optimistic *if* $\nu^1_{1,2} < \frac{\nu^0_{0,2}}{\nu^0_{0,1}}$.

Definition 3 (Sufficient pessimism) *A period 0 naive CEU household is* sufficiently pessimistic *if* $\frac{\nu^0_{0,1}}{\psi_{0,1}} < \left(\frac{m^{1,n}_1}{m_1}\right)^\theta < 1$.

Meanwhile, the HD agent has an additional short run discount factor γ, cf. equation (9) and uses objective survival rates ψ. Thus lifetime utility can be written as

$$U^0_{HD} = u(c_0) + \gamma\beta\psi_{0,1}\left(u(c_1) + \beta\psi_{1,2}u(c_2)\right) \tag{24}$$

We assume a $CRRA$ per-period utility function with $\theta \neq 1$ given by

$$u\left(c_t\right) = \Gamma + \frac{c_t^{1-\theta}}{1-\theta}, \tag{25}$$

for all t with preference shifter $\Gamma \geq 0$ such that condition (38) holds, cf. Appendix A.3.

5.1 Rational Expectations

The reference model is the standard solution to the rational expectations model (where $\delta_0 = \delta_1 = 0$):

Observation 6 *Policy functions of the rational expectations solution to the simple model are linear in total wealth, $w_t \equiv x_t + h_t$ (where $x_t \equiv a_t R + y_t$ is cash on hand and $h_t \equiv$*

[18] Notice that, despite equation (23), we may have that the household in period 0 is pessimistic with respect to survival from period 1 to 2, hence we may have that

$$\psi_{1,2} > \nu^0_{1,2} = \delta_0\lambda + (1-\delta_0)\psi_{1,2}.$$

This is so because $\delta_0 < \delta_1$ and therefore less weight is put on the optimism parameter λ.

$\sum_{s=t+1}^{T}\left(\frac{1}{1+r}\right)^{s-t} y_s$ *is human wealth):* $c_t = m_t w_t$ *where* $w_{t+1} = (w_t - c_t)R$ *and*

$$m_t = \begin{cases} \frac{b\psi_t^{-\frac{1}{\theta}} m_{t+1}}{1+b\psi_t^{-\frac{1}{\theta}} m_{t+1}} = \frac{1}{1+\frac{1}{b\psi_t^{-\frac{1}{\theta}} m_{t+1}}} & \text{for } t < T \\ 1 & \text{for } t = T \end{cases} \tag{26}$$

for $b \equiv \left(\beta R^{1-\theta}\right)^{-\frac{1}{\theta}}$.

Proof 3 *See, e.g., Deaton (1992).*

5.2 Naive Agent

To draw a distinction between the different agent types, we use superscript nh to denote policy functions (in terms of marginal propensities to consume) of $n$aive $h$yperbolic households while we simply use n for $n$aive CEU households. Given that the household consumes all outstanding wealth in the final period 2 (i.e. $m_2^n = 1$) in the following we describe the solution for all other periods.

Naive HD agent

Lifetime utility in period 0 for the naive HD agent is given by equation (24).The next self in period 1 solves

$$U_1^{1,nh} = u(c_1) + \beta\gamma\psi_1 u(c_2)$$

According to the assumption of naiveté for hyperbolic discounter, self 0 expects self 1 to be acting rational thus

$$U_1^{0,nh} = u(c_1) + \beta\psi_1 u(c_2)$$

Proposition 2 *The solution for the naive HD household is as follows:*

- *The solution to the problem in period 1 is:*

$$c_1^{1,nh} = m_1^{1,nh} w_1 \quad \text{where} \quad m_1^{1,nh} = \frac{1}{1+\frac{1}{\psi_1^{-\frac{1}{\theta}} b^n}}$$

  where

$$b^n \equiv \gamma^{-\frac{1}{\theta}}\left(\beta R^{1-\theta}\right)^{-\frac{1}{\theta}} \tag{27}$$

- *The plan in period 0 for period 1 is:*

$$c_1^{0,nh} = m_1^{0,nh} w_1 \quad \text{where} \quad m_1^{0,nh} = m_1,$$

  where m_1 *is given in (26).*

- *The solution in period 0 is:*

$$c_0^{0,nh} = m_0^{0,nh} w_0 \quad \textit{where} \quad m_0^{0,nh} = \frac{1}{1 + \frac{1}{b^n \psi_0^{-\frac{1}{\theta}} m_1}}$$

where m_1 and b^n are given by equations (26) and (27), respectively.

Proof 4 *See Appendix B.5.*

Naive CEU agent

For the naive HD agent we have for lifetime utility in period 0 is given by equation (21). Again, self 1 in period 1 solves

$$U_1 = u(c_1) + \beta \nu_{1,2}^1 u(c_2)$$

whereas self 0 expects self 1 to act according to

$$U_1^{0,n} = u(c_1) + \beta \frac{\nu_{0,2}^0}{\nu_{0,1}^0} u(c_2)$$

The naive CEU agent uses the ratio of beliefs $\nu_{0,2}^0/\nu_{0,1}^0$ for future planning ages and does not anticipate that self 1 will use $\nu_{1,2}^1 < \nu_{0,2}^0/\nu_{0,1}^0$.

The solution of the household's problem for the naive CEU agent are as follows:

Proposition 3 *The solution for the naive CEU household is as follows:*

- *The solution to the problem in period 1 is:*

$$c_1^{1,n} = m_1^{1,n} w_1 \quad \textit{where} \quad m_1^{1,n} = \frac{1}{1 + \frac{1}{b\left(\nu_{1,2}^1\right)^{-\frac{1}{\theta}}}}$$

- *The plan in period 0 for period 1 is:*

$$c_1^{0,n} = m_1^{0,n} w_1 \quad \textit{where} \quad m_1^{0,n} = \frac{1}{1 + \frac{1}{b\left(\frac{\nu_{0,2}^0}{\nu_{0,1}^0}\right)^{-\frac{1}{\theta}}}}$$

- *The solution in period 0 is:*

$$c_0^{0,n} = m_0^{0,n} w_0 \quad \textit{where} \quad m_0^{0,n} = \frac{1}{1 + \frac{1}{b(\nu_{0,1}^0)^{-\frac{1}{\theta}} m_1^{0,n}}}$$

Proof 5 *See Appendix B.6.*

Interpreting the propositions yields the following observation:

Observation 7 *Realization in period* 1*: Comparing the CEU with the RE agent we get* $m_1^{1,n} < m_1$ *under Assumption 1, equation (23). Comparing the HD with the RE agent we get* $m_1^{1,nh} > m_1$ *since* $b^n > b$ *as* $\gamma \in [0,1]$.

According to Observation 7 the naive CEU household, having optimistic survival beliefs, saves more out of accumulated wealth in period 1 than the household with rational expectations. On the contrary, the naive HD agent saves relatively *less* in period 1. The reason is that the HD agent in period 1 uses the discount factor $\gamma\beta$ when making her decision between consumption in period 1 and period 2.

Accumulated wealth in turn is an endogenous object. We shall see below that in period 0 both a *sufficiently pessimistic* naive CEU household as well as a naive HD household will save less out of initial wealth than households with rational expectations. While it is therefore clear that accumulated wealth of the naive CEU household in period 1 is lower than for rational expectations, relative wealth positions across the CEU household and the RE agent in period 2 depend on the relative strength of *sufficient pessimism* in period 0 vis-a-vis *optimism* in period 1. For the CEU agent, it is therefore ultimately a quantitative question whether accumulated wealth in period 2 exceeds wealth of households with rational expectations. On the contrary, wealth of the HD household in period 2 can never exceed wealth of the RE agent, as there is no analogue to the *optimism* present for CEU households.

Observation 8 *Plan for period* 1*: For the plan of the naive CEU agent we get* $m_1 > m_1^{1,n} > m_1^{0,n}$ *under* moderate optimism*, cf. Definition 2. For the HD agent we get* $m_1^{1,nh} > m_1^{0,nh} = m_1$*, cf. Opservation 7 and Propostition 2.*

Observation 8 implies that a naive CEU household with *moderate optimism* plans in period 0 to save more out of accumulated wealth in period 1 than he actually does, i.e. $m_1^{1,n} > m_1^{0,n}$. That is, only if optimism expressed in the data on subjective beliefs is not too large there is hope for our quantitative analysis to match the stylized fact that households, in the course of the life-cycle, save less than originally planned. Moreover, the CEU household plans in period 0 to save more out of wealth in period 1 than the rational expectations household. It implies that, to generate undersaving in our model, pessimism in the first period must be sufficiently large.

The HD agent unambiguously plans in period 0 to save more in period 1 than he actually does, where the HD plan for period 1 corresponds to RE case, i.e. $m_1^{0,n} = m_1$.

Observation 9 *Realization for period 0: For the naive CEU agent with* sufficient pessimism *we get that* $m_0^{0,n} > m_0$*, and for the naive HD agent we have* $m_0^{0,nh} > m_0$ *since again* $b^n > b$ *as* $\gamma \in [0,1]$.

Proof 6 *See Appendix B.7.*

Both the CEU and the HD household consume more than the RE household in period 0. But for the naive CEU agent the household needs to be sufficiently pessimistic in the sense of Definition 3 in order to generate undersaving (relative to the solution to the rational expectations model).

Table 1 summarizes the results for the naive agents compared to the RE agent.

Table 1: Marginal propensities to consume out of current wealth

	naive CEU	naive HD
Realization in $t=0$	$m_0^{0,n} > m_0$ [1)]	$m_0^{0,nh} > m_0$
Plan in $t=1$	$m_1^{1,n} > m_1^{0,n}$ [2)]	$m_1^{1,nh} > m_1^{0,nh}$
Realization in $t=1$	$m_1^{1,n} < m_1$	$m_1^{1,nh} > m_1$

m_0 and m_1 denotes the MPC of the RE agent

[1)] Under sufficient pessimism

[2)] Under moderate optimism

In period $t=0$ both agents exhibit undersaving compared to the RE agent. Similarly, both agents plan to save more in period 1 than they actually do. Note, that for the CEU agent certain conditions have to be met in order to generate these results. Most importantly, the CEU agent having optimistic survival beliefs saves more relative to the RE agent in period 1. In contrast, the naive HD agent exhibit undersaving in period 1 relative to the RE agent once more. A tendency for oversaving at older ages cannot be accounted for.

5.3 Sophisticated Agents

Sophisticated HD Agent

As described in the multiperiod model, cf. Section 4.5, the sophisticated agent is aware of future selves behavior and anticipates their present bias. Thus, the sophisticated agents takes the (over-)consumption of future selves into account when making her current savings plan.

In periods 2 and 1, the solution is identical to the respective solution for the naive agent. Equivalence in period 2 is trivial because $m_2 = 1$, for period 1 we have in the Euler equation that $m_2 = 1$ enters so that the Euler equation simplifies to

$$c_1^{-\theta} = \beta\gamma R\psi_1 c_2^{-\theta}$$

which is equivalent to the first-order condition of the naive agent.

For period 0 we get the following Euler-equation

$$c_0^{-\theta} = \beta\gamma R\psi_0 \left(m_1^{1,nh} + \frac{1}{\gamma}(1 - m_1^{1,nh}) \right) c_1^{-\theta}$$

where

$$d\left(m_1^{1,nh}\right) \equiv \left(m_1^{1,nh} + \frac{1}{\gamma}(1 - m_1^{1,nh}) \right) > 1 \tag{28}$$

can be interpreted as an adjustment term to the discount factor. This factor is larger one as long as $m_1^{1,nh} < 1$. The fact that $m_1^{1,nh}$ enters the Euler equation makes explicit that the sophisticated agent understands how next period's self will behave.

Proposition 4 *The solution to the sophisticated HD agent's problem in period 0 is given by*

$$c_0^{0,s} = m_0^s w_0 \quad \textit{where} \quad m_0^s = \frac{1}{1 + \frac{1}{b\left[d\left(m_1^{1,nh}\right)\cdot\gamma\psi_0\right]^{-\frac{1}{\theta}} m_1^{1,nh}}}$$

where $d\left(m_1^{1,n}\right)$ *is given by equation (28).*

Proof 7 *Relegated to Appendix B.8.*

Sophisticated CEU Household

The solution to the sophisticated agent's problem is identical to the naive agent in periods 2 and 1. In period 0 the sophisticated agent knows that his future period 1 self will have marginal propensity $m_1^{1,n}$. Again we use superscript sh to denote the *s*ophisticated *h*yperbolic households and s for the *s*ophisticated CEU household. The solution to the problem of the sophisticated CEU agent is as follows:

Proposition 5 *The solution to the sophisticated CEU agent's problem in period 0 is given by*

$$c_0^{0,s} = c_0 = m_0^{0,s} w_0 \quad \textit{where} \quad m_0^{0,s} = \frac{1}{1 + \frac{1}{b\left[d\left(m_1^{1,n}\right)\cdot\nu_{0,1}^0\right]^{-\frac{1}{\theta}} m_1^{1,n}}}$$

and

$$d\left(m_1^{1,n}\right) \equiv \left(m_1^{1,n} + \frac{\nu_{0,2}^0}{\nu_{0,1}^0\nu_{1,2}^1}(1 - m_1^{1,n}) \right) > 1$$

Proof 8 *Relegated to Appendix B.9.*

Note, that $d\left(m_1^{1,n}\right) > 1$ only under moderate optimism, cf. Definition 2, for the CEU agent. Observe that d is the counterpart to the adjustment of the discount factor Ψ_t in the multiperiod model in Proposition 16.

The interpretation for the sophisticated agents are similar for the HD and the CEU case. Compared to the naive agent two forces are at work.[19] On the one hand, $d > 1$ which reflects the sophisticated agent's high valuation of savings. On the other hand, self 0 anticipates the high marginal propensity to consume of self 1. Since d depends negatively on $m_1^{1,n}$, implying that when self 0 decides how much to save, she takes into account that future self 1 will consume more, than self one would like her to.

The propositions above lead to the following observation:

Observation 10 *The relationship between the marginal propensities to consume between the sophisticated and the naive agent is as follows:*

1. HD agent:

$$m_0^{0,sh} < m_0^{0,nh} \qquad \textit{iff} \qquad d\left(m_1^{1,nh}\right) > \left(\frac{m_1^{1,nh}}{m_1}\right)^{\theta} > 1$$

$$m_0^{0,sh} > m_0^{0,nh} \qquad \textit{iff} \qquad \left(\frac{m_1^{1,nh}}{m_1}\right)^{\theta} > d\left(m_1^{1,nh}\right) > 1.$$

2. CEU agent

$$m_0^{0,s} < m_0^{0,n} \qquad \textit{iff} \qquad d\left(m_1^{1,n}\right) > \left(\frac{m_1^{1,n}}{m_1^{0,n}}\right)^{\theta} > 1$$

$$m_0^{0,s} > m_0^{0,n} \qquad \textit{iff} \qquad \left(\frac{m_1^{1,n}}{m_1^{0,n}}\right)^{\theta} > d\left(m_1^{1,n}\right) > 1.$$

Proof 9 *Relegated to Appendix B.10*

Which out of the two effects dominates is again a quantitative question.

[**TBC: Compare sophisticated agent with Rational expectations agent:**

$$\begin{aligned} m_0^{0,nh} &> m_0 \\ m_0^{0,n} &> m_0 \end{aligned}$$

Also: discuss period 1 decisions (identical to naive case here)]

[19] Again, for the sophisticated CEU agent those two forces are only at work under moderate optimism, cf. Definition 2.

6 Calibration

6.1 Household Age

Households enter the model at age 20 (model age 0). The last working year is age 64, hence $t_r = 45$. We set the horizon to some maximum biological human lifespan at age 125, hence $T = 105$. This choice is motivated by Weon and Je (2009) who estimate a maximum human lifespan of around 125 years using Swedish female life-table data between $1950 - 2005$.

6.2 Objective Cohort Data

For objective survival rates we estimate cohort specific survival rates for US cohorts alive in 2007. Objective cross-sectional data is taken from the Social Security Administration (SSA) for $1890 - 1933$ and the Human Mortality Database (HMD) for the years $1934 - 2007$. To obtain complete cohort tables, future survival rates are predicted by the Lee and Carter (1992) procedure. Details are described in Ludwig and Zimper (2013).

Since data on survival rates is unreliable for ages past 100 we estimate survival rates assuming the Gompertz-Makeham law.[20] Accordingly, the mortality rate μ_t at age t is assumed to follow

$$\mu_t = \alpha_1 + \alpha_2 \cdot \exp(\alpha_3 \cdot t) + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma^2).$$

We estimate parameters $\{\alpha_i\}_{i=1}^3$ to get an out of sample prediction for ages past 100. The resulting predicted mortality rate function fits actual data very well and is used as objective cohort data in the simulation. According to our estimates, the average mortality rate approaches 1 at ages around age 110 ($t = 90$). For all ages $t = 91, \ldots, 105$, we set the objective survival rate to $\psi_{t,t+1} = \varepsilon = 0.01$.

6.3 Estimated Subjective Survival Beliefs

We follow Ludwig and Zimper (2013) and estimate parameters $\delta \equiv \delta_{h=0}$ and λ by pooling a sample of HRS data formed of HRS waves $\{2000, 2002, 2004\}$. Except for heterogeneity in sex and age, we ignore all other heterogeneity across individuals. This gives $\delta = 0.118$ and $\lambda = 0.406$.[21]

[20] See, e.g., Preston et al. (2001), p. 192.

[21] Estimation results are separately for men and women. We take an equally weighted average of the estimated parameters to get an approximation for λ and δ in the population.

6.4 Calibrated Hyperbolic Short-Term Discount Factor

The effective discount function of the CEU model at age h is $\nu_{h,t}^h \beta^{t-h}$, for $t = h, \ldots, T$. In the HD model it is $\psi_{h,t}\gamma\beta^{t-h}$. We calibrate the short term discount factor γ of the HD model by minimizing the Eucleadian distance of the two discount functions for every plan:

$$\min_{\gamma} \sum_{h=1}^{T} \sum_{t=h}^{T} \left[\nu_{h,t}^h \beta^{t-h} - \psi_{h,t}\gamma\beta^{t-h}\right]^2 \tag{29}$$

The resulting short-term preference factor is $\gamma = 0.77$

[**TBC: check value**].

This is just in the middle of the range cited in the literature. In numerical simulations the short term preference parameter γ is set to values between 0.6 and 0.85, see, e.g. Laibson, Repetto, and Tobacman (1998).

6.5 Preferences

As in the simple model of the previous section, per period utility is assumed to be CRRA, $u(c_t) = \Gamma + \frac{c_t^{1-\theta}-1}{1-\theta}$, at all ages t. As a benchmark, we choose $\theta = 3.0$—corresponding to an inter-temporal elasticity of substitution (IES) of one third—and consider as range for sensitivity analysis $\theta \in \{2, 4\}$.

Given $\theta > 1$, per period utility is negative. For the CEU model it is necessary that the discounted sum of utilities in case the agent survives until T is positive, cf. condition (38) in Appendix A.3. Therefore a preference shifter Γ must be calibrated such that condition (38) holds for all t, η^t. We set $\Gamma = 76.7$ for the naive CEU agent which turns out to be sufficiently high.[22] We further set the discount rate ρ to 5%.

6.6 Prices and Endowments

Wages are normalized to $w = 1$. We consider a three-state first-order Markov chain process for the income process in periods $t = 0, \ldots, t_r$ with state vector $E^w = [1+\epsilon, 1-\epsilon, 0]$ whereby the last entry reflects the state with zero income. Let $\zeta = 0.01$ be the small probability of receiving zero income. Then the transition matrix during the working period writes as

$$\Pi^w = \begin{bmatrix} (1-\zeta)\kappa & (1-\zeta)(1-\kappa) & \zeta \\ (1-\zeta)(1-\kappa) & (1-\zeta)\kappa & \zeta \\ 0.5\cdot(1-\zeta) & 0.5\cdot(1-\zeta) & \zeta \end{bmatrix}$$

[22] This relates to Hall and Jones (2007) who calibrate—in a different model setup—a preference shifter in the range of [22.1; 131.9]. Notice that this is just an arbitrary monotone transformation. Any choice of $\Gamma > 76.7$ ensures that the value of life is always higher than the value of death.

for $t = 0, \ldots, t_r$. We take as initial probability vector of the Markov chain $\pi_0 = [0.5, 0.5, 0]'$, i.e., households do not draw zero income in their first period of life.

Values of persistence and conditional variance of the income shock process are based on the estimates of Storesletten, Telmer, and Yaron (2004)yield $\kappa = 0.97$ and $\epsilon = 0.68$.

Age specific productivity of wages is estimated based on PSID data applying the method developed in Huggett, Ventura, and Yaron (2007).

In retirement, for $t = t_r + 1, \ldots, \bar{T}$, we take as state vector $E^r = [1, 0]$. We assume an even smaller probability to receive zero retirement income of $\zeta^r = 0.001$ which reflects default of the government on its pension obligations. We accordingly have

$$\Pi^r = \begin{bmatrix} 1 - \zeta^r & \zeta^r \\ 1 - \zeta^r & \zeta^r \end{bmatrix}$$

for $t = t_r + 1, \ldots, \bar{T}$ and we take as initial probability vector $\pi_{t_r+1} = [1 - \zeta^r, \zeta^r]'$.

The interest rate is taken to be $r = 0.042$ which is the average real rate of return of stocks and long-run bonds for the US between 1946 and 2001, estimated by Siegel (2002). For the social security contribution rate we take the US average of $\tau = 0.124$. The pension benefit level then follows from the social security budget constraint, cf. equation (6).

All parameters are summarized in table 2.

7 Results

7.1 Subjective Survival Beliefs and Effective Discounting

Predicted and Actual Subjective Survival Beliefs

Figure 3 compares predicted subjective survival rates resulting from the decision theoretic model with their empirical counterparts and corresponding objective survival rates. Predicted subjective beliefs fit data on subjective survival probabilities well. In particular, the model replicates underestimation of survival rates at younger ages and overestimation at older ages.[23]

Survival Belief Functions and Probability Weighting

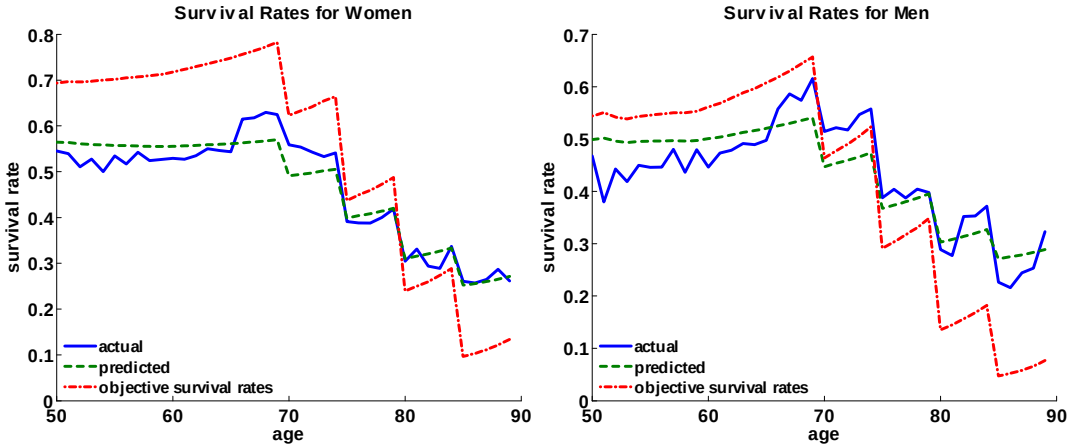
Figure 4 compares subjective survival belief functions of a CEU agent to the subjective beliefs implied by CPT as well as to objective cohort data. The panels in the figure show unconditional survival rates viewed from different planning ages where target age t is

[23] The different line segments are due to changes in target ages. Ludwig and Zimper (2013) perform an extensive sensitivity analysis with regard to focal point answers, the choice of initial age, and the specific form of the experience function. They show that results do not hinge on these aspects.

Table 2: Calibrated parameters

Technology and Prices	
$w = 1$	Gross wage
$r = 0.042$	Interest rate
$\tau = 0.124$	Social security contribution rate
$\chi = 0.322$	Net pension benefit level
Income Process	
$\kappa = 0.97$	Persistence of income
$\epsilon = 0.68$	Variance of income
$\zeta = 0.01$	Probability of receiving zero labor income
$\zeta^r = 0.001$	Probability of receiving zero pension benefits
$\{\phi_t\}$	Age specific productivity estimated from PSID
Preferences	
$\theta \in \{2, 3, 4\}$	Coefficient of relative risk aversion
$\rho = 0.05$	Subjective discount rate (CEU and HD)
$\gamma = 0.77$	Short-term discount factor (HD)
$\Gamma^{CEU} = 76.65$	Preference shifter (naive CEU)
$\Gamma^{HD} = XXX$	Preference Shifter (naive HD)
Subjective Survival Beliefs	
$\delta = 0.118$	Initial degree of ambiguity (CEU)
$\lambda = 0.406$	Degree of optimism (CEU)
Age Limits and Survival Data	
0	Initial model age (age 20)
$t_r = 45$	retirement (age 65)
$T = 105$	Maximum human lifespan (age 125)
$\{\psi_{k,t}\}$	Objective survival rates from SSA and HMD

Figure 3: Objective, subjective and predicted survival rates



Notes: In the HRS interviewees are asked about their survival belief to a specific target age depending on the age at interview: Respondents between ages 50-69 are asked their probability to survive to 80, while agents between 70-74 (and 75-79, 80-84, 85-89) are asked about their belief to survive until 85 (and 90, 95, 100). The figure shows these subjective survival beliefs for different target ages (solid blue line), the corresponding objective survival rates (dashed-dotted red line) and the simulated subjective survival beliefs from the estimated CEU model (dashed green line).

depicted on the abscissa. In each of the panels experience and thus likelihood insensitivity does not change.

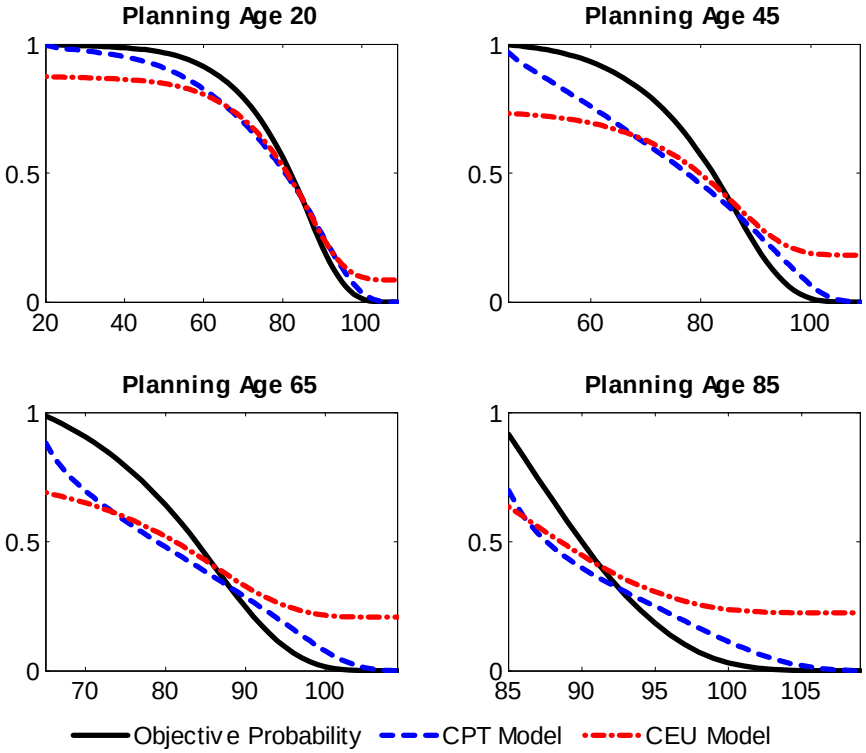
First note that both the CEU and the CPT model imply survival belief functions which are flatter than the objective counterpart, confirming previous findings from Hammermesh (1985) and several others. Thus, both models exhibit an underestimation of high survival rates ans an overestimation of low survival probabilities.

The main difference is present at the endpoints. The subjective survival belief function implied by the CPT model is close to the objective data at the endpoints where the survival rate is close to one or zero, respectively. On the other hand, the survival beliefs of the CEU model are characterized by an immediate downward jump from the first to the second period while at the same time the overestimation for old-age survival rates becomes more pronounced for higher target ages.

While we do not have data on the one-year-ahead survival beliefs of agents in the HRS we can only judge by anecdotal evidence, that the subjective belief to survive the current year is pretty close to the objective data on average implying the CPT model to be more in line with reality than the CEU model.

On the other hand, according to the HRS the overestimation of survival rates at older ages becomes even larger for higher ages. This can be well accounted for by the CEU model where the overestimation becomes more pronounced at older ages For example, our model predicts that a 85 year old agent overestimates her probability to become 100 but she overestimates the probability to become 110 by even more. For the probability

Figure 4: Unconditional probabilities



Notes: Unconditional objective and subjective survival probabilities viewed from different planning ages h.
The parameters ξ_h for the CPT model, cf. equation (), are calibrated by minimizing the Euclidean distance of $\varpi_h\left(\psi_{h,t}, \xi_h\right)$ with a non-linear root finder that best matches the neo-additive capacity $\nu^h_{h,t}$ of that age.

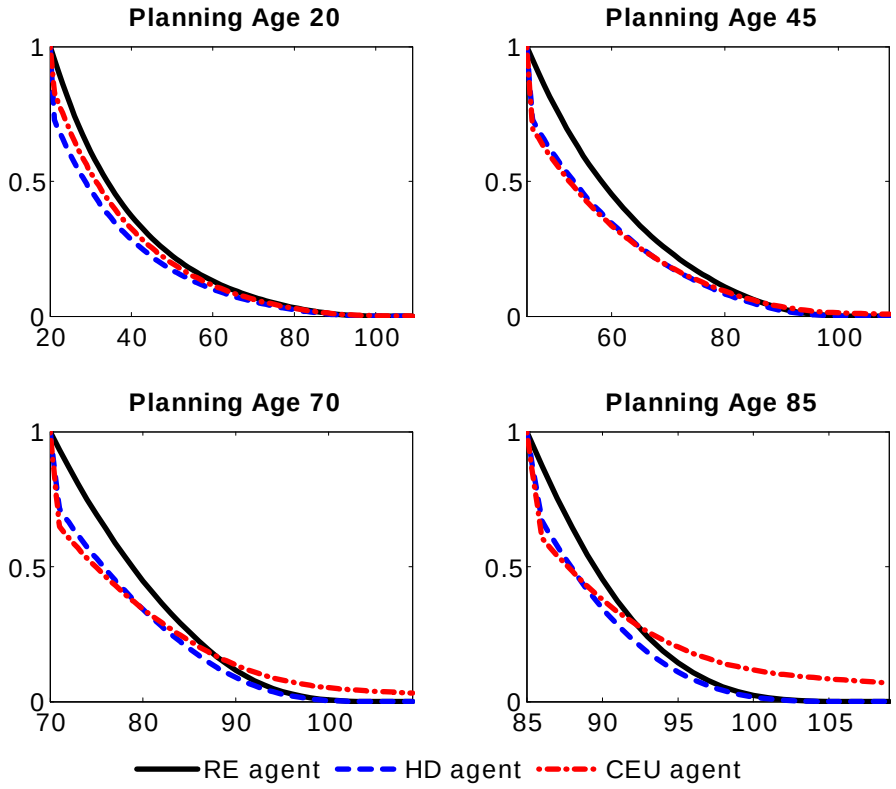
weighting model, the subjective belief of a 85 year old agent to become 100 is overestimated while the belief to become 110 is correctly anticipated and coincides almost with the objective probability. The increasing optimism observed in the data can thus only partially be accounted for by the CPT model.

Subjective CEU Beliefs and Quasi-Hyperbolic Discounting

Both the survival probability and the discount function form the effective discount factor which is depicted in Figure 5. The figure shows effective discount functions of the CEU, the HD and RE agent. Here we are using the parameter calibrated for the numerical version for our model, cf. Table 2. In particular, we assume the long-term discount factor β as a deep structural parameter and thus the same for all agent types throughout the quantitative analysis of our paper.

At a first glance, the effective discount functions of the CEU and the HD agent are quite similar. Since the functions for both models are generally lower than their objective counterpart at younger ages, this also translates into the effective discount function. Note, that at younger ages, the CEU function is closer to the RE agent than the HD function

Figure 5: Effective discount functions



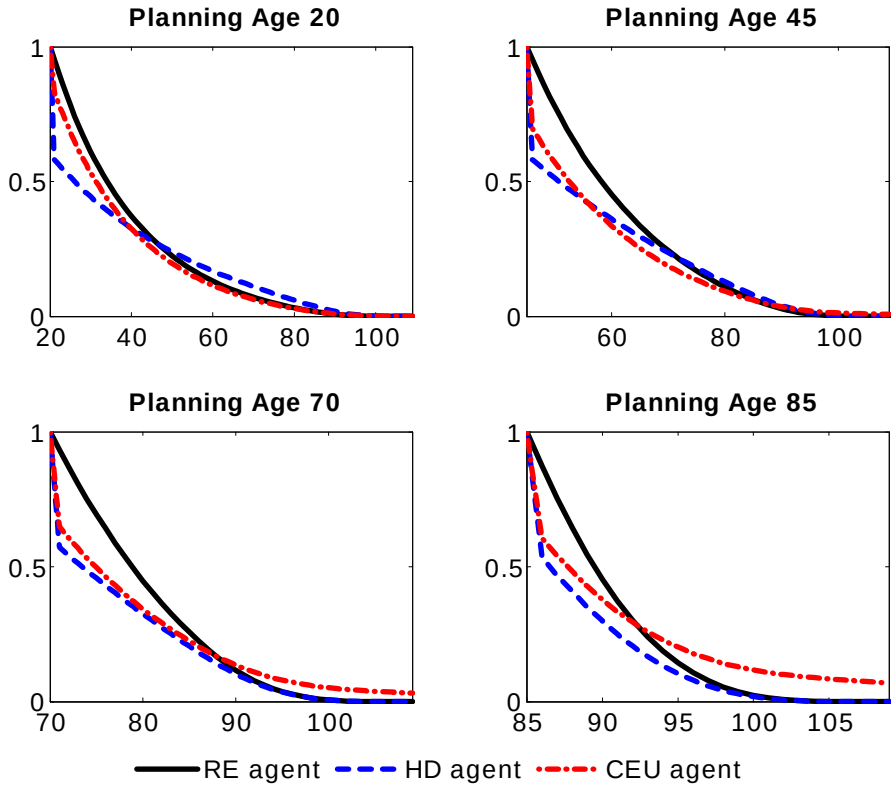
Notes: Effective discount functions of the RE, CEU and HD agent, 1st period is set to one. The parameter γ for the HD discount function is calculated by minimizing the Eucledian distance to the CEU discount function.

while this is reverse at older ages. The main difference between the models occur at later target ages. Here, the CEU function is higher than for the RE agent, and this is becoming more pronounced the older the agent is when making her plan. In contrast, the effective discount function of the HD agent always converges to the discount function of the RE agent from below. This is due to the fact that we use the same long-term discount factor β for all agents. The specific feature of relatively higher discount factors at old ages from our subjective survival belief model accounts for the empirical fact that older agents become even more optimistic concerning their survival prospects the older they get. This generates important life-cycle consequences namely high asset holdings of the elderly.

To see whether this result rests on the assumption of holding the discount factor constant, we plot the discount functions for different β in Figure 6. In particular we exogenously set different long-term discount factors of $\beta^{CEU} = \beta^{RE} = 0.95$ and $\beta^{HD} = 0.97$ and calibrated $\gamma = 0.6$ in order to minimize the Eucledian distance of the effective discount function, cf. equation (29) [**TBC: recalibrate**].

As a new feature compared to the results shown in Figure 5, the HD effective discount function is now also higher than the RE counterpart for high target ages. This

Figure 6: Effective discount functions for different β



Notes: Effective discount functions of the RE, CEU and HD agent, 1st period is set to one. We use the following values: $\gamma = 0.6$ and $\beta^{HD} = 0.97$ for the quasi-hyperbolic, $\gamma = 1$ and $\beta^{CEU} = 0.95$.

holds especially true for younger planning ages. But although the difference of the pure discount factor of the HD agent relative to the RE agent becomes larger for higher target ages, cf. the illustration in the right panel of Figure 2, the *effective* quasi-hyperbolic discount factor—consisting of the product of the pure time discount factor and the survival probability—approaches zero at old ages. The reason for this outcome is that the HD agent uses objective (unconditional) survival rates that reach values below 2% for agents past 100.

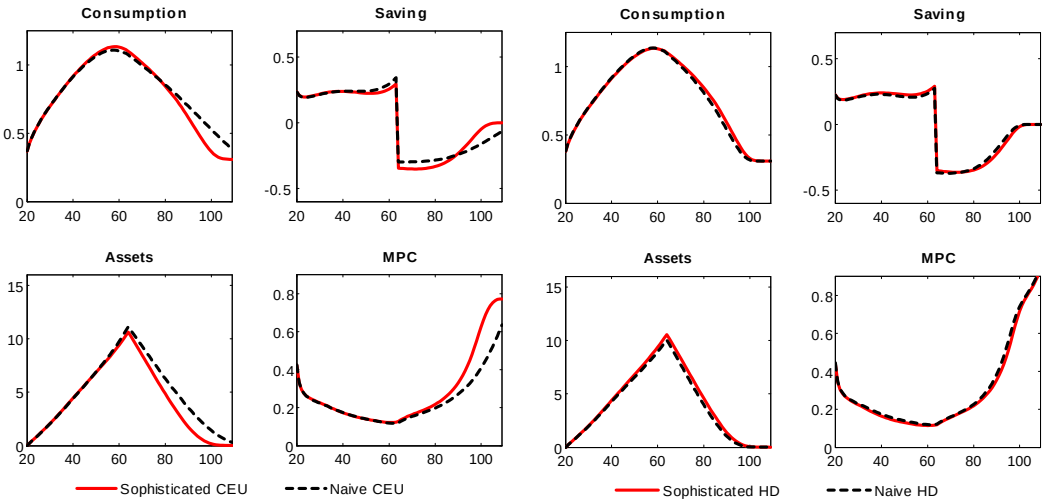
This is in stark contrast to our model where the CEU agent uses subjective beliefs (and exponential discounting) rendering the effective discount function persistently higher at older target ages than for the RE agent. Moreover, this higher effective discount function becomes more pronounced for higher target ages, in line with the data on subjective survival beliefs that shows higher deviations from the objective data the older the agent gets. As mentioned above, it is precisely this feature of our CEU model that generates high asset holdings at older ages compared to the RE agent, which will be shown in the next sections.

7.2 Life-Cycle Profiles

Naive vs. Sophisticated Agents

Figure 7 compares the sophisticated and the naive agent for the CEU and the HD model, respectively. Average life-cycle profiles of sophisticated and naive CEU agents are depicted panel (a) of the figure. On average, the two agents types' behave quite similar at the first half of the working life. At the age of around 40 sophisticated agents start to consume more and build up less assets on average than naives. The overconsumption continues during retirement, where sophisticates run down their assets more strongly than naifs. Finally, naive agents end up with more assets at older ages than sophisticated agents. Intuitively, it seems that sophisticates anticipate that later selves will spend a lot of the retirement savings of the current self. Thus, sophisticates tend to consume more than naifs right away.

Figure 7: Life-cycle profiles of the naive and sophisticated CEU agent



Notes: Profile of sophisticated CEU and HD agents and profile of the naive couterpart using ex-post realization for all planning ages.

In light of the simple model presented in Section 5 there are two opposing effects on the marginal propensity to consume of sophisticated agents. According to Observation 10 it seems that the marginal propensity to consume is higher for the sophisticated agent, if the deviation of the planned and the realized future marginal propensity of the naive agent is large. This is certainly more pronounced the case for the CEU agent. Recall that the agent plans with the ratio of subjective belief $\nu^h_{h,t+1}/\nu^h_{h,t}$ while the next period's self uses $\nu^{h+1}_{t,t+1}$. This ratio of beliefs is close to one while the subjective belief $\nu^{h+1}_{t,t+1}$ is rather low at older ages. Thus the different between the naive plan and the actual realization is large. In the simple model this would imply $\left(\frac{m_1^{1,nh}}{m_1}\right)^\theta > d\left(m_1^{1,nh}\right)$ and thus $m_0^{0,sh} > m_0^{0,nh}$, cf. Observation 10. On the contrary, the naive and sophisticated agent in the HD model—as

already shown in the simple model—become more alike, the older the agent gets. Finally, in period T they are solving the same problem. In line with the predictions of the theory, the profiles for the HD agents, depicted in panel (b) of Figure 7 behave almost identical with given parameters, an outcome confirming earlier studies, cf. Angeletos et al. (2001). [**TBC: Give intuition**]

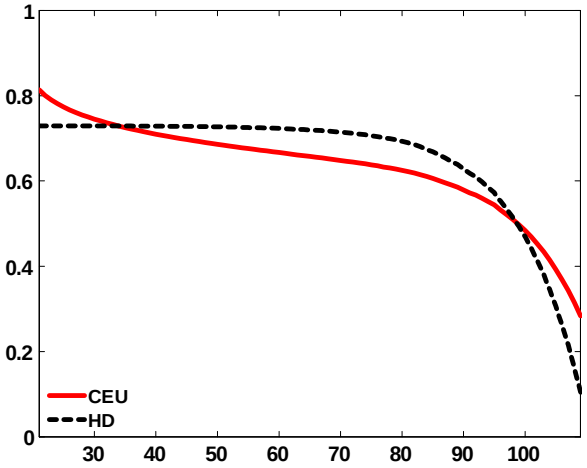
Comparing the sophisticated agents observe the very similar profiles for the CEU and the HD agents. As for the naive agent, the sophisticated CEU agents exhibit higher asset holdings at old ages, see also Figure 14. To see why, observe in the lower right panel of Figure 7 that also the marginal propensity to consume is lower for the CEU agents at older ages. To explain this outcome recall the main differences of the generalized Euler-equation of the CEU and the HD agents, cf. equations (16) and (18) which are restated again for convenience:

$$
\begin{aligned}
(CEU) \qquad \frac{\partial u}{\partial c_h} &= (1+r)\,\beta \nu^h_{h,h+1} \\
&\quad \cdot E_h\left[\left(\frac{\partial c_{h+1}}{\partial x_{h+1}} + \frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1}\cdot \nu^{h+1}_{h+1,h+2}}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\right)\frac{\partial u}{\partial c_{h+1}}\right] + \Lambda_t, \\
(HD) \qquad u_c(c^h_h) &= (1+r)\,\beta\gamma\psi_{t,t+1} \\
&\quad \cdot E_h\left[\left(\frac{\partial c_{h+1}}{\partial x_{h+1}} + \frac{1}{\gamma}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\right)\frac{du}{dc_{h+1}}\right],
\end{aligned}
$$

where Λ_t is given by equation (17).

First, the effective discount factor from period h to $h+1$ are, of course, different, as shown in Figure 8. While for the HD agent the factor ist just $\gamma\beta\psi_{h,h+1}$, i.e. a scaled-

Figure 8: Effective Discount Factors from age h to age h+1



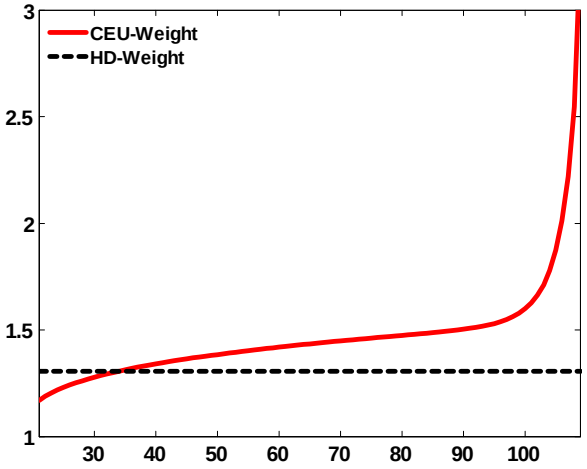
Notes: The figure shows the effective discount factors $\beta\nu^h_{h,h+1}$ for CEU agents and $\gamma\beta\psi_{h,h+1}$ for HD agents.

down objective survival rate function, for the CEU agent the discount factor is $\beta\nu^h_{h,h+1}$, i.e. depending on the subjective survival belief. Observe from the figure that the discount

factor of the CEU agent is higher early in life, then lower most of the lifetime but turns higher again at very old ages.

Second, the effective discount factor is scaled up by an adjustment factor depending on the marginal propensity to consume. Note that for the CEU agent the adjustment is age-dependent through the weight of the marginal propensity to save. The weights to the marginal propensity to save are given by $\nu^h_{h,h+2}/\nu^h_{h,h+1} \cdot \nu^{h+1}_{h+1,h+2}$ for CEU agents and $1/\gamma$ for HD agents. Observe that the latter is constant while the former is age-dependent. Figure 9 plots both weights.

Figure 9: Weights in the Generalized Euler-Equation of Sophisticated Agents



Notes: The figure shows the weights of the marginal propensity to save out of total wealth in the adjustment of the effective discount factors for the CEU agent, cf. equation (19) given by $\nu^h_{h,h+2}/\nu^h_{h,h+1} \cdot \nu^{h+1}_{h+1,h+2} > 1$ and for the HD agent, cf. equation (20), given by $\frac{1}{\gamma} > 1$.

First, note that for the CEU agents the weight is larger one, implying that the agents are *moderately optimistic* in the sense of Definition 1 throughout their lifetime. In addition, the weight increases with age implying that the sophisticated agent becomes more optimistic at older ages putting more weight on next periods marginal utility in the generalized Euler-equation. In contrast, for HD agents the weight is constant throughout.

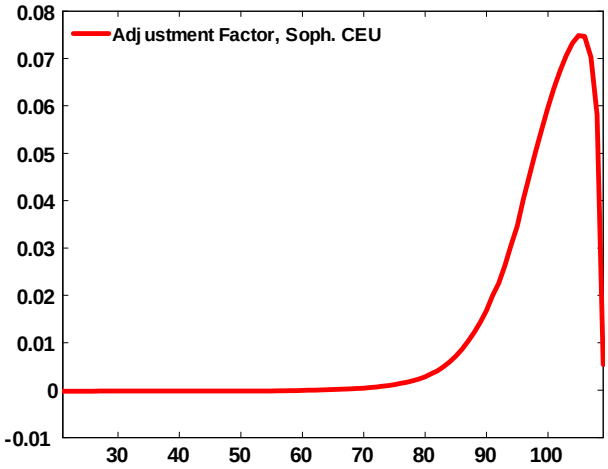
Third, the generalized Euler-equation for the CEU agent contains an adjustment factor Λ_t, cf. equation (17). The average factor is plotted in Figure 10.

Generally, the adjustment factor is slightly negative for young agents and turns positive for older agents. Again, there is a third reason for the sophisticated CEU agent to have higher asset holdings at older ages than the sophisticated HD agent.

Time Inconsistency of Naive Agents

This section compares the plan and the realized action of naive agents who behave time-inconsistent. Naive agents update their plan in each period. As a way to compare any

Figure 10: Adjustment Factor for the Sophisticated CEU Agents



Notes: The figure plots the adjustement factor of the sophisticated CEU agents given by equation (17).

gap between a plan made at age h and future plans for CEU agents we denote planned average consumption with superscripts and compute

$$\tilde{c}_t^h = \int c_t^h(a_t, \eta_t)\Phi_t^h(da_t \times d\eta_t) \tag{30}$$

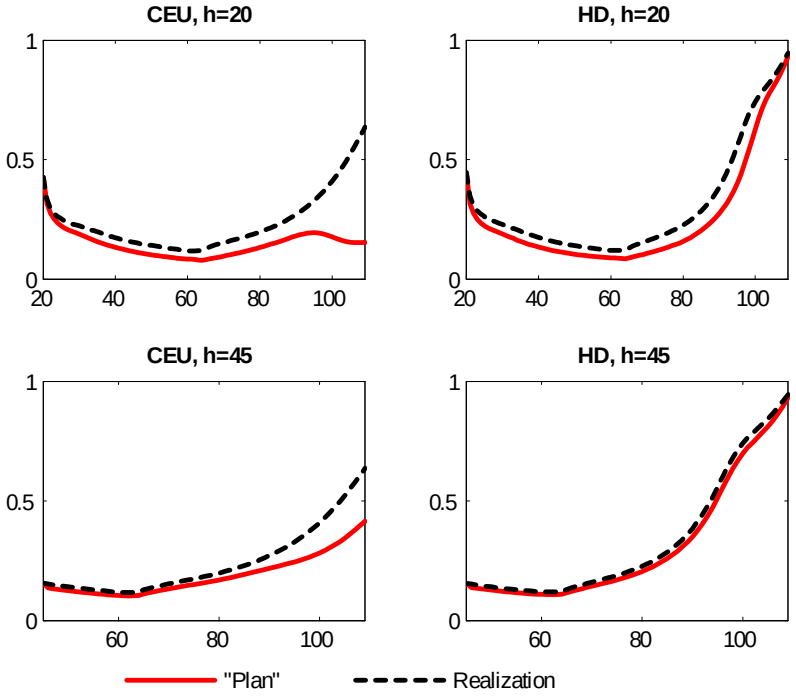
for all t. This gives us hypothetical average consumption profiles in the population if households would stick to their respective period-h plans in all periods $t = h, \ldots, T$. Observe that $\Phi_t^h(\cdot)$ is an artificial distribution generated by respective plans of households. We refer to (30) as (average) "planned" consumption (asset, ...) profile in the figures that follow.

Figure 11 compares the planned and the realized marginal propensity to consume out of cash-on-hand (MPC) of naive agents for the CEU and the HD model at two different planning ages. The notable difference between the two models is that within the CEU model the MPC for the naive agent is moving further away from the sophisticated counterpart the higher the target age. As mentioned above this is due to the fact, that the naive agents plan with the ratio of subjective belief $\nu_{h,t+1}^h/\nu_{h,t}^h$ instead of using simply $\nu_{t,t+1}^h$. On the contrary, the MPC for the HD agent are much more alike between the sophisticated and the naive agents. In addition, they end up being the same in the last period.

Figure 12 compares the CEU and the HD agent by showing average planned consumption, $\tilde{c}_t^h$, and ex-post realizations, $\bar{c}_t$, for the planning age 40. The figure confirms the similarity of the two models: both agents exhibit initial downward jumps in their plans. The main difference stand out at very old target ages, where the CEU plan keeps to be off the realized values.

To understand the nature of the initial blip in both models rewrite the Euler equation

Figure 11: Planned and realized MPC for CEU and HD agents



Notes: Marginal propensity to consume out of cash-on-hand (MPC) of CEU and HD agent at different planning ages $h = 20$ and $h = 45$ compared to life-cycle profile of the ex-post realization for all planning ages.
MPC is approximated by computing averages of $\Delta c/\Delta x$ from the associated policy functions.

of the CEU agent from (13) in absence of binding borrowing constraints as

$$(CEU) \qquad \frac{du}{dc_t^h} = \beta\,(1+r)\cdot E_t\left[\frac{du}{dc_{t+1}^h}\right]\cdot\begin{cases}\left(\delta_h\lambda+(1-\delta_h)\psi_{h,h+1}\right) & \text{for } t=h\\ \frac{\delta_h\lambda+(1-\delta_h)\psi_{h,t+1}}{\delta_h\lambda+(1-\delta_h)\psi_{h,t}} & \text{for } t>h.\end{cases}$$

Hence, initially (planned) consumption growth is inter alia determined by the subjective probability to survive to the next period. In later periods, it is the ratio of subjective probabilities between successive periods. The limit of the Euler equation for $h \to \infty$, hence $\delta_h \to 1$, is

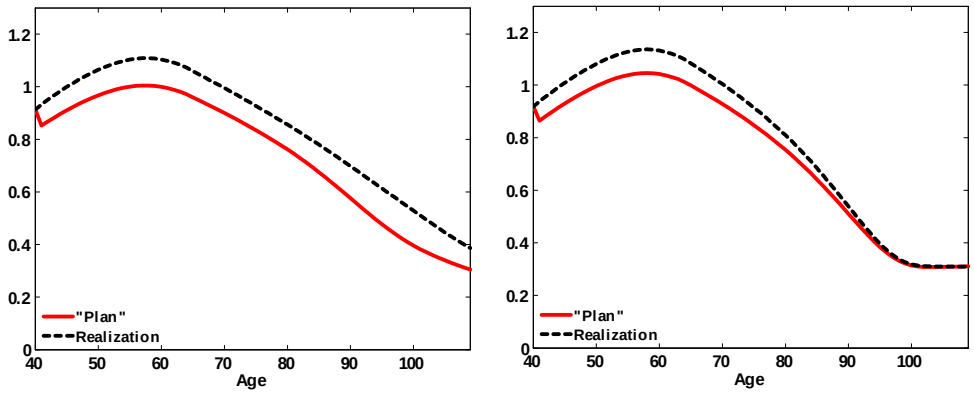
$$(CEU) \qquad \lim_{h\to\infty}\frac{du}{dc_t^h} = \beta\,(1+r)\cdot E_t\left[\frac{du}{dc_{t+1}^h}\right]\cdot\begin{cases}\lambda & \text{for } t=h\\ 1 & \text{for } t>h.\end{cases}$$

Because $\psi_{h,h+1}$ also monotonically decreases with increasing h, the initial drop in consumption plans increases in h.

In the limit this equation closely resembles the HD Euler-equation (14) which was

Figure 12: Planned and Realized Consumption at Planning Age 40

(a) Naive CEU (b) Naive HD



Notes: Planned and realized consumption profiles at planning age $h = 40$ for CEU and HD agent.

given by

$$(HD) \qquad \frac{du}{dc_t} \geq \beta\,(1+r)\,\psi_{t,t+1} \cdot E_t\left[\frac{du}{dc_{t+1}}\right] \cdot \begin{cases} \gamma, & \text{for } t = h \\ 1, & \text{for } t > h \end{cases} .$$

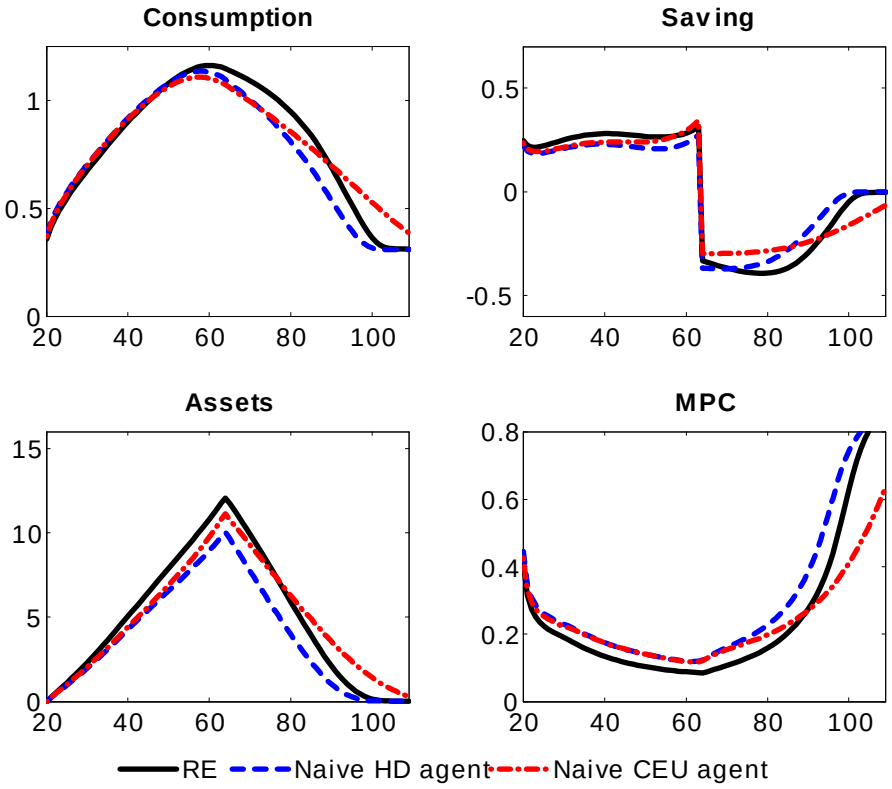
Comparison with Rational Expectations Benchmark Model

In this section we compare the CEU and the HD model with the rational expectations (RE) agent as a benchmark model. The RE agent uses objective survival probabilities and an exponential discount function with a time independent discount factor β. Figure 13 concentrates on the life-cycle profiles of the naive CEU and HD agents compared to the RE agent.

Both the naive CEU and naive HD agents value current consumption higher than RE agents. Thus, the naive agents save less during working ages which is in line with the empirical facts of undersaving outlined in the introduction. As a consequence, less assets are accumulated. For HD agents asset holdings are lower throughout their life which is in contrast to empirical findings indicating that elderly hold on to their assets. In contrast, our CEU model can account for both empirical facts at once. Naive CEU agents consume more and correspondingly undersave at working ages compared to RE agents. At the same time, assets are decumulated much slower. The naive CEU agents finally have higher asset holdings at ages 76+ compared to agents with rational expectations. This result stems from the overestimation of survival beliefs at older ages. The naive CEU model can thus simultaneously account for undersaving and high old-age asset holdings while the naive HD model can only account for the former.

Figure 14 show life-cycle profiles for the sophisticated agent types. Quantitatively, the sophisticated CEU and HD agents behave much more alike than the naive agents. Now,

Figure 13: Life-Cycle Profiles of Naive CEU and HD Agents compared to RE

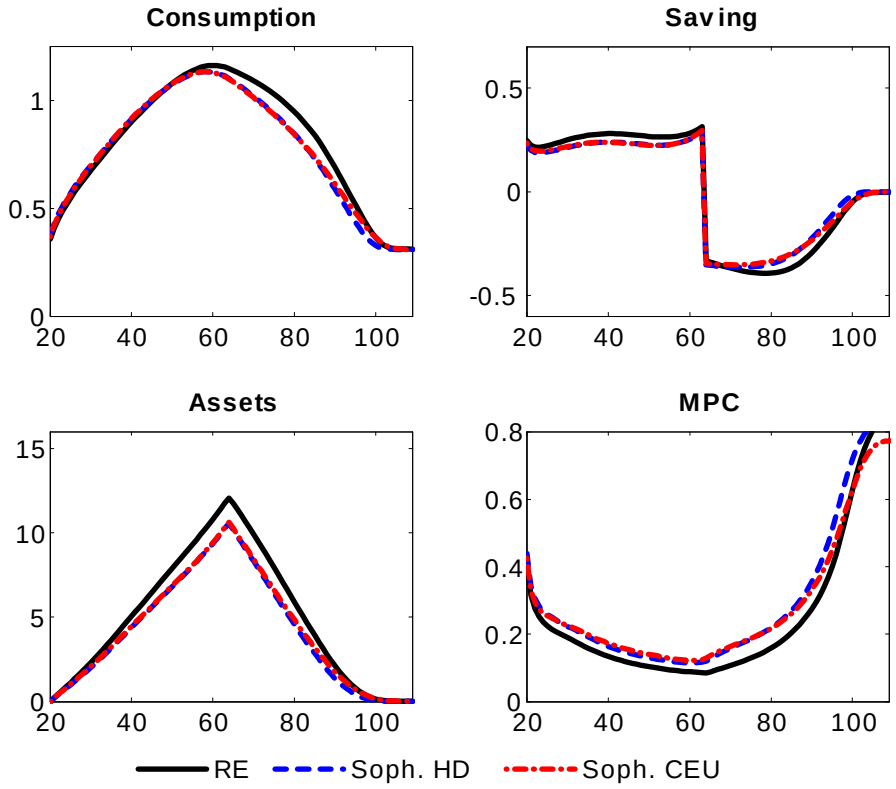


Notes: Realized life-cycle profiles of the CEU, the HD and the RE agent. MPC denotes the marginal propensity to consume out of cash-on-hand which is approximated by computing averages of $\Delta c/\Delta x$ from the associated policy functions.

both still exhibit undersaving but as a main difference, the distinct higher asset holdings of the elderly CEU relative to the RE agent is not present for sophisticates. Nevertheless, qualitatively the results still hold: the CEU agent has slightly higher assets at the very last years of life. This is reflected by a lower marginal propensity to consume (MPC) at older ages, cf. lower right panel of Figure 14.

Table 3 comprises these results by reporting summary statistics. The average saving rate of naive CEU agents, defined as the ratio of average savings to average income, during her working life is 1.5 percentage points lower than the average saving rate of RE agents. Generally, the saving rate is lower for HD agents under the current calibration. The saving rate of the naive HD agent is 3.3 percentage points lower than for the RE agent. On the other hand, the asset holdings of the elderly are even lower for the HD model than for the RE model, thus, the hyperbolic discounting model can not replicate the empirical fact of high asset holdings at older ages, reported by Hurd and Rohwedder (2010), and others. Assets of a naive HD agent at age 85 (95) relative to her assets at retirement entry are only 26.2% (0.76%) while these values are much higher for the CEU agent with 46.4 (23.5). This difference is especially strong for the very old agents of 90+. Note, that assets of the sophisticated agents are generally lower compared to the RE

Figure 14: Life-Cycle Profiles of Sophisticated CEU and HD Agents compared to RE



Notes: realized ex-post life-cycle profile of the CEU and the RE agent.

counterpart. Only at very old ages, assets are (still) higher for the sophisticated CEU agent: Assets at 95+ relative to assets at age 65 are 8.43% relative to 8.39% for the RE agent. Thus, quantitatively, the high asset holdings are only present for the naive CEU agents.

8 Conclusion

We compare a decision-theoretic model of subjective survival belief formation (CEU) proposed by Ludwig and Zimper (2013) with the well-known (quasi-)hyperbolic discounting model (HD) in a life-cycle model. We derive the Euler-equations for naive and sophisticated agents and calibrate the model to compare the quantitative implications of both models for consumption and saving behavior. As our main result, the two models differ with regard to old-age asset holdings. Due to an increasing optimism at older ages—which is in line with the data in subjective survival beliefs—CEU agents hold on to their assets late in life which accommodates an important empirical puzzle in the life-cycle literature. This stylized fact cannot be addressed by the hyperbolic discounting model.

The decision theoretic nature of the model of subjective survival beliefs entails important distinctions from the hyperbolic discounting model. First, (quasi-)hyperbolic discounting result from—rather ad hoc—assumptions on functional forms. In contrast,

Table 3: Summary Statistics

		RE	*CEU*		*HD*	
			Naive	Soph.	Naive	Soph.
Ratio of max. consumption[1)]		2.12	1.95	2.00	1.98	1.98
Age at max. consumption		60	58	58	58	58
Saving rate[2)]		23.0%	21.5%	20.5%	19.7%	19.3%
Assets at 85 to 65[3)]	Avg	35.5%	46.4%	33.0%	26.2%	29.9%
	Med	31.8%	36.2%	26.5%	17.6%	22.8%
Assets at 95 to 65[3)]	Avg	8.39%	23.5%	8.43%	3.9%	5.5%
	Med	5.9%	11.2%	2.3%	0.76%	0.72%
Assets at 85+ to lifetime avg[4)]		52.2%	106.9%			

[1)] Maximal consumption relative to consumption at age 25

[2)] We define the "average" saving rate as the ratio of averages during working life. We hence compute $\sum s_t / \sum y_t$

[3)] Assets of age 85 (95) relative to assets at retirement entry.

[4)] Percentage difference of average assets during ages 85-110 relative to average assets through whole life.

our model is based on an axiomatic decision theoretic model. Second, the effective hyperbolic discounting functions in our model are not stationary but evolve as the agent receives new information. Third, a key problem of calibrating quasi-hyperbolic discounting models is that additional preference parameters reflecting the degree of present-bias are not observable. There is not much consensus concerning their value. In contrast, our information based model exploits the information on subjective survival beliefs in the data.

Further differences of the model of subjective survival beliefs compared to the hyperbolic discounting model arise once we add an altruistic bequest motive in the standardly used warm-glow fashion as suggested by Andreoni (1989). In addition to receiving utility from consumption parents derive utility from leaving bequest to their children in case they die. Since potential bequests are weighed with the death rate we have a different weighting of the warm-glow of giving between the agent using subjective survival belief and a hyperbolic discounter. We leave this extension for future research.

References

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

Andreoni, J. (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. *Journal of Political Economy 97*(6), 1447–1458.

Angeletos, G.-M., D. Laibson, A. Repetto, J. Tobacman, and S. Weinberg (2001). The hyperbolic consumption model: Calibration, simulation, and empirical evaluation. *The Journal of Economic Perspectives 15*(3), 47–68.

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.

Bernheim, D. B. (1998). Financial Illiteracy, Education, and Retirement Saving. In O. S. Mitchell and S. J. Schieber (Eds.), *Living with Defined Contribution Pensions*, pp. 38–68. The Pension Research Council.

Carroll, C. D. (1992). The buffer-stock theory of saving: Some macroeconomic evidence. *Brookings Papers on Economic Activity 23*(2), 61–156.

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 137*(1), 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.

De Nardi, M. (2004). Wealth Inequality and Intergenerational Links. *Review of Economic Studies 71*, 743–768.

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

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

Eichberger, J., S. Grant, and D. Kelsey (2007). Updating Choquet beliefs. *Journal of Mathematical Economics 43*(7-8), 888–899.

Friedman, B. M. and M. J. Warshawsky (1990). The Cost of Annuities: Implications for Saving Behavior and Bequests. *The Quarterly Journal of Economics 105*(1), 135–154.

Gilboa, I. (1987). Expected Utility with Purely Subjective Non-Additive Probabilities. *Journal of Mathematical Economics 16*(1), 65–88.

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

Gourinchas, P.-O. and J. Parker (2002). Consumption over the Lifecycle. *Econometrica 70(1)*, 47–91.

Groneck, M., A. Ludwig, and L. Zimper (2013). A life-cycle model with ambiguous survival beliefs. *University of Cologne Working Paper Series in Econcomis 63.*

Gul, F. and W. Pesendorfer (2001). Temptation and Self-Control. *Econometrica 69*(6), 1403–1435.

Gul, F. and W. Pesendorfer (2004). Self-Control and the Theory of Consumption. *Econometrica 72*(1), 119–158.

Hall, R. E. and C. I. Jones (2007). The Value of Life and the Rise in Health Spending. *The Quarterly Journal of Economics 122*(1), 39–72.

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.

Hubbard, R. G., J. Skinner, and S. P. Zeldes (1995). Precautionary Saving and Social Insurance. *The Journal of Political Economy 103*(2), 360–399.

Huggett, M., G. Ventura, and A. Yaron (2007). Sources of Lifetime Inequality. *NBER Working Paper 13224*.

Hurd, M. (1989). Mortality Risk and Bequests. *Econometrica 57*(4), 173–209.

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

Imrohoroglu, A., S. Imrohoroglu, and D. H. Joines (1995). A Life Cycle Analysis of Social Securtiy. *Economic Theory 6*, 83–114.

Krusell, P. and A. A. Smith (2003). Consumption-savings decisions with quasi-geometric discounting. *Econometrica 71*(1), 365–375.

Laibson, D. (1997, May). 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, and Tobacman (1998). Self-Control and Saving for Retirement. *Brookings Papers on Economic Activity 1998*(1), 91–196.

Lee, R. D. and L. Carter (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association 87*(419), 659–671.

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

Loewenstein, G. and D. Prelec (1992). Anomalies in Intertemporal Choice: Evidence and an Interpretation. *The Quarterly Journal of Economics 107*(2), 573–597.

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. *Forthcoming in: Theory and Decision*.

Munnell, A., A. Webb, and F. Golub-Sass (2010). How will higher tax rates affect the National Retirement Risk Index? Brief of the Center for Retirement Research.

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

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.

Pires, C. P. (2002). A Rule for Updating Ambiguous Beliefs. *Theory and Decision 53*(2), 137–152.

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

Preston, S. H., P. Heuveline, and M. Guillot (2001). *Domography - Measuring and Modelling Population Processes*. Blackwell Publishers Ltd.

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

Savage, L. J. (1954). *The Foundations of Statistics*. New York, London, Sydney: John Wiley and Sons, Inc.

Schmeidler, D. (1986). Integral Representation Without Additivity. *Proceedings of the American Mathematical Society 97*, 255–261.

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

Siegel, J. J. (2002). *Stocks for the Long Run : The Definitive Guide to Financial Market Returns and Long-Term Investment Strategies*. New York: McGraw-Hill.

Siniscalchi, M. (2011). Dynamic Choice under Ambiguity. *Theoretical Economics 6*(3), 379–421.

Storesletten, K., C. Telmer, and A. Yaron (2004). Cyclical Dynamics in Idiosyncratic Labor Market Risk. *Journal of Political Economy 112*(3), 695–717.

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

Strotz, R. H. (1956). Myopia and Inconsistency in Dynamic Utility Maximization. *Review of Economic Studies 23 (3)*, 165–180.

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

Wakker, P. (2004). On the Composition of Risk Preference and Belief. *Psychological Review 111*(1), 236–241.

Wakker, P. (2010). *Prospect Theory for Risk and Ambiguity*. Cambridge University Press.

Weon, B. and J. Je (2009). Theoretical Estimation of Maximum Human Lifespan. *Biogerontology 10*, 65–71.

Wu, G. and R. Gonzalez (1996). Curvature of the probability weighting function. *Management Science 42*(12), 1676–1690.

Zimper, A. (2012). Asset pricing in a lucas fruit-tree economy with the best and worst in mind. *Journal of Economic Dynamics and Control 36*(4), 610–628.

A Appendix: The CEU model with Subjective Survival Beliefs

The following Appendix (i) provides decision theoretic preliminaries, (ii) presents the Bayesian Learning of ambiguous survival beliefs as introduced by Ludwig and Zimper (2013), and finally (iii) presents how to merge this model into a life-cycle, cf. Groneck, Ludwig, and Zimper (2013).

A.1 Decision Theoretic Preliminaries

A.1.1 Choquet Integration and Neo-additive Capacities

Consider a measurable space $(\Omega, \mathcal{F})$ with $\mathcal{F}$ denoting a σ-algebra on the state space Ω and a non-additive probability measure (=*capacity*) $\kappa : \mathcal{F} \to [0,1]$ satisfying

(i) $\kappa(\emptyset) = 0$, $\kappa(\Omega) = 1$

(ii) $A \subset B \Rightarrow \kappa(A) \leq \kappa(B)$ for all $A, B \in \mathcal{F}$.

The Choquet integral of a bounded $\mathcal{F}$-measurable function $f : \Omega \to \mathbb{R}$ with respect to capacity κ is defined as the following Riemann integral extended to domain Ω (Schmeidler 1986):

$$E[f,\kappa] = \int_{-\infty}^{0} (\kappa(\{\omega \in \Omega \mid f(\omega) \geq z\}) - 1)\, dz + \int_{0}^{+\infty} \kappa(\{\omega \in \Omega \mid f(\omega) \geq z\})\, dz. \quad (31)$$

For example, assume that f takes on m different values such that $A_1, ..., A_m$ is the unique partition of Ω with $f(\omega_1) > ... > f(\omega_m)$ for $\omega_i \in A_i$. Then the Choquet expectation (31) becomes

$$E[f,\kappa] = \sum_{i=1}^{m} f(\omega_i) \cdot [\kappa(A_1 \cup ... \cup A_i) - \kappa(A_1 \cup ... \cup A_{i-1})].$$

This paper focuses on non-additive probability measures that are defined as *neo-additive capacities* in the sense of Chateauneuf et al. (2007). Recall that the set of *null events*, denoted $\mathcal{N}$, collects all events that the decision maker deems impossible.

Definition 4 *Fix some set of null-events $\mathcal{N} \subset \mathcal{F}$ for the measurable space $(\Omega, \mathcal{F})$. The neo-additive capacity, ν, is defined, for some $\delta, \lambda \in [0,1]$ by*

$$\nu(A) = \delta \cdot \nu_\lambda(A) + (1-\delta) \cdot \mu(A) \quad (32)$$

for all $A \in \mathcal{F}$ such that μ is some additive probability measure satisfying

$$\mu(A) = \begin{cases} 0 & \text{if } A \in \mathcal{N} \\ 1 & \text{if } \Omega \backslash A \in \mathcal{N} \end{cases}$$

and the non-additive probability measure ν_λ is defined as follows

$$\nu_\lambda(A) = \begin{cases} 0 & \textit{iff } A \in \mathcal{N} \\ \lambda & \textit{else} \\ 1 & \textit{iff } \Omega \backslash A \in \mathcal{N}. \end{cases}$$

In this paper, we are exclusively concerned with the empty set as the only null event, i.e., $\mathcal{N} = \{\emptyset\}$. In this case, the neo-additive capacity ν in (32) simplifies to

$$\nu(A) = \delta \cdot \lambda + (1 - \delta) \cdot \mu(A)$$

for all $A \neq \emptyset, \Omega$. The following observation extends a result (Lemma 3.1) of Chateauneuf et al. (2007) for finite random variables to the more general case of random variables with a bounded range (see Zimper 2012 for a formal proof).

Observation 11 *Let $f : \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable function with bounded range. The Choquet expected value (31) of f with respect to a neo-additive capacity (32) is then given by*

$$E[f, \nu] = \delta(\lambda \sup f + (1 - \lambda) \inf f) + (1 - \delta) E[f, \mu]. \tag{33}$$

According to Observation 11, the Choquet expected value of a random variable f with respect to a neo-additive capacity is a convex combination of the expected value of f with respect to some additive probability measure μ and an ambiguity part. If there is no ambiguity, i.e., $\delta = 0$, then the Choquet expected value (33) reduces to the standard expected value of a random variable with respect to an additive probability measure. In case there is some ambiguity, however, the second parameter λ measures how much weight the decision maker puts on the least upper bound of the range of f. Conversely, $(1 - \lambda)$ is the weight he puts on the greatest lower bound.

A.1.2 The Generalized Bayesian Update Rule

CEU theory has been developed in order to accommodate paradoxes of the Ellsberg type which show that real-life decision-makers violate Savage's sure thing principle Savage (1954). Abandoning of the sure thing principle has two important implications for conditional CEU preferences. First, in contrast to Bayesian updating of additive probability measures, there exist several perceivable Bayesian update rules for non-additive probability measures (Gilboa and Schmeidler 1993; Pires 2002; Eichberger, Grant, and Kelsey 2007; Siniscalchi 2011). Second, if CEU preferences are updated in accordance with an updating rule that universally satisfies the principle of consequentialism, then these CEU preferences violate the principle of dynamic consistency (in a universal sense) whenever they do not reduce to EU preferences (cf. Zimper 2012 and references therein).

In the present paper we assume that the agents form conditional capacities in accordance with the Generalized Bayesian update rule such that, for all non-null $A, B \in \mathcal{F}$,

$$\kappa(A \mid B) = \frac{\kappa(A \cap B)}{\kappa(A \cap B) + 1 - \kappa(A \cup \neg B)}. \tag{34}$$

An application of (34) to a neo-additive capacity ν gives rise to the following observation.

Observation 12 *If the Generalized Bayesian update rule (34) is applied to a neo-additive capacity (32), we obtain, for all non-null $A, B \in \mathcal{F}$,*

$$\nu(A \mid B) = \delta_B \cdot \lambda + (1 - \delta_B) \cdot \mu(A \mid B)$$

such that

$$\delta_B = \frac{\delta}{\delta + (1 - \delta) \cdot \mu(B)}.$$

A.2 Bayesian Learning of Ambiguous Survival Beliefs

This appendix briefly recalls the learning model of ambiguous survival beliefs as introduced in Ludwig and Zimper (2013). We consider an h-old agent, with $0 \leq h \leq k$, who observes the random sample information $\tilde{I}_{n(h)}$ which counts how many individuals out of a sample of size $n(h)$ have survived from age k to t. By assumption, these individuals have the same i.i.d. objective survival probability as the agent.

The Benchmark Case of Additive Survival Beliefs

At first, consider a standard Bayesian decision maker whose additive estimator for the chance of surviving from k to t conditional on $\tilde{I}_{n(h)}$ is defined as the conditional expected value

$$E\left[\tilde{\theta}, \mu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right]$$

where the random variable $\tilde{\theta}$ stands for the agent's survival chance with support on $(0, 1)$. By the i.i.d. assumption of individual survivals, $\tilde{I}_{n(h)}$ is, conditional on the true survival probability $\tilde{\theta} = \theta$, binomially distributed with probabilities

$$\mu\left(\tilde{I}_{n(h)} = j \mid \theta\right) = \binom{n(h)}{j} \theta^j (1 - \theta)^{n-j} \text{ for } j \in \{0, ..., n(h)\}.$$

We further assume that the agent's prior over $\tilde{\theta}$ is given as a Beta distribution with parameters $\alpha, \beta > 0$, implying $E\left[\tilde{\theta}, \mu\left(\tilde{\theta}\right)\right] = \frac{\alpha}{\alpha+\beta}$. That is, we assume that

$$\mu\left(\tilde{\theta} = \theta\right) = K_{\alpha,\beta} \theta^{\alpha-1} (1 - \theta)^{\beta-1}$$

where $K_{\alpha,\beta} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$ is a normalizing constant.[24]

By Bayes' rule we obtain the following conditional distribution of $\tilde{\theta}$

$$\begin{aligned}\mu\left(\tilde{\theta}=\theta \mid \tilde{I}_{n(h)}=j\right) &= \frac{\mu\left(\tilde{I}_{n(h)}=j \mid \theta\right)\mu(\theta)}{\int_{(0,1)}\mu\left(\tilde{I}_{n(h)}=j \mid \theta\right)\mu(\theta)\,d\theta} \\ &= K^{\alpha+j-1}_{\alpha+j,\beta+n(h)-k}\theta^{\alpha+j-1}(1-\theta)^{\beta+n(h)-j-1} \text{ for } \theta\in(0,1)\end{aligned}$$

Note that $\mu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}=j\right)$ is itself a Beta distribution with parameters $\alpha+j, \beta+n(h)-j$. The agent's subjective survival belief conditional on information $\tilde{I}_{n(h)}=j$ is thus given as

$$\begin{aligned}E\left[\tilde{\theta}, \mu\left(\tilde{\theta} \mid j\right)\right] &= \frac{\alpha+j}{\alpha+\beta+n(h)} \\ &= \left(\frac{\alpha+\beta}{\alpha+\beta+n(h)}\right)E\left[\tilde{\theta}, \mu\left(\tilde{\theta}\right)\right]+\left(\frac{n(h)}{\alpha+\beta+n(h)}\right)\frac{j}{n(h)}, \\ \text{for } j &\in \{0,...,n(h)\}.\end{aligned}$$

That is, the posterior estimator $E\left[\tilde{\theta}, \mu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right]$ is a weighted average of her prior survival probability $E\left[\tilde{\theta}, \mu\left(\tilde{\theta}\right)\right]$, not including any sample information, and the observed sample mean $\frac{j}{n(h)}$.

Ambiguous Survival Beliefs in a parsimonious model

Turn now to a Choquet decision maker with neo-additive capacity

$$\nu\left(\tilde{\theta}\right) = \delta\cdot\lambda+(1-\delta)\cdot\mu\left(\tilde{\theta}\right)$$

such that the conditional neo-additive capacity $\nu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)$ results from an application of the Generalized Bayesian update rule. Instead of the additive estimator $E\left[\tilde{\theta}, \mu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right]$ we now suppose that the agent's estimator for her survival chance is given as the conditional Choquet expected value

$$E\left[\tilde{\theta}, \nu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right] = \delta_{\tilde{I}_{n(h)}}\left(\lambda\sup\tilde{\theta}+(1-\lambda)\inf\tilde{\theta}\right)+\left(1-\delta_{\tilde{I}_{n(h)}}\right)E\left[\tilde{\theta}, \mu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right].$$

For a Beta distribution $\mu\left(\tilde{\theta}\right)$, Ludwig and Zimper (2013) prove the following result:

[24] The gamma function is defined as $\Gamma(y) = \int_0^\infty x^{y-1}e^{-x}dx$ for $y>0$.

Observation 13 *The Choquet decision maker's ambiguous survival belief is given as*

$$E\left[\tilde{\theta}, \nu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right] = \delta_{\tilde{I}_{n(h)}} \cdot \lambda + \left(1 - \delta_{\tilde{I}_{n(h)}}\right) \cdot E\left[\tilde{\theta}, \mu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right], \tag{35}$$

with

$$\delta_{\tilde{I}_{n(h)}} = \frac{\delta}{\delta + (1-\delta) \cdot \mu\left(\tilde{I}_{n(h)}\right)}$$

where the unconditional distribution of $\tilde{I}_{n(h)}$ *is given by*

$$\begin{aligned} \mu\left(\tilde{I}_{n(h)} = j\right) &= \binom{n(h)}{j} \frac{(\alpha + j - 1) \cdot ... \cdot \alpha \cdot (\beta + n(h) - j - 1) \cdot ... \cdot \beta}{(\alpha + \beta + n(h) - 1) \cdot ... \cdot (\alpha + \beta)}, \\ \text{for } j &\in \{0, ..., n(h)\}. \end{aligned} \tag{36}$$

Finally, to derive from (35) the parsimonious characterization of ambiguous survival beliefs in Observation 1, we employ several simplifying assumptions:

Assumption 2 *The additive part* $E\left[\tilde{\theta}, \nu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right]$ *is, for any information* $\tilde{I}_{n(h)}$*, given as the objective probability, denoted* $\psi_{k,t}$*, to survive from age* k *to* t*.*

Assumption 3 *The agent's additive prior over the parameter space is given as a uniform distribution, i.e., a Beta distribution with parameters* $\alpha = \beta = 1$*, implying for (36) that* $\mu\left(\tilde{I}_{n(h)} = j\right) = \frac{1}{1+n(h)}$*.*

Assumption 4 *The age-dependent sample-size function is given as*

$$n(h) = \sqrt{h} \text{ for } h \leq T$$

which implies, by Assumption 3, that

$$\mu\left(\tilde{I}_{n(h)} = j\right) = \frac{1}{1 + n(h)}, \text{ for } j \in \{0, ..., n(h)\}.$$

Assumption 2 is an extreme version of the rational Bayesian learning part of the model developed in Appendix A.2. It specifies a fully informed prior and hence simplifies upon Ludwig and Zimper (2013).[25] By this assumption any difference between subjective survival beliefs and objective survival probabilities are exclusively driven by the ambiguity part of the agent's belief. Assumption 3 allows for an explicit expression of the unconditional probability $\mu\left(\tilde{I}_{n(h)}\right)$ which only depends on age h, i.e., it is identical for every observed sample information $\tilde{I}_{n(h)}$ if h is fixed. By assumption 4, the agent observes a strictly increasing sample while growing older.

[25] Ludwig and Zimper (2013) are more explicit about the rational Bayesian learning part of the model and assume a proportional bias in prior additive beliefs.

Given this assumption Ludwig and Zimper (2013) derive the subjective survival belief given in equation (1).

A.3 CEU Preferences in the Life-Cycle Setting

It remains to translate the notion of ambiguous survival beliefs, cf. equation (1) into the construction of the relevant conditional neo-additive probability space $(\Omega, \mathcal{F}, \nu(\cdot \mid \cdot))$. To this purpose define the finite state space $\Omega = \{0, 1, ..., T\}$ and let the σ-algebra $\mathcal{F}$ be the powerset of Ω. We interpret $D_t = \{t\}$, $t \in \Omega$ as the event in $\mathcal{F}$ that the agent dies at the end of period t. Define age h of the agent as the following event in $\mathcal{F}$: $h = D_h \cup ... \cup D_T$. Further, formally define $Z_{k,t} = D_t \cup ... \cup D_T$ as the event in $\mathcal{F}$ that the agent survives from period k to the beginning of period t.

The relevant information filtration of our model is simply given by $\mathcal{F}_1 \subset ... \subset \mathcal{F}_T = \mathcal{F}$ such that, for each age h, $\mathcal{F}_h$ is generated by the following partition of Ω: $\{\{0\}, .., \{h-1\}, \{h, ..., T\}\}$. That is, if the agent turns age h she (trivially) observes that she has not died in any previous period but will die at the end of either period h or $h+1$ or ... or T.

Finally, we assume that $\nu(\cdot \mid \cdot)$ satisfies (i) $\nu(\emptyset \mid \cdot) = 0$, $\nu(\Omega \mid \cdot) = 1$, (ii) $\nu(\cdot \mid \cdot)$ is a conditional neo-additive capacity in the sense of Chateauneuf et al. (2007) which is updated in accordance with the Generalized Bayesian update rule where (iii) for all h, $Z_{k,t} \neq \emptyset$ and $Z_{k,t} \neq \Omega$ with $h \leq k < T$ and $k < t \leq T$, $\nu(Z_{k,t} \mid h) \equiv \nu^h_{k,t}$ with $\nu^h_{k,t}$ given by (1).

We assume additive separability and discounting at rate β. For any $s \in \{h, h+1, ..., T\}$ and survival until period s, the agent's von Neumann Morgenstern utility (vNM) from a consumption plan $\mathbf{c}$ is then defined as

$$U(\mathbf{c}(s)) = u(c_h) + \sum_{t=h+1}^{s} \beta^{t-h} E_t[u(c_t)].$$

Denote by $\nu^h \equiv \nu(\cdot \mid h)$ the agent's age-conditional neo-additive capacity. In order to formalize utility maximization over life-time consumption with respect to neo-additive probability measures, we henceforth describe an h-old agent as a CEU decision maker who maximizes her Choquet expected utility from life-time consumption with respect to ν^h. By Observation 11 in Appendix A.1, this agent's CEU from consumption plan $\mathbf{c}$ with respect to ν^h is given as

$$\begin{aligned} E\left[U(\mathbf{c}), \nu^h\right] &= \delta_h \left[\lambda \sup_{s \in \{h, h+1, ...\}} U(\mathbf{c}(s)) + (1-\lambda) \inf_{s \in \{h, h+1, ...\}} U(\mathbf{c}(s))\right] \\ &\quad + (1-\delta_h) \cdot \sum_{s=h}^{T} [U(\mathbf{c}(s)), \psi(D_s)] \end{aligned} \tag{37}$$

where we have for any $\mathbf{c}$ that

$$\begin{aligned}\sup_{s\in\{h,h+1,\ldots\}} U\left(\mathbf{c}\left(s\right)\right) &= u(c_h) + \sum_{t=h+1}^{T} \beta^{t-h} E_t\left[u\left(c_t\right)\right], \\ \inf_{s\in\{h,h+1,\ldots\}} U\left(\mathbf{c}\left(s\right)\right) &= u(c_h).\end{aligned}$$

Observe that we require

$$\sum_{t=h+1}^{T} \beta^{t-h} E_h\left[u\left(c_t\right)\right] > 0, \qquad \forall\, h = 0, \ldots \tag{38}$$

The objective probability to survive until period t is given as

$$\psi_{h,t} = \prod_{s=h}^{t-1} \psi_{s,s+1}$$

implying

$$\psi_{h,t} = \sum_{s=t+1}^{T} \psi^h(D_s)$$

where D_t denotes the event that the agent dies at the end of period t. Consequently, (37)

can be equivalently written in terms of survival beliefs as

$$
\begin{aligned}
E\left[U\left(\mathbf{c}\right),\nu^{h}\right] &= \delta_{h}\left(\lambda\left(u(c_{h})+\sum_{t=h+1}^{T}\beta^{t-h}E\left[u\left(c_{t}\right),\pi\left(\eta_{t}|\eta_{h}\right)\right]\right)+(1-\lambda)u(c_{h})\right)\\
&\quad+(1-\delta_{h})\left(u(c_{h})+\sum_{t=h+1}^{T}\psi^{h}(D_{t})\sum_{s=h+1}^{t}\beta^{s-h}E\left[u\left(c_{s}\right),\pi\left(\eta_{s}|\eta_{h}\right)\right]\right)\\
&= u(c_{h})+\delta_{h}\lambda\sum_{t=h+1}^{T}\beta^{t-h}E\left[u\left(c_{t}\right),\pi\left(\eta_{t}|\eta_{h}\right)\right]\\
&\quad+(1-\delta_{h})\sum_{t=h+1}^{T}\psi^{h}(D_{t})\sum_{s=h+1}^{t}\beta^{s-h}E\left[u\left(c_{s}\right),\pi\left(\eta_{s}|\eta_{h}\right)\right]\\
&= u(c_{h})+\delta_{h}\lambda\sum_{t=h+1}^{T}\beta^{t-h}E\left[u\left(c_{t}\right),\pi\left(\eta_{t}|\eta_{h}\right)\right]\\
&\quad+(1-\delta_{h})\sum_{t=h+1}^{T}\psi_{h,t}\cdot\beta^{t-h}E\left[u\left(c_{t}\right),\pi\left(\eta_{t}|\eta_{h}\right)\right]\\
&= u(c_{h})+\sum_{t=h+1}^{T}\left(\delta_{h}\lambda+(1-\delta_{h})\psi_{h,t}\right)\beta^{t-h}E\left[u\left(c_{t}\right),\pi\left(\eta_{t}|\eta_{h}\right)\right]\\
&= u(c_{h})+\sum_{t=h+1}^{T}\nu_{h,t}^{h}\beta^{t-h}E\left[u\left(c_{t}\right),\pi\left(\eta_{t}|\eta_{h}\right)\right].
\end{aligned}
$$

which gives equation (8) in Observation 2.

B Appendix: Formal Proofs

B.1 Proof of Naive CEU Euler-equation

Maximization Problem

$$
\begin{aligned}
V_{t}^{h}\left(x_{t},\eta_{t}\right) &= \max_{c_{t},x_{t+1}}\left\{u\left(c_{t}\right)+\beta\frac{\nu_{h,t+1}^{h}}{\nu_{h,t}^{h}}E_{t}\left[V_{t+1}^{h}\left(x_{t+1},\eta_{t+1}\right)\right]\right\} \qquad (39)\\
&\quad s.t.\\
x_{t+1} &= \left(x_{t}-c_{t}\right)\left(1+r\right)+y_{t+1}\geq 0, \qquad (40)
\end{aligned}
$$

Taking the first order condition with respect to the choice variable c_t gives

$$
\frac{du}{dc_{t}}-\beta\left(1+r\right)\frac{\nu_{h,t+1}^{h}}{\nu_{h,t}^{h}}\sum_{\eta^{t+1}}\pi\left(\eta^{t+1}\mid\eta^{t}\right)\frac{\partial V_{t+1}^{h}}{\partial x_{t+1}}=0 \qquad (41)
$$

Taking the partial derivative of the value function w.r.t. x_t gives:

$$\frac{\partial V_t^h}{\partial x_t} = \beta\left(1+r\right)\frac{\nu_{h,t+1}^h}{\nu_{h,t}^h}\sum_{\eta^{t+1}}\pi\left(\eta^{t+1}\mid\eta^t\right)\frac{\partial V_{t+1}^h}{\partial x_{t+1}}, \tag{42}$$

where (42) follows from the enveloppe theorem ensuring that $dc_t/dx_t = 0$ in optimum.

From (41) and (42) it follows that for every period t

$$\frac{du}{dc_t} = \frac{\partial V_t^h}{\partial x_t}$$

Thus, combining equations leads to

$$\frac{du}{dc_t} - \beta(1+r)\frac{\nu_{h,t+1}^h}{\nu_{h,t}^h}\sum_{\eta^{t+1}}\pi\left(\eta^{t+1}\mid\eta^t\right)\frac{du}{dc_{t+1}} = 0$$

$$\frac{du}{dc_t} = \beta(1+r)\frac{\nu_{h,t+1}^h}{\nu_{h,t}^h}E_t\left[\frac{du}{dc_{t+1}}\right] \tag{43}$$

.□

B.2 Proof of naive HD agent

Recall the discount function given by

$$\varphi_t^h = \begin{cases}\delta\beta, \text{ for } t = h\\ \beta, \text{ for } t > h\end{cases} \tag{44}$$

Maximization Problem

$$V_t^h\left(x_t,\eta_t\right) = \max_{c_t,x_{t+1}}\left\{u\left(c_t\right)+\varphi_t^h\psi_{t,t+1}\sum_{\eta^{t+1}}\pi\left(\eta^{t+1}|\eta^t\right)V_{t+1}^h\left(x_{t+1},\eta_{t+1}\right)\right\} \tag{45}$$

$$s.t.$$

$$x_{t+1} = \left(x_t-c_t\right)\left(1+r\right)+y_{t+1}\geq 0, \tag{46}$$

Taking the first order condition with respect to the choice variable c_t gives

$$\frac{du}{dc_t} - \varphi_t^h\left(1+r\right)\psi_{t,t+1}\sum_{\eta^{t+1}}\pi\left(\eta^{t+1}|\eta^t\right)\frac{\partial V_{t+1}^h}{\partial x_{t+1}} = 0 \tag{47}$$

Taking the partial derivative of the value function w.r.t. x_t gives:

$$\frac{\partial V_t^h}{\partial x_t} = \varphi_t^h \left(1+r\right) \psi_{t,t+1} \sum_{\eta^{t+1}} \pi\left(\eta^{t+1}|\eta^t\right) \frac{\partial V_{t+1}^h}{\partial x_{t+1}}, \tag{48}$$

where (48) follows from the enveloppe theorem ensuring that $dc_t/dx_t = 0$.

From (47) and (48) it follows that for every period t

$$\frac{du}{dc_t} = \frac{\partial V_t^h}{\partial x_t}$$

Thus, combining equations leads to

$$u_c(c_h^h) \geq \beta \left(1+r\right) \psi_{t,t+1} \sum_{\eta^{t+1}} \pi\left(\eta^{t+1}|\eta^t\right) u_c(c_{t+1}^h) \cdot \begin{cases} \gamma, \text{ for } t = h \\ 1, \text{ for } t > h \end{cases}$$

.□

B.3 Proof of sophisticated CEU

The first order condition of self $h+1$ derived analoguous to equation (41) is given by

$$\frac{du}{dc_{h+1}} = \beta \left(1+r\right) \nu_{h+1,h+2}^{h+1} \sum_{\eta^{h+2}} \pi\left(\eta^{h+2} \mid \eta^{h+1}\right) \frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}} \tag{49}$$

The derivative of the value function for $h+2$ is

$$\frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}} = \frac{du}{dc_{h+2}} \cdot \frac{\partial c_{h+2}}{\partial x_{h+2}} + \beta \left(1+r\right) \frac{\nu_{h+1,h+3}^{h+1}}{\nu_{h+1,h+2}^{h+1}} \left(1 - \frac{\partial c_{h+2}}{\partial x_{h+2}}\right) \cdot \sum_{\eta^{h+3}} \pi\left(\eta^{h+3} \mid \eta^{h+2}\right) \frac{\partial V_{h+3}^{h+1}}{\partial x_{h+3}} \tag{50}$$

Note, that the enveloppe theorem does not hold so that $\frac{du}{dc_{h+2}} \cdot \frac{\partial c_{h+2}}{\partial x_{h+2}} \neq 0$.

Rearrange (49) such that

$$\begin{aligned}
\frac{du}{dc_{h+1}} &= \beta\left(1+r\right)\nu_{h+1,h+2}^{h+1}\sum_{\eta^{h+2}}\pi\left(\eta^{h+2}\mid\eta^{h+1}\right)\frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}} \\
&\quad +\beta\left(1+r\right)\nu_{h+1,h+2}^{h}\sum_{\eta^{h+2}}\pi\left(\eta^{h+2}\mid\eta^{h+1}\right)\frac{\partial V_{h+2}^{h}}{\partial x_{h+2}} \\
&\quad -\beta\left(1+r\right)\nu_{h+1,h+2}^{h}\sum_{\eta^{h+2}}\pi\left(\eta^{h+2}\mid\eta^{h+1}\right)\frac{\partial V_{h+2}^{h}}{\partial x_{h+2}} \\
\beta\left(1+r\right)\sum_{\eta^{h+2}}\pi\left(\eta^{h+2}\mid\eta^{h+1}\right)\frac{\partial V_{h+2}^{h}}{\partial x_{h+2}} &= \frac{1}{\nu_{h+1,h+2}^{h+1}}\frac{\partial u}{\partial c_{h+1}} \\
&\quad +\beta\left(1+r\right)\sum_{\eta^{h+2}}\pi\left(\eta^{h+2}\mid\eta^{h+1}\right)\left[\frac{\partial V_{h+2}^{h}}{\partial x_{h+2}}-\frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}}\right]
\end{aligned}$$

Now plug derivative of the value function for self h (cf. eq. (15) in the main text) to get

$$\begin{aligned}
\frac{\partial V_{h+1}^{h}}{\partial x_{h+1}} &= \frac{du}{dc_{h+1}}\cdot\frac{\partial c_{h+1}}{\partial x_{h+1}}+\beta\left(1+r\right)\frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h}}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\sum_{\eta^{h+2}}\pi\left(\eta^{h+2}\mid\eta^{h+1}\right)\frac{\partial V_{h+2}^{h}}{\partial x_{h+2}} \\
&= \frac{du}{dc_{h+1}}\cdot\frac{\partial c_{h+1}}{\partial x_{h+1}}+\frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h}}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\left(\frac{1}{\nu_{h+1,h+2}^{h+1}}\frac{du}{dc_{h+1}}\right. \\
&\quad \left.+\beta\left(1+r\right)\sum_{\eta^{h+2}}\pi\left(\eta^{h+2}\mid\eta^{h+1}\right)\left[\frac{\partial V_{h+2}^{h}}{\partial x_{h+2}}-\frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}}\right]\right)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial V_{h+1}^{h}}{\partial x_{h+1}} &= \left[\frac{\partial c_{h+1}}{\partial x_{h+1}}+\frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h}\nu_{h+1,h+2}^{h+1}}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\right]\cdot\frac{du}{dc_{h+1}} \\
&\quad +\beta\left(1+r\right)\frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h}}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)E_{h+1}\left[\frac{\partial V_{h+2}^{h}}{\partial x_{h+2}}-\frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}}\right]
\end{aligned}$$

Plug this into (41) from h to $h+1$ yields

$$\begin{aligned}
\frac{du}{dc_{h}} &= \beta\left(1+r\right)\nu_{h,h+1}^{h}\sum_{\eta^{h+1}}\pi\left(\eta^{h+1}\mid\eta^{h}\right)\left\{\left[\frac{\partial c_{h+1}}{\partial x_{h+1}}+\frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h}\nu_{h+1,h+2}^{h+1}}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\right]\cdot\frac{du}{dc_{h+1}}\right\} \\
&\quad +\left[\beta\left(1+r\right)\right]^{2}\nu_{h,h+2}^{h}\sum_{\eta^{h+1}}\pi\left(\eta^{h+1}\mid\eta^{h}\right)\left\{\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\sum_{\eta^{h+2}}\pi\left(\eta^{h+2}\mid\eta^{h+1}\right)\left[\frac{\partial V_{h+2}^{h}}{\partial x_{h+2}}-\frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}}\right]\right\}
\end{aligned}$$

□

B.4 Proof of Sophisticated HD

The first order conditions and the derivative of the value function for self h is given by

$$\begin{aligned} u_c(c_h^h) &= \beta\gamma\left(1+r\right)\psi_{t,t+1}\sum_{\eta^{t+1}}\pi\left(\eta^{t+1}|\eta^t\right)\frac{\partial V_{t+1}^h}{\partial x_{t+1}} \\ \frac{\partial V_{h+1}^h}{\partial x_{h+1}} &= \frac{du}{dc_{h+1}}\cdot\frac{\partial c_{h+1}}{\partial x_{h+1}}+\beta\left(1+r\right)\psi_{t+1,t+2}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\sum_{\eta^{t+2}}\pi\left(\eta^{t+2}|\eta^{t+1}\right)\left[\frac{\partial V_{h+2}^h}{\partial x_{h+2}}\right] \end{aligned}$$

The sophisticated HD agent knows that futures selves will also have a strong present bias and face the same problem (while the naive agent thinks that future selves will behave rational). Thus self h knows that selve $h+1$ will have the same FOC given by

$$\frac{du}{dc_{h+1}} = \beta\gamma\left(1+r\right)\psi_{t+1,t+2}\sum_{\eta^{h+2}}\pi\left(\eta^{t+2}|\eta^{t+1}\right)\frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}} \tag{51}$$

Now observe that the derivative of the value function for future periods are identical because both discount any future period after the successive period with β, i.e. they discount future periods identical, thus for self h and $h+1$ we have:

$$\begin{aligned} \frac{\partial V_{h+2}^h}{\partial x_{h+2}} &= \frac{du}{dc_{h+2}}\cdot\frac{\partial c_{h+2}}{\partial x_{h+2}}+\beta\left(1+r\right)\psi_{t+2,t+3}\left(1-\frac{\partial c_{h+2}}{\partial x_{h+2}}\right)\cdot\sum_{\eta^{h+3}}\pi\left(\eta^{t+3}|\eta^{t+2}\right)\frac{\partial V_{h+3}^h}{\partial x_{h+3}} \\ \frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}} &= \frac{du}{dc_{h+2}}\cdot\frac{\partial c_{h+2}}{\partial x_{h+2}}+\beta\left(1+r\right)\psi_{t+2,t+3}\left(1-\frac{\partial c_{h+2}}{\partial x_{h+2}}\right)\cdot\sum_{\eta^{h+3}}\pi\left(\eta^{t+3}|\eta^{t+2}\right)\frac{\partial V_{h+3}^{h+1}}{\partial x_{h+3}} \end{aligned}$$

Thus we get using $\frac{\partial V_{h+2}^h}{\partial x_{h+2}} = \frac{\partial V_{h+2}^{h+1}}{\partial x_{h+2}}$:

$$\begin{aligned} u_c(c_h^h) &= \beta\gamma\left(1+r\right)\psi_{t,t+1}\sum_{\eta^{t+1}}\pi\left(\eta^{t+1}|\eta^t\right)\left[\frac{du}{dc_{h+1}}\cdot\frac{\partial c_{h+1}}{\partial x_{h+1}}\right. \\ &\quad \left.+\beta\left(1+r\right)\psi_{t+1,t+2}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\sum_{\eta^{t+2}}\pi\left(\eta^{t+2}|\eta^{t+1}\right)\left[\frac{\partial V_{h+2}^h}{\partial x_{h+2}}\right]\right] \\ &= \beta\gamma\left(1+r\right)\psi_{t,t+1}\sum_{\eta^{t+1}}\pi\left(\eta^{t+1}|\eta^t\right)\left[\frac{du}{dc_{h+1}}\cdot\frac{\partial c_{h+1}}{\partial x_{h+1}}\right. \\ &\quad \left.+\frac{\beta\left(1+r\right)\psi_{t+1,t+2}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\sum_{\eta^{t+2}}\pi\left(\eta^{t+2}|\eta^{t+1}\right)\frac{du}{dc_{h+1}}}{\beta\gamma\left(1+r\right)\psi_{t+1,t+2}\sum_{\eta^{h+2}}\pi\left(\eta^{t+2}|\eta^{t+1}\right)}\right] \\ &= \beta\gamma\left(1+r\right)\psi_{t,t+1}\sum_{\eta^{t+1}}\pi\left(\eta^{t+1}|\eta^t\right)\left[\frac{\partial c_{h+1}}{\partial x_{h+1}}+\frac{1}{\gamma}\left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\right]\frac{du}{dc_{h+1}} \end{aligned}$$

$$u_c(c_h^h) = \beta\,(1+r)\,\psi_{t,t+1}\sum_{\eta^{t+1}}\pi\left(\eta^{t+1}|\eta^t\right)\left[\gamma\frac{\partial c_{h+1}}{\partial x_{h+1}} + \left(1-\frac{\partial c_{h+1}}{\partial x_{h+1}}\right)\right]\frac{du}{dc_{h+1}}$$

□

B.5 Proof of Proposition 2

Naive HD agent:

1. In period 2 we have $m_2^n = 1$.

2. In period 1 we have that the naive agent's first-order condition writes as

$$\begin{aligned} c_1^{-\theta} &= \beta\gamma\psi_1 R c_2^{-\theta} \\ &= \beta\gamma\psi_1 R(m_2^n(w_1-c_1)R)^{-\theta} \\ &= \beta\gamma\psi_1 R^{1-\theta}(w_1-c_1)^{-\theta} \\ \Leftrightarrow \qquad c_1 &= b^n\psi_1^{-\frac{1}{\theta}}(w_1-c_1)) \\ \Leftrightarrow \qquad c_1 &= m_1^n w_1 = \frac{\psi_1^{-\frac{1}{\theta}}b^n}{1+\psi_1^{-\frac{1}{\theta}}b^n} = \frac{1}{1+\frac{1}{\psi_1^{-\frac{1}{\theta}}b^n}}w_1 \end{aligned}$$

   where

$$b^n \equiv \gamma^{-\frac{1}{\theta}}\left(\beta R^{1-\theta}\right)^{-\frac{1}{\theta}} = \gamma^{-\frac{1}{\theta}}b > b$$

   because $\gamma \in [0,1]$. Consequently $m_1^n > m_1$.

3. In period 0 we have that the naive agent's first-order condition writes as

$$\begin{aligned} c_0^{-\theta} &= \beta\gamma\psi_0 R c_1^{-\theta} \\ &= \beta\gamma\psi_0 R(m_1(w_0-c_0)R)^{-\theta} \end{aligned}$$

   whereby the second line reflects that the naive agent assumes that period 1 agent is going to have marginal propensity m_1 (and not m_1^n).

   Working on this equation we get

$$\begin{aligned} c_0^{-\theta} &= \beta\gamma\psi_0 R(m_1(w_0-c_0)R)^{-\theta} \\ &= \beta\gamma R^{1-\theta}\psi_0 m_1^{-\theta}(w_0-c_0)^{-\theta} \\ \Leftrightarrow \qquad c_1 &= b^n\psi_0^{-\frac{1}{\theta}}m_1(w_0-c_0) \\ \Leftrightarrow \qquad c_0 &= m_0^n w_1 = \frac{b^n\psi_0^{-\frac{1}{\theta}}m_1}{1+b^n\psi_0^{-\frac{1}{\theta}}m_1}w_1 \end{aligned}$$

Analyzing m_0^n further we get

$$\begin{aligned} m_0^n &= \frac{b^n \psi_0^{-\frac{1}{\theta}} m_1}{1 + b^n \psi_0^{-\frac{1}{\theta}} m_1} \\ &= \frac{1}{1 + \frac{1}{b^n \psi_0^{-\frac{1}{\theta}} m_1}} \end{aligned}$$

For the rational expectations agent we have the corresponding expression $m_0 = \frac{1}{1+\frac{1}{b\psi_0^{-\frac{1}{\theta}} m_1}}$. Because $b^n > b$ we have that $m_0^n > m_0$, just as expected.

4. Period 0: Plan for period 1 gives first-order condition:

$$c_1^{-\theta} = \beta \psi_1 R c_2^{-\theta}$$

because the naive HD agent in peiord 0 thinks that self 1 will be acting rational (using the discount factor β). Thus, the solution planned for period one coincides with the RE solution:

$$c_1^n = m_1 w_1 \quad \text{where} \quad m_1 = \frac{1}{1 + \frac{1}{\psi_1^{-\frac{1}{\theta}} b}} w_1 =$$

B.6 Proof of Proposition 3

Naive CEU agent:

1. First-order condition in period 1 is:

$$u_c(c_1) = \beta R \nu_{1,2}^1 u_c(c_2)$$

which gives

$$c_1^{1,n} = m_1^{1,n} w_1 \quad where \quad m_1^{1,n} = \frac{1}{1 + \frac{1}{b\left(\nu_{1,2}^1\right)^{-\frac{1}{\theta}}}}.$$

2. Period 0: Plan for period 1 gives the first-order condition:

$$u_c(c_1) = \beta R \frac{\nu_{0,2}^0}{\nu_{0,1}^0} u_c(c_2)$$

which yields

$$c_1^{1,n} = m_1^{1,n} w_1 \quad \text{where} \quad m_1^{1,n} = \frac{1}{1 + \frac{1}{b\left(\frac{\nu_{0,2}^0}{\nu_{0,1}^0}\right)^{-\frac{1}{\theta}}}}$$

3. First-order condition in period 0 is:

$$u_c(c_0) = \beta R \nu_{0,1}^0 u_c(c_1)$$

yielding

$$c_0 = m_0^{0,n} w_0 = \frac{1}{1 + \frac{1}{b(\nu_{0,1}^0)^{-\frac{1}{\theta}} m_1^{0,n}}} w_0$$

B.7 Proof of Observation 9

We have already established above that

$$m_1^{0,n} < m_1,$$

i.e., that the CEU household plans to save more in the second period than the RE household. In order to get that

$$m_0^{0,n} > m_0,$$

i.e., that the CEU household consumes more than the RE household in period 1 we need that

$$\begin{aligned} & \frac{1}{1 + \frac{1}{b(\nu_{0,1}^0)^{-\frac{1}{\theta}} m_1^{0,n}}} > \frac{1}{1 + \frac{1}{b(\psi_{0,1}^0)^{-\frac{1}{\theta}} m_1}} \\ \Leftrightarrow \quad & (\nu_{0,1}^0)^{-\frac{1}{\theta}} m_1^{0,n} > (\psi_{0,1}^0)^{-\frac{1}{\theta}} m_1 \\ \Leftrightarrow \quad & \frac{\nu_{0,1}^0}{\psi_{0,1}} < \left(\frac{m_1^{0,n}}{m_1} \right)^{\theta} \end{aligned}$$

which is *suficient pessimism*, cf. Definition 3.

B.8 Proof of Proposition 4

Sophisticated HD agent:

1. In period 2 we have $m_2^s = 1$, thus $c_2 = w_2$

2. In period 1 the sophisticated agent's first-order condition is the same than for the naive agent and is given by

$$c_1^{-\theta} = \beta \gamma \psi_1 R c_2^{-\theta}$$

Thus

$$c_1 = m_1^n w_1 = \frac{1}{1 + \frac{1}{\psi_1^{-\frac{1}{\theta}} b^n}} w_1$$

where

$$b^n \equiv \gamma^{-\frac{1}{\theta}} \left(\beta R^{1-\theta}\right)^{-\frac{1}{\theta}} = \gamma^{-\frac{1}{\theta}} b > b$$

because $\gamma \in [0,1]$. Consequently $m_1^n > m_1$.

3. In period 0 we have as a maximization problem for the sophisticated HD agent:

$$\begin{aligned}
L_0 &= u(c_0) + \beta\gamma\psi_0 \left[u(c_1(w_1)) + \frac{1}{\gamma}\beta\psi_1 u(c_2)\right] + \lambda_0 \left[w_1 - (w_0 - c_0) R\right] + \lambda_1 \left[w_2 - (w_1 - c_1) R\right] + \lambda_2 [\\
\frac{\partial L}{\partial c_0} &= \frac{du}{dc_0} + \lambda_0 R = 0 \\
\frac{\partial L}{\partial c_1} &= \beta\gamma\psi_0 \frac{du}{dc_1} + \lambda_1 R = 0 \\
\frac{\partial L}{\partial c_2} &= \beta^2\psi_0\psi_1 \frac{du}{dc_2} - \lambda_2 = 0 \\
\frac{\partial L}{\partial w_1} &= \beta\gamma\psi_0 \frac{du}{dc_1}\frac{dc_1}{dw_1} + \lambda_0 + \lambda_1 \left(1 - \frac{dc_1}{dw_1}\right) R = 0 \\
\frac{\partial L}{\partial w_2} &= \beta^2\psi_0\psi_1 \frac{du}{dc_2}\frac{dc_2}{dw_2} + \lambda_1 + \lambda_2 \left(1 - \frac{dc_2}{dw_2}\right) R = \beta^2\psi_0\psi_1 \frac{du}{dc_2} + \lambda_1
\end{aligned}$$

Thus

$$\begin{aligned}
\beta\gamma\psi_0 \frac{du}{dc_1}\frac{dc_1}{dw_1} + \lambda_0 + \lambda_1 \left(1 - \frac{dc_1}{dw_1}\right) R &= 0 \\
\beta\gamma\psi_0 \frac{du}{dc_1}\frac{dc_1}{dw_1} - \frac{du}{dc_0}\frac{1}{R} - \beta\gamma\psi_0 \frac{du}{dc_1}\frac{1}{R}\left(1 - \frac{dc_1}{dw_1}\right) R &= 0 \\
\beta R\gamma\psi_0 \left(\frac{dc_1}{dw_1} - \left(1 - \frac{dc_1}{dw_1}\right)\right) \frac{du}{dc_1} &= \frac{du}{dc_0}
\end{aligned}$$

Assuming a linear policy function $c_1 = m_1^n w_1$ and using $w_2 = (w_1 - c_1)R$ we get $\frac{dc_1}{dw_1} = m_1^n$ and thus:

$$c_0^{-\theta} = \beta\gamma\psi_0 R \left(m_1^n + \frac{1}{\gamma}(1 - m_1^n)\right) c_1^{-\theta}$$

From the FOC we get the policy function which is indeed linear:

$$
\begin{aligned}
c_0^{-\theta} &= \beta\gamma\psi_0 R\left(m_1^n + \frac{1}{\gamma}(1-m_1^n)\right) c_1^{-\theta} \\
c_0^{-\theta} &= \beta\gamma\psi_0 R\left(m_1^n + \frac{1}{\gamma}(1-m_1^n)\right) (m_1^{1,n}(w_0-c_0)R)^{-\theta} \\
c_0^{-\theta} &= \beta\gamma\psi_0 R^{1-\theta}\left(m_1^n + \frac{1}{\gamma}(1-m_1^n)\right) (m_1^{1,n}(w_0-c_0))^{-\theta} \\
c_0 &= \left(\beta R^{1-\theta}\right)^{-\frac{1}{\theta}} (\psi_0)^{-\frac{1}{\theta}} \left(\gamma m_1^n + (1-m_1^n)\right)^{-\frac{1}{\theta}} (w_0-c_0) m_1^{1,n} \\
c_0 &= b\cdot(\gamma\psi_0)^{-\frac{1}{\theta}} \underbrace{\left(m_1^n + \frac{1}{\gamma}(1-m_1^n)\right)}_{=d\geq 1 \text{ because } \gamma\in[0,1]}^{-\frac{1}{\theta}} (w_0-c_0) m_1^{1,n}
\end{aligned}
$$

$$
\Leftrightarrow \qquad c_0 = m_0^s w_0 \quad \text{where} \quad m_0^s = \frac{1}{1+\frac{1}{b\left[d\left(m_1^n\right)\cdot\gamma\psi_0\right]^{-\frac{1}{\theta}} m_1^{1,n}}}
$$

B.9 Proof of Proposition 5

FOC of the sophisticated agent:

$$
\begin{aligned}
u_c(c_0) &= \beta R\nu_{0,1}^0\left(m_1^{1,n} + \frac{\nu_{0,2}^0}{\nu_{0,1}^0\nu_{1,2}^2}(1-m_1^{1,n})\right) u_c(c_1) \\
\Leftrightarrow \qquad c_0 &= b(\nu_{0,1}^0)^{-\frac{1}{\theta}} \underbrace{\left(m_1^{1,n} + \frac{\nu_{0,2}^0}{\nu_{0,1}^0\nu_{1,2}^1}(1-m_1^{1,n})\right)}_{\equiv d>1 \text{ for moderate optimism (cf. Definition 2)}}^{-\frac{1}{\theta}} (w_0-c_0) m_1^{1,n} \\
\Leftrightarrow \qquad c_0 &= \frac{1}{1+\frac{1}{b(d\cdot\nu_{0,1}^0)^{-\frac{1}{\theta}} m_1^{1,n}}} w_0
\end{aligned}
$$

B.10 Proof of Observation 10

For the HD agent we get that

$$
\begin{aligned}
m_0^{sh} &< m_0^{0,nh} \\
\frac{1}{1+\frac{1}{b\left[d\left(m_1^{1,nh}\right)\cdot\gamma\psi_0\right]^{-\frac{1}{\theta}}m_1^{1,nh}}} &< \frac{1}{1+\frac{1}{b^n\psi_0^{-\frac{1}{\theta}}m_1}} \\
b\left[d\left(m_1^{1,nh}\right)\cdot\gamma\psi_0\right]^{-\frac{1}{\theta}}m_1^{1,nh} &< b^n\psi_0^{-\frac{1}{\theta}}m_1 \\
b\left[d\left(m_1^{1,nh}\right)\cdot\gamma\psi_0\right]^{-\frac{1}{\theta}}m_1^{1,nh} &< \gamma^{-\frac{1}{\theta}}b\psi_0^{-\frac{1}{\theta}}m_1 \\
\left[d\left(m_1^{1,nh}\right)\right]^{-\frac{1}{\theta}}\frac{m_1^{1,nh}}{m_1} &< 1 \\
\left(\frac{m_1^{1,nh}}{m_1}\right)^{\theta} &< d\left(m_1^{1,nh}\right)
\end{aligned}
$$

We know that $m_1^{1,nh} > m_1^{0,nh} = m_1$, cf. 8. For the CEU agent we have

$$
\begin{aligned}
m_0^{0,s} &< m_0^{0,n} \\
\frac{1}{1+\frac{1}{b\left[d\left(m_1^{1,n}\right)\cdot\nu_{0,1}^0\right]^{-\frac{1}{\theta}}m_1^{1,n}}} &< \frac{1}{1+\frac{1}{b(\nu_{0,1}^0)^{-\frac{1}{\theta}}m_1^{0,n}}} \\
\frac{1}{b(\nu_{0,1}^0)^{-\frac{1}{\theta}}m_1^{0,n}} &< \frac{1}{b\left[d\left(m_1^{1,n}\right)\cdot\nu_{0,1}^0\right]^{-\frac{1}{\theta}}m_1^{1,n}} \\
\left(\frac{m_1^{1,n}}{m_1^{0,n}}\right)^{\theta} &< d\left(m_1^{1,n}\right)
\end{aligned}
$$