House of Finance
Working Paper Series

Max Groneck - Alexander Ludwig - Alexander Zimper

A Life-Cycle Model with Ambiguous Survival Beliefs

SAFE Working Paper No. 73

SAFE | Sustainable Architecture for Finance in Europe
A cooperation of the Center for Financial Studies and Goethe University Frankfurt

House of Finance | Goethe University
Grüneburgplatz 1 | 60323 Frankfurt am Main

Tel. +49 69 798 33684 | Fax +49 69 798 33910
info@safe-frankfurt.de | www.safe-frankfurt.de

Non-Technical Summary

Optimal tax and transfer systems are key for the design of modern economics. One of the workhorse models used by economists to evaluate the welfare benefits of (reforms to) these systems is the life-cycle model of consumption and savings. The model can, e.g., be used to investigate the reactions of households to savings subsidies or any other kind of reform to old-age insurance, or, more general, any institutional feature of the tax-transfer system.

Yet, from a quantitative perspective, it is well known that the standard model produces several "puzzles" in a sense that the standard model cannot match certain facts in the data. It is well established that, relative to an optimal saving rate according to the model, households save too little in the data. Furthermore, the decumulation speed of assets in old-age is much lower in the data than predicted by the standard model. Finally, households behave dynamically inconsistent, in a sense that they generally save less during working life for retirement than they originally planned. Such inconsistencies cannot be accommodated by the standard model.

In order to generate correct quantitative predictions it is therefore important to modify the standard model in order to account for these three empirical regularities. This is the aim of the present paper.

The specific model element under investigation is the life-expectancy of households which is one of the most important ingredients of the model. Obviously, survival beliefs are of high relevance for savings behavior. The standard model uses objective data on survival beliefs, traced out from population wide survival tables. However, in several datasets that explicitly ask for subjective survival beliefs, substantial biases in survival beliefs relative to such objective data can be observed. E.g., young people strongly underestimate whereas old people (after retirement) strongly overestimate their chances to survive into the future.

This paper addresses the question how these biases in survival beliefs may alter model savings behavior, thereby bringing model predictions closer to the data on household savings. On the one hand, underestimation of survival beliefs may lead to lower savings than in the standard model. On the other hand, overestimation in old-age may lead to the fact that households hold on to their assets longer in life than predicted by the standard model.

To test whether the observed biases in survival beliefs have quantitatively important implications for the household model, we proceed in two steps. First, we develop a model of survival belief formation. We base our model on a decision theoretic framework which enables us to be explicit about psychological effects such as optimism, pessimism and doubt. A parsimonious representation of survival beliefs enables us to match pessimism with regard to survival for young and optimism for old households, as in the data on subjective survival beliefs. Furthermore, households continuously update their survival beliefs as they age in light of objective information towards which they express doubt. Second, we combine this model with an otherwise standard household life-cycle consumption-savings model. We analytically derive conditions under which this combination gives rise to less savings and a lower speed of asset accumulation, just as observed in the data. Furthermore, we show that the continuous updating of survival beliefs combined with our notion of doubt leads to dynamically inconsistent household behavior. We also show that a calibrated version of our model indeed matches well the key trends of asset accumulation and decumulation as observed in the data.

Therefore, our model kills three birds (=undersaving, too little asset decumulation and dynamic inconsistency) with one stone (=a model of survival beliefs with psychological biases).

Our results are of high relevance for future research because they show that biases in survival beliefs have strong implications for household savings. An immediate policy implication is that households must be provided with appropriate information about their survival prospects. Specifically, young households tend to underestimate the improvements to survival by medical progress. This may lead to too little old-age provision.

A Life-Cycle Model with Ambiguous Survival Beliefs*

Max Groneck† Alexander Ludwig‡ Alexander Zimper§

October 20, 2014

Abstract

On average, "young" people underestimate whereas "old" people overestimate their chances to survive into the future. We adopt a Bayesian learning model of ambiguous survival beliefs which replicates these patterns. The model is embedded within a non-expected utility model of life-cycle consumption and saving. Our analysis shows that agents with ambiguous survival beliefs (i) save less than originally planned, (ii) exhibit undersaving at younger ages, and (iii) hold larger amounts of assets in old age than their rational expectations counterparts who correctly assess their survival probabilities. Our ambiguity-driven model therefore simultaneously accounts for three important empirical findings on household saving behavior.

JEL Classification: D91, D83, E21.

Keywords: Cumulative prospect theory; Choquet expected utility; Dynamic inconsistency; Life-cycle hypothesis; Saving puzzles

*Financial support from the Network for Studies on Pensions, Aging and Retirement (Netspar), the Research Center SAFE, funded by the State of Hessen initiative for research LOEWE, the State North Rhine-Westphalia and Economic Research Southern Africa (ERSA) is gratefully acknowledged. We thank seminar participants at various places and Hippolyte d'Albis, Ben Heijdra, Christian Hellwig, Marciano Siniscalchi as well as two anonymous referees for their helpful comments. We also thank Annegret Hoch for excellent research assistance.

†CMR, University of Cologne; Netspar; Albertus-Magnus-Platz; 50923 Köln; Germany; E-mail: groneck@wiso.uni-koeln.de

‡SAFE, Goethe University Frankfurt; MEA; Netspar; House of Finance; Grüneburgplatz 1; 60323 Frankfurt am Main; Germany; E-mail: ludwig@safe.uni-frankfurt.de

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

1 Introduction

Expectations about future survival are important for numerous economic decisions. Forming such expectations is a very difficult task. In fact, substantial biases of subjective survival beliefs relative to objective data have been documented in the literature. This paper investigates the implications of such biases for life-cycle saving behavior.

Our work builds on two empirical regularities. First, according to the Health and Retirement Study (HRS), 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.[1] For example, women of age 65 on average report a 53 percent chance to survive to age 80 while their corresponding estimated cohort survival rate is 75 percent. In contrast, men of age 85 on average overestimate their probability to become 100 by 19 percentage points.

Second, recent empirical findings on household saving behavior proved to be puzzling for the standard "workhorse"-life-cycle model à la Modigliani and Brumberg (1954) and Ando and Modigliani (1963). For example, Laibson et al. (1998) and Bernheim and Rangel (2007) report large gaps between self-reported behavior and self-reported plans. People save less for retirement than actually planned (Choi et al. 2006; Barsky et al. 1997; Lusardi and Mitchell 2011). They behave in a dynamically inconsistent manner. Another well-known puzzle is that people hold large amounts of assets still late in life and dissave less in old age than predicted by the standard model (see, e.g., De Nardi et al. 2010; Hurd and Rohwedder 2010; Lockwood 2013).

We ask whether the observed biases in survival rates add to explanations for these empirical findings on saving behavior. Our approach to this research question comprises of two buildings blocks: We merge a model of learning of subjective survival beliefs with an otherwise standard life-cycle consumption-saving model.

As to the first building block, our point of departure is that the observed biases in survival beliefs cannot be reconciled with Muth (1961)'s rational expectations paradigm. According to this paradigm an agent would always hold correct objective beliefs. These biases are also inconsistent with standard models of Bayesian learning. According to such models subjective additive beliefs converge to objective probabilities when people gain more experience, i.e., grow older. We therefore follow Ludwig and Zimper (2013) who argue that psychological factors such as ambiguity, relative pessimism when young and relative optimism when old are drivers of survival belief formation.

Specifically, we adopt a simplified version of the Ludwig and Zimper (2013) Bayesian learning model of ambiguous survival beliefs. The model is based on Choquet ex-

[1] Similar patterns have been documented in numerous other datasets, see Section 2 for a review.

pected utility (CEU) theory (Schmeidler 1989; Gilboa 1987) or, equivalently, cumulative prospect theory (CPT) (Tversky and Kahneman 1992; Wakker and Tversky 1993).[2] Ludwig and Zimper (2013) characterize ambiguous survival beliefs as conditional neo-additive capacities in the sense of Chateauneuf et al. (2007) giving rise to a very parsimonious parametrization. This model covers a whole family of not necessarily additive probability measures. It nests the rational expectations paradigm as a special case.[3] Out of this large family of neo-additive capacities, our subsequent quantitative analysis employs the one capacity that gives the best fit to the data on survival beliefs reported in the HRS. Furthermore, Bayesian updating of beliefs results in age-specific conditional survival beliefs for all future ages. In contrast, for a static scenario—in which agents do not learn over the life-cycle—our notion of ambiguous survival beliefs reduces to an age-independent probability weighting function. While the special case of static non-additive survival beliefs has already been studied in the literature (Bleichrodt and Eeckhoudt 2006; Halevy 2008), the more general concept of Bayesian learning of ambiguous survival beliefs over the life-cycle adds to the plausibility of our approach.

To investigate whether biases in survival beliefs can partially resolve saving puzzles we combine the model of Bayesian learning of ambiguous survival beliefs with an otherwise standard canonical stochastic life-cycle consumption and saving model developed as our second building block. We compare consumption and saving behavior of CEU agents with the nested special case of rational expectations (RE) agents. Whenever CEU agents do not reduce to RE agents, life-cycle maximization gives rise to dynamically inconsistent behavior. We study both 'naive' and 'sophisticated' CEU agents. While the former do not anticipate that their future-selves deviate from ex ante optimal consumption plans, the latter are fully aware of their dynamically inconsistent behavior.

Qualitative analysis for a simplified three-period model, presented in a Supplementary Appendix[4], shows that naive as well as sophisticated CEU agents exhibit undersaving relative to their RE counterparts if they sufficiently underestimate objective survival probabilities at young ages. Furthermore, naive CEU agents have to only moderately overestimate their survival chances in order to save less in the intermediate model period than originally planned (otherwise they would save more). At the same time, naive CEU agents save more out of cash on hand in the intermediate model period than the

[2] Restricted to gains (as in our model), CPT is identical to CEU theory.

[3] The other extreme special case corresponds to the *Hurwicz criterion* according to which the agent only cares about best and worst possible outcomes whereas she is insensitive with respect to the likelihood of any non-extreme outcome. In particular, a model of beliefs is nested that would describe a *likelihood insensitive* agent who attaches a "50% chance" to the occurrence of any uncertain event.

[4] The Supplementary Appendix is available online at http://www.wiwi.uni-frankfurt.de/fileadmin/user_upload/dateien_abteilungen/abt_ewf/LS_Ludwig/SubjBeliefs_SuppApp.pdf.

corresponding RE agent. However, whether asset holdings in the final period are higher for the CEU agent depends on the interplay between underestimation at young ages and overestimation at older ages. Whether these conditions hold and how relevant the biases in beliefs are for generating saving puzzles are quantitative questions.

To address these questions we calibrate the stochastic quantitative life-cycle model to the data. The calibrated RE model gives rise to the aforementioned savings puzzles: The average saving rate for prime age savers of age $25-54$ is at 13.5%, compared to 9.5% in the data. Average asset holdings at ages 75, 85 and 95 relative to asset holdings at the average retirement age of 62 are 70.0%, 37.0% and 9.1%, compared to 72.4%, 53.0% and 47.9% in the data. Hence, through the lens of the RE model, the data are puzzling: the young save too little and the old decumulate assets too fast. The calibrated naive CEU agents model partially resolves these puzzles. The average saving rate is at 9.2% and relative asset holdings at ages 75, 85 and 95 at 77.9%, 57.1% and 34.9%. These statistics are remarkably close to the data. In addition, the realized saving rate is 5.5 percentage points lower than the planned saving rate. Predictions on asset holdings for the sophisticated agent CEU model are similar. They save a bit more than naive CEU agents and hence feature slightly higher asset holdings in old age. Overall, the fit to the data is better for naive than for sophisticated agents. Our analysis therefore suggests that our notion of ambiguous survival beliefs combined with naivety provides an accurate quantitative picture of saving behavior until about age 85.

The intuition for these quantitative findings is as follows: The calibrated model gives rise to sufficient underestimation at young age so that naive CEU agents save less than their RE counterparts. At the same time, naive CEU households only moderately overestimate their future survival chances so that they end up saving less in each period than originally planned in the past. As agents get older, overestimation of future survival beliefs eventually dominates so that the speed of asset decumulation is lower to the effect that the level of old age asset holdings is eventually higher than for RE agents. Finally, sophisticated agents correctly anticipate the more optimistic beliefs of their future selves. For reasons of consumption smoothing they therefore save more which leads them to have higher old-age asset holdings than their naive counterparts.

The standard model to explain dynamic inconsistency and undersaving is the hyperbolic time-discounting model. Building on the early work by Strotz (1955) and Pollak (1968), Laibson et al. (1998) find that exponential consumers save more than hyperbolic consumers, cf. also Angeletos et al. (2001). This standard model cannot account for high old-age asset holdings because long-run discounting is as in the rational expectations model. In contrast, optimistic beliefs in our CEU model imply lower long-run effective discount rates which leads to higher old-age asset holdings. In this respect our work relates to Halevy (2008) as well as Epper et al. (2011) who argue that hyperbolic time discounting is actually generated by ambiguous survival beliefs. Motivated by this insight, we show in our companion paper (Groneck et al. 2014) that quasi-hyperbolic time-discounting over the life-cycle is formally equivalent to a static CEU life-cycle model in which agents hold purely pessimistic neo-additive survival beliefs that are not subject to Bayesian learning.[5] The CEU model of the present paper is thus formally different from any quasi-hyperbolic time-discounting model because, first, it considers neo-additive survival beliefs that can express both, pessimistic as well as optimistic, ambiguity attitudes (cf., the inverse S-shaped probability weighting function of CPT) and, second, it allows for Bayesian learning over the life-cycle.

Similarly, standard explanations for insufficient old-age asset decumulation such as a bequest motive (Hurd 1989; Lockwood 2013) and precautionary savings behavior (Palumbo 1999; De Nardi et al. 2010) cannot generate undersaving at young ages. Our model of ambiguous survival beliefs therefore adds to existing explanations for saving behavior by simultaneously generating all three stylized findings: (i) time inconsistency, (ii) undersaving at young and (iii) high asset holdings at old age.

The remainder of our paper is organized as follows. Section 2 motivates and presents the parsimonious model of ambiguous survival beliefs. In Section 3 we combine it with a multi-period stochastic life-cycle model. Calibration is outlined in Section 4 and results of the quantitative analysis are presented in Section 5. Finally, Section 6 concludes. Appendix A recalls formal definitions from Choquet decision theory. Appendix B sketches the construction of ambiguous survival beliefs through Bayesian learning under ambiguity. Appendix C lays out the proof to the Propositions. Appendix D describes the construction of the asset data used for calibration.

2 Ambiguous Survival Beliefs

2.1 Biases in Survival Beliefs

As point of departure we consider subjective survival beliefs elicited in the Health and Retirement study (HRS). Respondents are asked about their assessment of the probability to survive from some interview age up to a specific target age. Target age is mostly 10 to 15 years in advance, see Table 1.

[5] On the one hand, the static model in Groneck et al. (2014) is more restrictive than the approach in Halevy (2008) because it only considers neo-additive but not general non-additive probability measures. On the other hand, Groneck et al. (2014) is more general than Halevy (2008) in that we consider arbitrary and not only constant survival rates over the life-cycle.

Table 1: Interview and Target Age

Age at Interview	Target Age
≤ 69	80
$70 - 74$	85
$75 - 79$	90
$80 - 84$	95
$85 - 89$	100

Source: RAND HRS Data Documentation.

Figure 1 shows aggregated data from the HRS by plotting average age-specific biases in survival beliefs—the difference between the respective average subjective belief and the average objective data—for three waves of the HRS between 2000 and 2004.[6] We observe that relatively "young"—younger than age 70—respondents underestimate whereas relatively "old"—above age 80—respondents overestimate their chances to survive into the future. Younger respondents between ages 50 and 70 underestimate their survival chances by about 10 to 20 percentage points on average. Older respondents around the age of 85 overestimate their survival chances by 15-20 percentage points. Also notice that the overestimation is getting more pronounced with increasing age.

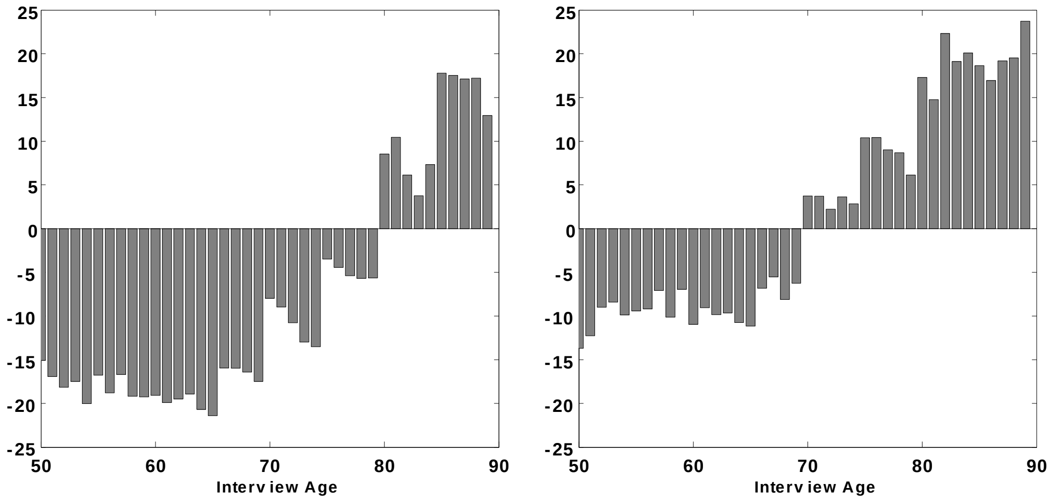
This age-specific pattern of subjective survival beliefs is a well-established stylized fact that has been confirmed by various other studies using different data sets. Hammermesh (1985) documents that subjective survival rate functions are generally flatter than their objective counterpart implying underestimation at younger ages and overestimation at older ages. Similar findings have been described by Elder (2013) for the US and by Peracchi and Perotti (2010) for European countries using the Survey of Health, Ageing and Retirement in Europe (SHARE). Wu et al. (2013) highlight a related fact in the Australian "Retirement Plans and Retirement Incomes: Pilot Survey": respondents underestimate survival probabilities in the near future whereas they overestimate survival rates for the distant future.

The biases of subjective survival perceptions from objective life-table data shown in Figure 1 can be expected to have significant implications for household's consumption and saving decisions. A number of recent studies confirms this. For example, Salm (2010) estimates that a 1 percent increase in the subjective probability of mortality reduces annual future consumption of non-durable goods by around 1.8 percent. Bloom

[6] Objective data are based on cohort life-tables so that future trends in life expectancy are appropriately taken into account.

Figure 1: Difference of Subjective Survival Probabilities and Cohort Data

(a) Women (b) Men



Notes: 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-Carter procedure (Lee and Carter 1992). Each bar depicts the difference of unconditional probabilities to survive to a specific target age, cf. Table 1.

Source: Own calculations based on HRS, Human Mortality Database and Social Security Administration data.

et al. (2006) find that an increased subjective survival probability leads to higher wealth accumulation thereby confirming results of Hurd et al. (1998).

Our contribution is to study the implications of these biases in a structural life-cycle model of consumption and saving. Incorporating subjective survival beliefs in such a model requires knowledge of the entire distribution of survival beliefs while, in general, there are only few data points available. In the HRS, for example, only a specific average subjective survival probability for each interview age is observed, cf. Table 1. To overcome this, a number of recent studies therefore estimate subjective survival belief functions by assuming specific hazard functions, cf., e.g., Gan et al. (2005), Bissonnette et al. (2011), Khwaja et al. (2007) and Wu et al. (2013). We add to this literature by adopting a decision theoretic model in order to interpret as well as to inter- and extrapolate the data. The next subsection describes this approach.

2.2 Bayesian Learning of Ambiguous Survival Beliefs

Life-cycle models with rational expectations use objective survival probabilities, denoted $\psi_{k,t}$, to model the representative agent's beliefs to survive from age k to t, with $k < t$.

These beliefs are independent of the agent's age because there is no learning of survival beliefs over the life-cycle: the rational expectations agent always already knows her true survival chances. Figure 1 demonstrates that real people do not know their true survival chances. In absence of such knowledge, it is plausible that some learning of survival beliefs happens over the representative agent's life-cycle. The agent receives more and more information about her survival chances as she grows older.

Instead of a rational expectations agent therefore consider, for the moment, an agent whose belief to survive from age k to t is given by the subjective additive probability $\varphi^h_{k,t} \neq \psi_{k,t}$. This belief is formed conditional on the information that the agent has collected up to her current age $h \leq k$. If this agent is a standard Bayesian decision maker, there is a straightforward way to model her learning behavior. Namely, in the absence of any information (i.e., at $h = 0$) the agent would hold a unique additive prior over an index set of distributional parameters. This corresponds to possible probabilities to survive from age k to t so that the (unconditional) subjective survival belief $\varphi_{k,t} \equiv \varphi^0_{k,t}$ is given as her Bayesian estimate. That is, it is the expected value of all possible survival probabilities with respect to this prior. In light of new information received at age $h > 0$, the agent would then update her prior to a posterior by Bayes' rule. Her conditional subjective survival belief $\varphi^h_{k,t}$ is now the expected value of all possible survival probabilities with respect to this posterior. Consistency results for standard Bayesian estimates establish convergence of the subjective survival belief $\varphi^h_{k,t}$ in the age h to the objective probability $\psi_{k,t}$. Convergence is due to the fact that posteriors will (almost surely) concentrate at the agent's true survival probability for a sufficiently large data sample.[7] However, the age-specific pattern of the biases in Figure 1 suggests that such converging learning behavior does, in reality, not happen over the life cycle; instead of convergence to objective survival probabilities, strong underestimation of objective survival probabilities is followed by strong overestimation as the representative agent grows older. Under the assumption that the agents are Bayesian learners who receive an increasing amount of statistical information over their life-cycle, the reported survival beliefs of Figure 1 can thus not be plausibly modeled as subjective additive probabilities.

The situation is different for Bayesian learning under ambiguity. Ambiguous survival beliefs do not necessarily converge to objective survival probabilities if the agent receives more information. In a first step, we model ambiguous survival beliefs as *neo-additive capacities* (Chateauneuf et al. 2007). Neo-additive capacities are an analytically very

[7] Convergence to the true parameter value only occurs if the prior is well-specified, i.e., has this true value in its support; (the seminal contribution is Doob 1949). For a more general convergence result–including misspecified priors–in terms of minimization of the Kullback-Leibler divergence, see Berk (1966).

tractable class of non-additive probability measures which are used in the literature[8] to approximate inverse S-shaped probability weighting functions as typically elicited for CPT (cf., e.g., Tversky and Kahneman 1992; Wu and Gonzalez 1996; 1999). For the given (unconditional) subjective survival probability $\varphi_{k,t} \in (0,1)$, the corresponding (unconditional) neo-additive probability to survive from k to t is formally defined as

$$\nu_{k,t} = \delta \cdot \lambda + (1-\delta) \cdot \varphi_{k,t}. \tag{1}$$

The parameter $\delta \in [0,1]$ is interpreted as a degree of ambiguity. If there is no ambiguity ($\delta = 0$), the (unconditional) ambiguous survival belief (1) reduces to the agent's subjective survival probability. If there is ambiguity ($\delta > 0$), the optimism parameter $\lambda \in [0,1]$ measures in how far the agent resolves this ambiguity through over- (high values of λ) versus under-estimation (low values of λ) of her subjective survival probability.

In a next step, we allow for the possibility that some notion of Bayesian learning happens over the agent's life-cycle. To this purpose we adopt a simplified version of the Choquet Bayesian learning model of Ludwig and Zimper (2013). These authors describe a conditional ambiguous survival belief, denoted $\nu^h_{k,t}$, as the Choquet expected value of all possible survival probabilities with respect to a neo-additive posterior. The posterior results from updating in accordance with the *Generalized Bayesian update rule* (Pires 2002; Eichberger et al. 2007) with respect to the survival chance information that the agent has obtained up to age h. Relative to Ludwig and Zimper (2013) our simplification here is that we ignore the initial bias in prior beliefs so that any bias between objective survival probabilities and subjective beliefs is exclusively ambiguity-driven. Hence, we set $\varphi_{k,t} = \psi_{k,t}$ for all k, t. In Appendix A we present a mathematically rigorous review of Choquet decision theory with neo-additive capacities. We derive our simplified version of the Ludwig and Zimper (2013) learning model in Appendix B where we also prove the following proposition:

Proposition 1. *Denote the objective probability to survive from k to t by $\psi_{k,t} \in (0,1)$ and fix age-independent parameters $\delta, \lambda \in [0,1]$. The h-old agent's age-dependent ambiguous belief to survive from age k to target age t is given by*

$$\nu^h_{k,t} = \delta_h \cdot \lambda + (1-\delta_h) \cdot \psi_{k,t} \tag{2}$$

where

$$\delta_h = \frac{\delta}{\delta + (1-\delta) \cdot \frac{1}{1+e(h)}} \tag{3}$$

for some non-decreasing experience function $e(h)$.

[8] See, e.g., Wakker (2010), Abdellaoui et al. (2011), and Ludwig and Zimper (2013).

To interpret our concept of age-conditional ambiguous survival beliefs (2) let us consider three different scenarios: First, in absence of any initial ambiguity, i.e., $\delta = 0$, we have for all h that $\nu_{k,t}^{h} = \psi_{k,t}$ so that all ambiguous survival beliefs (2) reduce to objective survival probabilities. Consequently, the standard rational expectations approach is nested as a (degenerate) special case within our notion of age-dependent ambiguous survival beliefs.

Second, if there is initial ambiguity, i.e., $\delta > 0$, but no learning over the agent's life-cycle, i.e., $e(h) = 0$ for all h, we speak of the "static" scenario. In this case, (2) reduces to the age-independent ambiguous survival belief

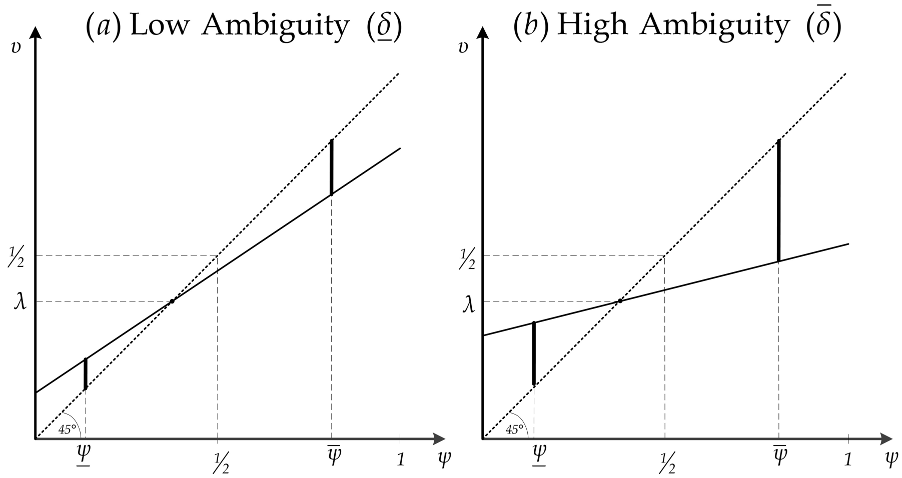
$$\nu_{k,t}^{h} = \delta \cdot \lambda + (1 - \delta) \cdot \psi_{k,t} \text{ for all } h, \tag{4}$$

which can be interpreted as the transformation of the objective survival probability by a neo-additive probability weighting function. Note that Bleichrodt and Eeckhoudt (2006) as well as Halevy (2008) already consider non-additive survival beliefs such that some age-independent probability weighting function is applied to an additive survival probability as in (4). Since the static scenario is nested within our general notion (2) as a special case, it is straightforward to investigate the sensitivity of our results with regard to this feature of the model.

Because the agents of the static scenario never revise their survival beliefs over the life-cycle, the static model is equivalent to a Bayesian model in which agents do not receive any new information as they grow older. This is in stark contrast to everyday experience according to which real-life people increasingly receive news about the deaths (or critical illnesses) of acquainted people thereby providing them with new information about their own survival chances. The third and, in our opinion, most plausible scenario is therefore a combination of initial ambiguity with Bayesian learning over the life-cycle such that the agent's experience function $e(h)$ strictly increases in her age h, i.e., $e_h > 0$. We further assume decreasing marginal experience, i.e., $e_{hh} < 0$.

If there is initial ambiguity, i.e., $\delta > 0$, the age-dependent ambiguous survival beliefs (2) do not converge through Bayesian learning over the life-cycle to the objective survival probabilities. As a consequence, our notion of ambiguous survival beliefs can replicate the pattern of Figure 1. We illustrate the model's main properties by use of Figure 2. Interior objective survival rates $\psi_{k,t} \in (0, 1)$ are mapped into corresponding subjective survival rates ν^h by a linear transform which is (i) flatter than the 45-degree line and (ii) becomes flatter with increasing age h. Furthermore, the ν^h-line intersects with the 45-degree line at the relative optimism parameter λ. Non-decreasing ambiguity in the agent's age together with lower objective survival rates for older agents—lower than λ, which is consistent with our empirical estimates, cf. Subsection 4.3—implies an increasing importance of overestimation of survival probabilities.

Figure 2: Dynamics of the Subjective Survival Belief Model



Notes: Subjective survival belief ν compared to the objective counterpart ψ. Panel (a) shows the case for low ambiguity $\underline{\delta}$ (young agent). The deviation from the 45-degree line is only modest implying that both underestimation of the high objective probability $\bar{\psi}$ and overestimation of the low objective probability $\underline{\psi}$ are small. Panel (b) shows high ambiguity $\bar{\delta}$ (older agent). The subjective survival line is more horizontal implying that under- and overestimation is more pronounced.

Explanatory Discussion: Age-increasing Ambiguity

Whenever the experience function $e(h)$ strictly increases in the agent's age h, her ambiguity parameter δ_h, given by (3), increases. In other words, the agent's ambiguity increases in the amount of statistical information she receives. This feature might seem to be counter-intuitive because one possible interpretation of ambiguity is the lack of sufficient statistical information to form a unique additive belief. Although an in-depth discussion of the ongoing (and fascinating) research on Bayesian learning under ambiguity is beyond the scope of this paper, the following remarks briefly discuss this feature within the context of the existing literature.

First, within the Choquet Bayesian learning model of Ludwig and Zimper (2013), age-increasing ambiguity is not an ad hoc assumption. It is a formal implication which turns out to be remarkably robust with respect to alternative Bayesian update rules for Choquet decision makers (cf. Gilboa and Schmeidler 1993; Zimper 2011). One reason is that ambiguity in the Ludwig and Zimper (2013) model applies to the joint probability of the parameter and sample space so that the updating process itself—and not only the distribution over parameters—is subject to ambiguity. As a consequence, the agent does not perceive the data generating process as i.i.d. as in the standard (non-ambiguous) Bayesian set-up. Furthermore, a large amount of statistical information refers to an event which has an ex ante small likelihood to be ex post observed within this non-i.i.d.

environment. Bayesian updating under ambiguity "punishes" small ex ante likelihoods in the sense that the decision maker's ambiguity increases if she observes information which she considered ex ante as unlikely.

Second, while thus a rather mechanical consequence of the updating process, age-increasing δ_h captures the intuitive notion that, as the objective risk of survival becomes less likely, agents attach less and less weight to this objective probability. According to our estimates of δ and λ, presented in Section 4, objective survival probabilities $\psi_{k,t}$ decrease with age to values lower than λ, cf. $\underline{\psi}$ in Figure 2. The model's convergence property hence implies that survival rates are overestimated eventually even when the initial degree of ambiguity, δ, is low. Overestimation at old age may result from the fact that people have survived the gamble against death several times before. Consequently, one possible heuristic interpretation of age-increasing δ_h might be that people want to avoid a realistic assessment of their encounter with death.[9]

Third, the concept of *likelihood-insensitivity* (cf., Wakker 2004; 2010; Abdellaoui et al. 2011), may provide an alternative heuristic interpretation for the age-increasing δ_h of our model. These authors interpret δ_h not as an ambiguity but rather as a cognitive parameter which reflects the empirical observation that people do not sufficiently distinguish between non-degenerate probabilities. For instance, an extreme example for likelihood insensitivity are "fifty-fifty" probability assessments for any uncertain event and its complement. Under this cognitive interpretation, likelihood insensitivity—and not necessarily ambiguity—would increase with age. Given that old people increasingly suffer from cognitive impairments, this alternative interpretation has some intuitive appeal. Despite this, we continue to interpret δ_h as age-dependent ambiguity in the remainder of our analysis.

Fourth, Nicholls et al. (2014) investigate whether violations of Savage's (1954) sure-thing principle (STP), typically interpreted as the expression of ambiguity attitudes, decrease or increase if the subjects receive an increasing amount of statistical information. As their main finding, these authors conclude that "[...] statistical learning has, at best, no impact on STP violations. At worst, it might even be causing STP violations to increase." (p. 14). This empirical finding suggests that conventional wisdom about Bayesian learning might not be adequate for situations with ambiguity.

Finally, there exist alternative models of Bayesian learning under ambiguity such that ambiguity might decrease in the amount of statistical information.[10] Within a multiple-

[9]This interpretation is consistent with the observation of Kastenbaum (2000) who summarizes the insights of psychological research on the reflection about personal death as follows: "There are divergent theories and somewhat discordant findings, but general agreement that most of us prefer to minimize even our cognitive encounters with death."

[10]For a detailed discussion of these models we refer the interested reader to Zimper and Ma (2014).

priors set-up, Marinacci (2002) restricts ambiguity to the parameter space whereas Bayesian updating happens with respect to a standard (i.e., non-ambiguous) i.i.d. data-generating process. The convergence behavior of the Bayesian learning process in the Marinacci (2002) model crucially depends on the support of the priors held by the decision maker. If not all priors have the same support, ambiguity does not necessarily vanish when an unlimited amount of statistical information becomes available. Epstein and Schneider (2007) consider two dimensions of ambiguity. First (as in Marinacci 2002), ambiguity with respect to prior beliefs is expressed through multiple priors; second, ambiguity with respect to the updating process is expressed through *multiple likelihoods*. Furthermore, these authors impose a specific *expected maximum likelihood criterion* as a prior-selection rule. This may reject initially plausible priors in the light of new information. Ambiguity with respect to posterior beliefs vanishes in the original Epstein and Schneider (2007) model if, and only if, there is no ambiguity with respect to the updating process. However, even if there is no ambiguity with respect to the updating process, ambiguity might not vanish in a modified—and ad hoc equally plausible—version of the Epstein and Schneider (2007) model. This model would consider—instead of the expected maximum likelihood criterion—some alternative prior selection rule such as the *minimal Kullback-Leibler divergence criterion* (cf. Zimper and Ma 2014). The analysis in Marinacci (2002) and in Epstein and Schneider (2007) thus suggests that it requires quite strong ad hoc assumptions on the priors' support, on the updating process as well as on the prior-selection rule for ambiguity to vanish in alternative theoretical models of Bayesian learning under ambiguity.

3 Quantitative Life-Cycle Model

This section merges our notion of ambiguous survival beliefs with a life-cycle model. One model period corresponds to one age year. We model a realistic life-cycle income profile including stochastic and age-specific labor productivity. In addition, a PAYG pension system is modeled with a fixed date of retirement. We assume no annuity markets and a self-imposed borrowing constraint (because there is always a small positive probability of drawing zero income). These elements are included only in order to generate realistic endogenous life-cycle consumption profiles. (Self-imposed) borrowing constraints, stochastic labor income in combination with impatience give a hump-shaped consumption profile, as in the data. Positive pension income implies that savings for retirement are not too large.

3.1 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 $\psi_{k,t} \in (0, 1)$ for all $t \leq T$ and $\psi_{k,t} = 0$ for $t = T+1$. We think of survival risk as an idiosyncratic risk that washes out at the aggregate level. Total population is therefore constant and dynamics of the population are correspondingly given by $N_{t+1} = \psi_{t,t+1} N_t$, for N_0 given.

3.2 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 modelling stochastic labor productivity is to impose discipline on calibration. For sake of comparability, our fully rational model features standard elements as used in numerous structural empirical studies on life-cycle models, cf., e.g., Laibson et al. (1998), Gourinchas and Parker (2002) and references therein. 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. E^{t_r-1} are σ-algebras generated by $\mathsf{E}, \mathsf{E}, ...$ We assume that there is an objective probability space $\left(\times_{t=0}^{t_r-1} \mathsf{E}^{t_r-1}, \pi\right)$ 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 is zero (zero income).[11] 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 continuously differentiable policy functions. (Self-imposed) borrowing constraints are required to generate realistic paths of life-time consumption, saving and asset accumulation. Continuous differentiability is convenient when we model a sophisticated agent. By thereby avoiding technicalities as addressed in Harris and Laibson (2001) we keep our analysis

[11] This is standard practice within the life-cycle literature. It is also applied by, e.g., Gourinchas and Parker (2002).

focused. Since the zero income probability is small, results are virtually unaffected by this assumption, relative to a model with a fixed zero borrowing limit which would result in a kink in each policy function. In fact, we obtain almost identical numerical results for such a model. In addition, we assume productivity to vary by age. Accordingly, ϕ_t denotes age-specific productivity which is estimated from the data and results in a hump-shaped life-cycle earnings profile.

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. Pension contributions are levied at contribution rate, τ. To achieve a self-imposed borrowing constraint and continuous policy functions also during the retirement period, we assume that there is a small i.i.d. probability of default of the government on its pension obligations. Accordingly, $\eta_t \in E^r = [1, 0]$ during retirement. Correspondingly, let E^r be the powerset of the finite set E^r. $\mathsf{E}^{\mathsf{r}T-t_r+1}$ are σ-algebras generated by $\mathsf{E}^\mathsf{r}, \mathsf{E}^\mathsf{r}, ...$ and $\left(\times_{t=t_r}^{T} \mathsf{E}^\mathsf{r}, \pi^r\right)$ is the objective probability space in the retirement period.

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

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

We abstract from private annuity markets.[12] The interest rate, r, is assumed to be fixed. With cash-on-hand given as $x_t \equiv a_t \left(1 + r\right) + y_t$ the budget constraint writes as

$$x_{t+1} = \left(x_t - c_t\right)\left(1 + r\right) + y_{t+1}. \tag{5}$$

Finally, define total income as $y_t^{tot} \equiv y_t + r a_t$, and gross savings as assets tomorrow, a_{t+1}.

3.3 Government

We assume a pure PAYG public social security system. 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

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

[12]Hence, we do not address the annuity puzzle in this paper, i.e., the observed small size of private annuity markets, see Friedman and Warshawsky (1990) for an overview. On the one hand, underestimation of survival beliefs extenuates the annuity puzzle. On the other hand, overestimation at old age reinforces the puzzle. However, overestimation of survival rates only sets in after the age of 70 and the average underestimation in our total sample is around 27 percentage points.

In addition, accidental bequests—arising because of missing annuity markets—are taxed away at a confiscatory rate of 100%. Revenue from this source is used for government consumption which is otherwise neutral. Also, in the unlikely event of default of the government on its pension obligations, the government uses contributions to the pension system for otherwise neutral government consumption.

3.4 CEU Preferences

Households face two dimensions of uncertainty, respectively risk, about period t consumption. First, due to our assumption of productivity shocks, agents face a risky labor income. Second, agents are uncertain with respect to their life expectancy. 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 ambiguous survival beliefs as stated in Proposition 1.

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 to every history of shocks $\eta^t|\eta^h$ some non-negative amount of period t consumption. Denote by $u(c_t)$ the agent's strictly increasing utility from consumption at age t, i.e., $u'(c_t) > 0$. We normalize $u(0) = 0$. We assume that the agent is strictly risk-averse, i.e., $u''(c_t) < 0$. Expected utility of an h-old agent from consumption in period $t > h$ contingent on the observed history of productivity shocks η^h is then given as $\mathbb{E}_h[u(c_t)] \equiv \mathbb{E}\left[u(c_t), \pi\left(\eta^t|\eta^h\right)\right] = \sum_{\eta^t|\eta^h} u(c_t)\, \pi\left(\eta^t|\eta^h\right)$.

We assume additive time-separability and add a raw time discount factor $\beta = \frac{1}{1+\rho}$.[13] Fix some $s \in \{h, h+1, ..., T\}$ with the interpretation that the agent survives until period s and dies afterwards. Zero consumption in periods of death implies that $u(c_t) = 0$ for all $t > s$. Given s, the agent's von Neumann Morgenstern utility from a consumption plan $\mathbf{c}$ is then defined as

$$U(\mathbf{c}(s)) = u(c_h) + \sum_{t=h+1}^{s} \beta^{t-h} \mathbb{E}_h [u(c_t)]. \tag{7}$$

To model survival uncertainty of an agent of age h with respect to ambiguous survival beliefs, we use the sequence of conditional neo-additive probability spaces $(\Omega, \mathcal{F}, \nu(\cdot \mid h))$, $h = 1, ..., T$, which is mathematically rigorously constructed in Appendix B.3. Denote

[13] In line with Halevy (2008) and Andreoni and Sprenger (2012), we assume that time-preferences cannot be reduced to preferences under uncertainty. To keep the formalism as transparent as possible, we simply consider standard exponential time-discounting.

by $\nu^h \equiv \nu(\cdot \mid h)$ the agent's age-conditional neo-additive capacity and by $\psi_s = \psi(\cdot)$ the objective probability to survive until age s.

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 1 in Appendix A.1, this agent's CEU from consumption plan $\mathbf{c}$ with respect to ν^h is given as

$$\begin{aligned} \mathbb{E}\left[U(\mathbf{c}), \nu^h\right] &= \delta_h \left[\lambda \sup_{s\in\{h,h+1,\ldots\}} U(\mathbf{c}(s)) + (1-\lambda) \inf_{s\in\{h,h+1,\ldots\}} U(\mathbf{c}(s))\right] \\ &\quad + (1-\delta_h) \cdot \sum_{s=h}^{T} \left[U(\mathbf{c}(s)), \psi_s\right]. \end{aligned} \tag{8}$$

The Choquet expected value of lifetime utility $U(\mathbf{c})$ with respect to a neo-additive capacity ν^h is a convex combination of the expected value of U with respect to some additive probability measure ψ_s and an ambiguity part. In case there is some ambiguity, i.e., $\delta > 0$, parameter λ measures how much weight the decision maker puts on the least upper bound of the range of U. Conversely, $(1-\lambda)$ is the weight she puts on the greatest lower bound. For these bounds we have for any $\mathbf{c}$ that

$$\begin{aligned} \sup_{s\in\{h,h+1,\ldots\}} U(\mathbf{c}(s)) &= u(c_h) + \sum_{t=h+1}^{T} \beta^{t-h} \mathbb{E}_h\left[u(c_t)\right], \\ \inf_{s\in\{h,h+1,\ldots\}} U(\mathbf{c}(s)) &= u(c_h), \end{aligned} \tag{9}$$

i.e., the least upper bound consists of the discounted sum of utilities if survival probabilities were equal to one in every period. The greatest lower bound is the utility if the agent does not survive to the following period. The following technically convenient characterization of (8) is derived in the appendix.

Proposition 2. *Consider an agent of age h. The agent's Choquet expected utility from consumption plan $\mathbf{c}$ is given by*

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

where the subjective belief to survive from age h to $t \geq h$ is given by

$$\nu_{h,t}^h = \begin{cases} \delta_h \cdot \lambda + (1-\delta_h) \cdot \psi_{h,t} & \text{for } t > h \\ 1 & \text{for } t = h \end{cases} \tag{11}$$

with δ_h given by (3).

3.5 Recursive Problem and Dynamic Inconsistency

In contrast to a sequence of conditional additive probability spaces $(\Omega, \mathcal{F}, \psi(\cdot \mid h))$, $h = 1, ..., T$, the age-dependent sequence of conditional neo-additive probability spaces $(\Omega, \mathcal{F}, \nu(\cdot \mid h))$, $h = 1, ..., T$, (generically) violates dynamic consistency of the agents' life-cycle utility maximization problem whenever $\delta > 0$.[14] To characterize actual behavior in the presence of dynamic inconsistency, we analyze both naive and sophisticated agents, cf. Strotz (1955) or inter alia O'Donoghue and Rabin (1999) for procrastination models.

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. 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 the behavior of their future selves.

Although there exists some empirical evidence which suggests that naive rather than sophisticated decision making might be more relevant (cf. O'Donoghue and Rabin (1999) and the literature cited therein), there also exists evidence according to which several investment and contractual arrangements (e.g., investment in rather illiquid assets such as real estate financed by long-term loans) serve as commitment devices through which sophisticated agents restrain the consumption behavior of their future selves (cf., e.g,. Ludwig and Zimper (2006) and references therein). In the present paper, we take the pragmatic stand to consider both types of behavior.

We further assume that income risk is first-order Markov such that $\pi(\eta^t \mid \eta^{t-1}) = \pi(\eta^t \mid \eta_{t-1})$. It is then straightforward to set up the recursive formulation of lifetime utility (10). The value function of age $t \geq h$ viewed from planning age h is given by

$$V_t^h(x_t, \eta_t) = \max_{c_t, x_{t+1}} \left\{ u(c_t) + \beta \frac{\nu_{h,t+1}^h}{\nu_{h,t}^h} \mathbb{E}_t \left[V_{t+1}^h \left(x_{t+1}, \eta_{t+1} \right) \right] \right\}.$$

Maximization of the above is subject to (5).

[14] We refer the interested reader to the axiomatic treatment of the relationship between violations of dynamic consistency and violations of Savage's (1954) sure-thing principle (as in CEU theory) to Epstein and Le Breton (1993), Ghirardato (2002), Siniscalchi (2011) and the Appendix in Zimper (2012).

Naive Agents

The naive CEU agent's first order condition is given by the standard Euler equations.

Proposition 3. *The consumption plan* $\mathbf{c} = (c_h, c_{h+1}, ...)$ *of a naive CEU agent must satisfy, for all* $t \geq h$,

$$\frac{du}{dc_t} = \beta(1+r) \cdot \frac{\nu^h_{h,t+1}}{\nu^h_{h,t}} \cdot \mathbb{E}_t \left[\frac{du}{dc_{t+1}} \right], \tag{12}$$

where

$$\frac{\nu^h_{h,t+1}}{\nu^h_{h,t}} = \begin{cases} \nu^h_{h,h+1} = \delta_h \psi_{h,h+1} + (1-\delta_h)\lambda & \text{for } t = h \\ \frac{\delta_h \psi_{h,t+1} + (1-\delta_h)\lambda}{\delta_h \psi_{h,t} + (1-\delta_h)\lambda} & \text{for } t > h. \end{cases}$$

By (12), the expected growth of marginal utility from h to $h+1$ is higher than under rational expectations if the household underestimates the probability of survival to the next period, i.e., if $\nu^h_{h,h+1} < \psi_{h,h+1}$, and vice versa for overestimation. From (12) we can also directly verify that the CEU life-cycle maximization problem is dynamically inconsistent if and only if the ambiguous survival beliefs do not reduce to additive probabilities. To see this formally let us compare the optimal consumption choice of an $h+1$ old agent, first, from the perspective of an h old and, second, from her actual perspective when she turns $h+1$. By Proposition 3, the optimal consumption plan for age $h+1$ from the perspective of age h requires that

$$\frac{du}{dc_{h+1}} = \beta(1+r) \cdot \frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1}} \cdot \mathbb{E}_{h+1} \left[\frac{du}{dc_{h+2}} \right], \tag{13}$$

whereas the optimal consumption choice at age $h+1$ from the perspective of age $h+1$ requires that

$$\frac{du}{dc_{h+1}} = \beta(1+r) \cdot \frac{\nu^{h+1}_{h+1,h+2}}{\nu^{h+1}_{h+1,h+1}} \cdot \mathbb{E}_{h+1} \left[\frac{du}{dc_{h+2}} \right]. \tag{14}$$

Dynamic consistency with respect to the optimal consumption choice at age $h+1$ thus holds if and only if the two first order conditions (13) and (14) coincide. Because of $\nu^{h+1}_{h+1,h+1} = 1$, this is the case if and only if

$$\frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1}} = \nu^{h+1}_{h+1,h+2},$$

which holds for $\delta = 0$, implying

$$\frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1}} = \frac{\psi_{h,h+2}}{\psi_{h,h+1}} = \psi_{h+1,h+2} = \nu^{h+1}_{h+1,h+2},$$

but which is violated for $\delta > 0$ since (generically)

$$\frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1}} = \frac{\delta_h \lambda + (1-\delta_h)\,\psi_{h,h+2}}{\delta_h \lambda + (1-\delta_h)\,\psi_{h,h+1}} \neq \delta_{h+1}\lambda + (1-\delta_{h+1})\,\psi_{h+1,h+2} = \nu^{h+1}_{h+1,h+2}.$$

As in the static CPT model of Halevy (2008), the life-cycle maximization problem of naive CEU agents is thus dynamically inconsistent. While dynamic inconsistency in Halevy (2008) results from a fixed non-additive probability weighting function, dynamic inconsistency in our model comes with a sequence of non-additive probability weighting functions.

Sophisticated Agents

Sophisticated agents are fully aware of their dynamic inconsistency which is reflected in their first order conditions. These agents try to influence future selves' behavior via the choice of x_{t+1}. Hence, the usual Envelope conditions which are standard in rational expectations problems no longer apply. As a result, the marginal propensities to consume out of cash-on-hand (MPC), $m_{h+1} \equiv \frac{\partial c_{h+1}}{\partial x_{h+1}}$, show up explicitly in the first-order conditions. This reflects how self h influences future self's $h+1$ choices with her choice of savings, x_{h+1}.

Combining first order conditions of optimality for the CEU agent results in a "generalized Euler equation with adjustment factor".

Proposition 4. *The generalized Euler equation with adjustment factor for the sophisticated CEU agent is given by*

$$\frac{du}{dc_h} = \beta\,(1+r)\,\nu^h_{h,h+1}\cdot \mathbb{E}_h\left[\Theta_{h+1}\cdot\frac{du}{dc_{h+1}} + \Lambda_{h+1}\right] \tag{15}$$

where

$$\Theta_{h+1} \equiv m_{h+1} + \frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1}\cdot\nu^{h+1}_{h+1,h+2}}\,(1-m_{h+1}) \tag{16}$$

and

$$\Lambda_{h+1} \equiv \beta(1+r)\frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1}}\,(1-m_{h+1})\left(\frac{\partial V^h_{h+2}}{\partial x_{h+2}} - \frac{\partial V^{h+1}_{h+2}}{\partial x_{h+2}}\right). \tag{17}$$

Proof: See Appendix C.

Relative to the naive agent, the FOC of the sophisticated agent (15) hence features two additional terms, Θ_{h+1} and Λ_{h+1}. To interpret this condition, first assume that $\Lambda_{h+1} = 0$. Then (15)-(16) are analogous to the generalized Euler equation derived in the (quasi-)hyperbolic time discounting literature, cf., e.g., Harris and Laibson

(2001). The latter refer to (the analogue of) expression $\beta\nu^h_{h,h+1}\Theta_{h+1}$ as the "effective discount factor". The condition is easiest to interpret by noticing that $\Theta_{h+1} > 1$ iff $\omega_h \equiv \frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1}\cdot\nu^{h+1}_{h+1,h+2}} > 1$, which holds in our calibration of the CEU model. In this case the marginal propensity to save (MPS) next period, $1 - m_{h+1}$, receives a higher value than the MPC, m_{h+1}, and self h correspondingly expresses higher patience than according to the pure short-run discount factor $\beta\nu^h_{h,h+1}$. To gain further intuition observe that, as long as $\omega_h > 1$, the effective discount factor varies inversely with next period's MPC, just as in the hyperbolic time discounting model. If self $h+1$ values consumption more—by consuming more out of cash on hand—then self h compensates this overconsumption of her own future self by increasing impatience, hence by consuming more today and saving less.

Next, turn to the general case where $\Lambda_{h+1} \neq 0$. For sophisticated CEU agents the value functions of selves h and $h+1$ in periods $h+2$ are age-dependent. A positive difference $\frac{\partial V^h_{h+2}}{\partial x_{h+2}} - \frac{\partial V^{h+1}_{h+2}}{\partial x_{h+2}}$ means that self h's marginal valuation of cash-on-hand in period $h+2$ is higher than self $h+1$'s. Under such a positive difference self h accordingly values savings from $h+1$ to $h+2$ more than self $h+1$. This increases the RHS of (15) thereby increasing savings already at age h.

3.6 Aggregation over Households

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 (encompassing both, the working and the retirement period). 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 $(\mathcal{A} \times \mathcal{E}) \in \mathcal{Y}$.

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:	$\bar{a}_t = \int a_t \Phi_t(da_t \times d\eta_t)$,
income:	$\bar{y}_t = \int y_t(\eta_t)\Phi_t d\eta_t$,
total income:	$\bar{y}_t^{tot} = \bar{y}_t + r\bar{a}_t$,
saving rate:	$\bar{s}_t = \int s_t(a_t, \eta_t)\Phi_t(da_t \times d\eta_t)$, where $s_t(a_t, \eta_t) = 1 - \frac{c_t(a_t,\eta_t)}{y_t(\eta_t)+r\cdot a_t}$.

In the quantitative section we also study average saving plans of naive CEU agents. By dynamic inconsistency, these agents update their plans in each period. As a way to compare any gap between plans made at age h and realizations at $t \geq h$ for CEU agents we denote the planned average saving rate with superscripts and compute

$$\tilde{s}_t^h = \int s_t^h(a_t, \eta_t)\Phi_t^h(da_t \times d\eta_t), \tag{18}$$

for all t. This gives hypothetical average profiles of the saving rate 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 (18) as the (average) "planned" saving rate. By dynamic consistency, we have for both RE and sophisticated CEU agents that

$$s_t^h(a_t, \eta_t) = s_t^1(a_t, \eta_t) \qquad \text{hence} \qquad \tilde{s}_t^h = \tilde{s}_t,$$

for all $h = 1, \ldots, T$. These equalities hold for naive CEU agents only for $t = h$ and, independent of current age h, for $t = T$.

4 Calibration

With the exception of the discount rate, all parameters are calibrated without using the life-cycle model. We refer to these parameters, summarized in Table 2, as (exogenous) first-stage parameters. The discount rate is accordingly referred to as (endogenous) second-stage parameter, cf. Table 3. The remainder of the section provides the details of our approach.

4.1 Household Age

Households enter the model at age 20 (model age 0). The retirement age is 62, hence $t_r = 42$, according to the average retirement age reported in the Survey of Consumer Finance (SCF).[15] We set the horizon to some maximum biological human lifespan at age 125,

[15] We compute the average retirement age by pooling the SCF waves 1992-2007 and exclude respondents younger than 45.

hence $T = 105$. This choice is motivated by estimates based on Swedish female life-table data by Weon and Je (2009).

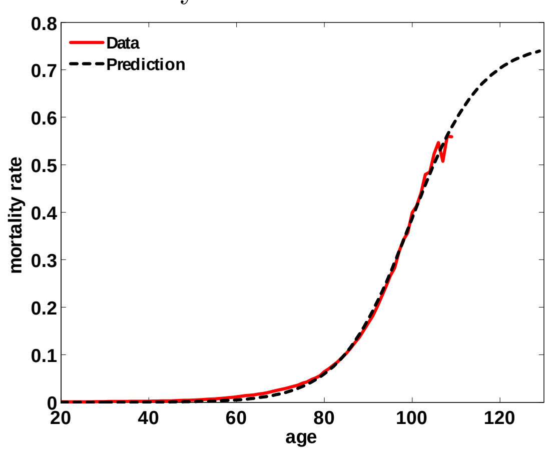
4.2 Objective Cohort Data

For objective survival rates we use average cross-sectional survival rates for the US between 2000-2010 taken from the Human Mortality Database (HMD). Data on survival rates becomes unreliable for ages past 100 as age-specific sample-size is low. Bebbington et al. (2011) argue, that a standard Gompertz-Makeham law, cf., e.g., Preston et al. (2001), is ill-suited for estimating human survival rates at high ages. This is due to the fact that human mortality, while first increasing exponentially with age, finally decelerates for high ages past 95. To account for this mortality deceleration we follow Bebbington et al. (2011) by applying the logistic frailty model. Accordingly, the mortality rate μ_t at age t obeys

$$\mu_t = \frac{A \exp(\alpha \cdot t)}{1 + s^2 (\exp(\alpha \cdot t) - 1) \frac{A}{\alpha}} + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma^2), \tag{19}$$

where the term in the denominator corresponds to the standard Gompertz-Makeham law. We estimate parameters to get an out of sample prediction for ages past 100. The resulting predicted mortality rate function fits actual data very well, cf. Figure 3. We use it as objective cohort data in the simulation.

Figure 3: Mortality Rates: Data vs. Estimation



Notes: Average mortality rates from 2000-2010 using HMD data (red solid line) and predicted mortality rates (black dashed line) using the logistic frailty model given in (19). Parameter estimates are provided in Table 2.

According to our parameter estimates reported in Table 2, the implied average mortality rate converges to a value of 0.57 at ages around 110 ($t = 90$). This is well in line with Gampe (2010) who reports an annual mortality rate of around 0.5 for persons past age 110 using data for a series of OECD countries on mortality rates of supercentenarians.

4.3 Estimated Subjective Survival Beliefs

We follow Ludwig and Zimper (2013) and estimate parameters δ and λ, cf. equations (2) and (3), to match the HRS data. Subjective survival rates are obtained by pooling a sample of HRS waves $\{2000, 2002, 2004\}$. Except for heterogeneity in sex and age, we ignore all other heterogeneity across individuals. Before proceeding with the estimation, the experience function $e(h)$ remains to be specified. It would be desirable to do so with a flexible functional form. Unfortunately, as Ludwig and Zimper (2013) demonstrate in their online appendix, identifying $e(h)$ is not straightforward because of the interplay with the other model parameters δ and λ. That said, setting $e(h) \equiv \sqrt{h}$ seems to be a good choice because the corresponding model calibration gives rise to parameter values of δ and λ that are in line with experimental evidence (see, e.g., Wakker 2010; Abdellaoui et al. 2011). Specifically, the estimation yields $\delta = 0.135$ and $\lambda = 0.418$.[16] As documented in Ludwig and Zimper (2013), these parameters are estimated with very high precision.

The predicted subjective survival rates resulting from our model of ambiguous survival beliefs fit their empirical counterparts, i.e., the average subjective survival beliefs for each interview age h, from the HRS quite well, cf. Figure 4 in Section 5. The R^2 of the regression is around $0.8 - 0.95$.[17]

4.4 Prices and Endowments

Wages are normalized to $w = 1$. We take a three-state first-order Markov chain for the income process in periods $t = 0, \ldots, t_r - 1$ with state vector $E^w = [1 + \epsilon, 1 - \epsilon, 0]$. The last entry reflects the state with zero income. Following the estimates of Carroll (1992) we set the probability of receiving zero labor income to $\zeta = 0.005$. Then the transition

[16] Estimation results are calculated 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. Differences in point estimates from those reported in Ludwig and Zimper (2013) are due to the fact that we use a simplified version of their model, cf. Subsection 2.2.

[17] The fit is slightly better for women than for men, cf. Ludwig and Zimper (2013). They further perform sensitivity analyses with regard to the choice of the initial age, the specific form of the experience function and focal point answers. This shows that results do not hinge on these aspects. Finally, they document that biases in beliefs are neither due to cohort effects nor selection biases.

matrix during the working period writes as

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

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 et al. (2004) yielding $\kappa = 0.97$ and $\epsilon = 0.68$. Age specific productivity $\{\phi_t\}$ of wages is estimated based on PSID data applying the method developed in Hugget et al. (2007).

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

The interest rate is set to $r = 0.042$ based on Siegel (2002). For the social security contribution rate we take the US contribution rate of $\tau = 0.124$. The pension benefit level then follows from the social security budget constraint (6).

4.5 Preferences

Recall that we normalize utility from death to zero, i.e., if the household dies at the end of period $t-1$ we let $u(c_t) = u(0) = 0$. As to utility from survival we take a $CRRA$ per period utility function with coefficient of relative risk aversion θ. For the intertemporal elasticity of substitution (IES), $1/\theta$, we take a standard value chosen in the literature of $1/3$, i.e., $\theta = 3$. This choice implies that a standard CRRA utility function of the form $u(c_t) = \frac{c_t^{1-\theta}}{1-\theta}$ is negative for all $c_t > 0$. This would violate our assumption that utility from survival is positive thereby exceeding utility from death. We cure this by two additional modifications of the utility function. First, we add an additive preference shifter to the per period utility function, denoted by $\Upsilon > 0$. With this monotone transformation we can ensure (via calibration) that the utility from survival is always positive. Of course, this does not affect optimal choices. Second, we take a Stone-Geary specification of the utility function and accordingly let $c_t - \underline{c}$ be its argument for some very small $\underline{c} > 0$. To understand this second modification observe that our specification of the income process with a positive zero income probability achieves differentiability

Table 2: First-Stage Parameters

Parameter		*Source*
Technology and Prices		
$w = 1$	Gross wage	normalized
$r = 0.042$	Interest rate	Siegel (2002)
$\tau = 0.124$	Social security contribution rate	irs.gov
$\chi = 0.322$	Net pension benefit level	model outcome
Income Process		
$\kappa = 0.97$	Persistence of income	Storesletten et al. (2004)
$\epsilon = 0.68$	Variance of income	Storesletten et al. (2004)
$\{\phi_t\}$	Age specific productivity	PSID
$\zeta = 0.005$	Probability of zero labor income	Carroll (1992)
$\zeta^r = 0.001$	Probability of zero retirement income	
Preferences		
$\theta = 3$	Coefficient of relative risk aversion	
$\underline{c} = 1.0e-08$	Minimum consumption level	
Subjective Survival Beliefs		
$\delta = 0.135$	Initial degree of ambiguity	HRS
$\lambda = 0.418$	Degree of relative optimism	HRS
Age Limits and Survival Data		
0	Initial model age (age 20)	
$t_r = 42$	Retirement (age 62)	SCF
$T = 105$	Maximum human lifespan (age 125)	Weon and Je (2009)
$\{\psi_{k,t}\}$	Cohort survival rates	Predictions based on HMD
$s = 0.41$	Logistic frailty model	
$\alpha = 0.13$	Logistic frailty model	
$A = 2.9e-06$	Logistic frailty model	

Notes: First-stage parameters that are calibratied outside the life-cycle model.

Table 3: Second-Stage Preference Parameter: The Subjective Discount Rate

Target (Source)	Asset profile (SCF)
RE	$\rho^{RE} = 0.0344$
naive CEU	$\rho^{CEU,n} = 0.0340$
sophisticated CEU	$\rho^{CEU,s} = 0.0427$

Notes: Second-stage parameters are calibratied such that asset moments from the model best match corresponding data moments.

of policy functions and positive asset holdings (and hence consumption) throughout but very low consumption levels have positive probability. This makes it very hard to assign values to Υ such that utility from survival is always positive. With a Stone-Geary-CRRA utility function we have that optimal consumption choices satisfy $c_t > \underline{c}$ because of the lower Inada condition. Accordingly, setting $\Upsilon = -u(\underline{c})$ achieves strictly positive utility in case of survival. Of course, for very small $\underline{c}$, the effect of this modification on optimal choices as well as on the IES is negligibly small. Collecting elements, the per-period utility function reads as

$$u\left(c_t\right) = \Upsilon + \frac{\left(c_t - \underline{c}\right)^{1-\theta}}{1-\theta}$$

for $\Upsilon = -\frac{\underline{c}^{1-\theta}}{1-\theta}$.

A key preference parameter of the model is the discount rate which we take as the only second stage parameter of our model. We calibrate it such that the average asset-to-permanent-income ratio from the model best matches the empirical counterpart. This approach is in the spirit of Gourinchas and Parker (2002) and De Nardi et al. (2010). Data on assets and permanent income is taken from the SCF. Appendix D describes in more detail how the data is constructed.

Denote by $\bar{a}_t^{data}$ average age-specific net-worth and by $\bar{y^p}_t^{data}$ average permanent-income constructed by pooling SCF data from 1992 to 2007. As defined in Section 3.6, $\bar{a}_t$ is the model counterpart. Correspondingly, we denote model permanent income by $\bar{y}_t^p$ which is calculated as the constant annuity payment from the net present value of average (labor, respectively retirement) income $\bar{y}_t$ over the life-cycle discounted with the riskfree interest rate $r = 0.042$.

We target the life-cycle profile between ages 30 ($t_0 = 10$) and 90 ($T_0 = 70$). A starting age of 30 is motivated by the fact that we do not explicitly model education decisions so that our model does not match the data well at very young ages. Our choice of the terminal age at 90 is due to data limitations at very high ages. Beyond age 90 there are too few observations on assets in the data so that (smoothed) asset age profiles get rather wiggly. Accordingly we search for ρ to solve

$$\min_{\rho} \frac{1}{2} \sum_{t=t_0}^{T_0} \left(\frac{\bar{a}_t^{data}}{(\bar{y}_t^p)^{data}} - \frac{\bar{a}_t\left(\rho\right)}{\bar{y}_t^p\left(\rho\right)} \right)^2 . \tag{20}$$

For our baseline results, we calibrate a different subjective time discount rate ρ for each of the three models, the RE, the naive and the sophisticated CEU model. Parameter estimates in Table 3 document that the difference between subjective discount factors calibrated for the RE and the naive CEU model is small whereas the difference to the sophisticated CEU model is large. In Section 5.2 we explain the reason for these differences. Importantly, we also investigate how results are affected by recalibration.

In these experiments we hold the discount rate constant at its calibrated value for the RE agent.

5 Results

5.1 Ambiguous versus Rational Survival Beliefs

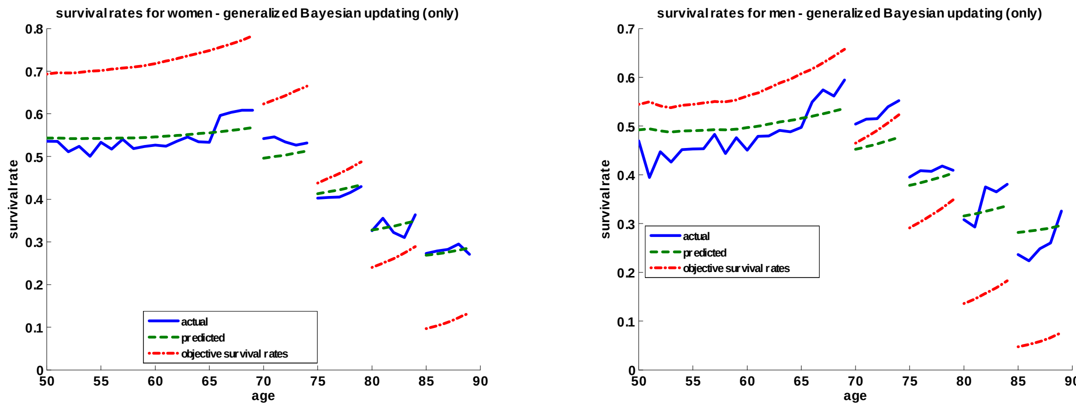
Figure 4 compares predicted subjective survival rates resulting from our model of ambiguous survival beliefs with their empirical counterparts and corresponding objective survival rates for men in Panel (a) and for women in Panel (b). Interview age, cf. Table 1, is shown on the abscissa. Actual subjective survival beliefs are depicted in the figure as a blue solid line and corresponding objective beliefs as a red dashed-dotted line. To understand this figure, recall that actual subjective survival beliefs are elicited in the HRS only for a combination of interview ages and target ages as shown in Table 1. The step function of corresponding objective beliefs follows from changes in the interview age / target age assignment. For example, a 69 year old person is asked about her subjective assessment to live until age 80 whereas a 70 year old is asked about her probability to reach age 85. The chance to live from 69 to 80 is much higher than the chance to live from 70 to 85. Therefore, objective survival beliefs drop discretely between interview ages 69 and 70. Furthermore, within each interview age / target age bin, objective survival rates generally increase. For example, the chance to survive from age 60 to 80 is lower than the chance to survive from age 61 to 80.[18] Finally, the figure shows as a green dashed line the predicted subjective survival rates from our model for the parameter estimates of δ and λ as given in Table 2. Overall, we can conclude from this figure that the fit of predicted to actual subjective survival rates is very good. In particular, the model replicates underestimation of survival rates at younger ages and overestimation at older ages.

Figure 5 shows the age-specific degree of ambiguity as a function of planning age h. While the initial degree of ambiguity at age 20 is $\delta = 0.135$, cf. Table 2, ambiguity at age 50 is roughly at 0.51. As predicted by the theoretical model, cf. Proposition 1, ambiguity is monotonically increasing and, as a consequence of the square-root experience function, $e(h) = \sqrt{h}$, concave in planning age h. This pattern, in particular the age-increasing ambiguity, is important to understand the survival rate functions of our

[18] On the other hand, our cohort based prediction of objective survival rates incorporates trends in life-expectancy. In particular at relatively "young" ages it may therefore be that the objective survival rate curve is downward sloping within interview age / target age bins. For example, the objective survival rate of a 52 year old man to live to age 80 turns out to be slightly higher than of a 53 year old man because the 52 year old man belongs to a younger cohort.

Figure 4: Objective, Subjective and Predicted Subjective Survival Rates

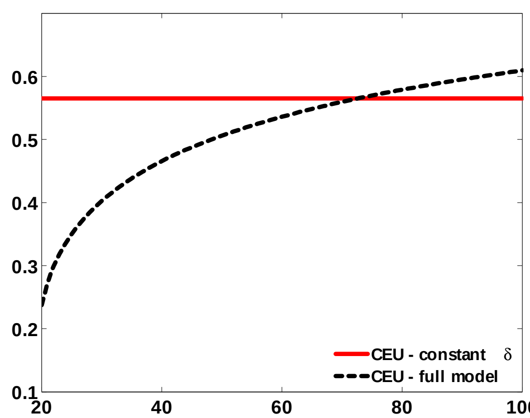
(a) Women (b) Men



Notes: Unconditional survival probabilities to different specific target ages according to the questions in the HRS, cf. Table 1. Interview age is on the abscissa. The solid blue line are subjective survival beliefs, the dashed-dotted red line are the corresponding objective survival rates and the dashed green line are simulated subjective survival beliefs from the estimated CEU model.

model to which we turn next. The figure also shows a line with constant ambiguity which will be relevant for our sensitivity analysis, cf. Section 5.2.3.

Figure 5: Ambiguity over the Life-Cycle



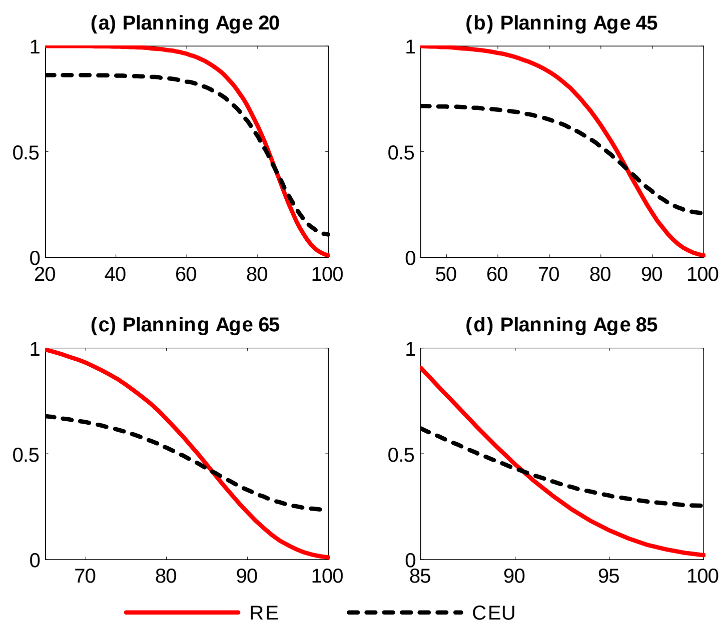
Notes: Degree of ambiguity δ as a function of planning age h. The average ambiguity of men and women is used in the calibrated life-cycle model.

Figure 6 compares ambiguous subjective survival functions—i.e., the subjective hazard rates—as black dashed lines to their objective counterparts as red solid lines. The four panels of the figure represent different planning ages h. Panel (a) is for planning age 20 ($h = 0$) and Panel (d) for planning age 85, ($h = 65$). In each of the four panels of the figure, target age t is depicted on the abscissa. Within each panel, experience—

and therefore the ambiguity parameter δ_h—is unaltered. Across panels, ambiguity is increasing according to the pattern of Figure 5. The initial point of survival functions at age $t = h$ is driven by ambiguity at that age. As planning age h increases, i.e., as we move from Panel (a) to Panel (d), the distance of this point to a survival rate of 1 increases. The size of this initial blip is large. It results from the parsimonious structure of our model but otherwise does not affect our results much, cf. Section 5.2.3.

The key observation from the figure is that subjective survival functions are flatter than their objective counterparts which is in line with Hammermesh (1985), Peracchi and Perotti (2010), Elder (2013) and several others. Furthermore, ambiguous survival beliefs match the stylized fact described by Wu et al. (2013): People at a specific planning (or interview) age underestimate their chances of survival to the nearer future and overestimate survival probabilities to the more distant future. Also notice that the overestimation of survival probabilities becomes more pronounced as the agent gets older. I.e., the point at which the subjective and the objective survival curves intersect moves to the left when moving across the figure from Panel (a) to Panel (d).

Figure 6: Survival Functions



Notes: Unconditional objective and subjective probabilities viewed from different planning ages h. Target age t is depicted on the abscissa.

5.2 Life-Cycle Profiles with Ambiguous Beliefs

5.2.1 Baseline Calibration

To highlight the effects of modeling subjective survival beliefs on life-cycle savings we conveniently compress all information by showing average asset holdings of CEU agents compared to RE agents who use objective survival data. We focus on the average asset-to-permanent income ratio as described in the calibration section. We scale assets with the same annuity value as the one used for estimating preference parameters, cf. Section 4.5.

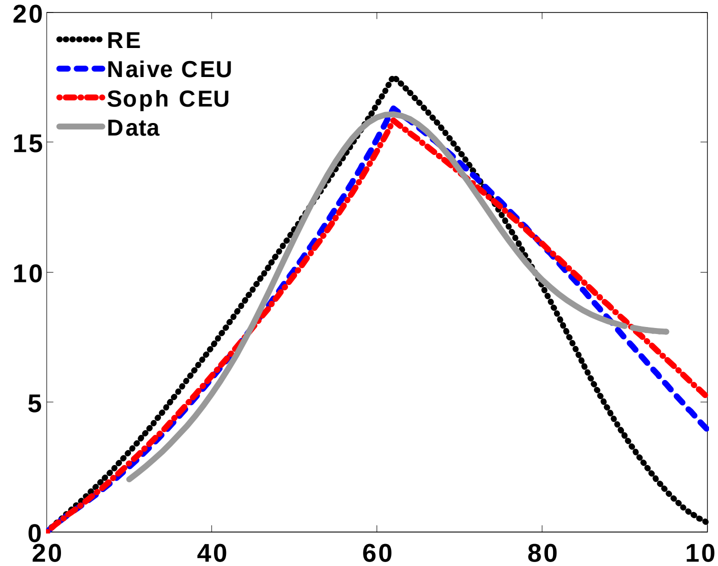
Results are shown in Figure 7. The figure shows average asset holdings over the life-cycle for the three types of agents, RE agents as the black dotted line, naive CEU agents as the blue dashed line and sophisticated CEU agents as the red dashed-dotted line. The profiles of our calibrated models are compared to the data (grey line). Assets steadily increase until retirement entry and fall thereafter. This implies positive saving rates during working life while agents dissave during retirement.

The overall shape of life-cycle asset holdings is explained as follows: due to prudence, agents save for precautionary reasons to self-insure against future income risk. There are two forces triggering precautionary saving. One is the standard income risk, the second is the risk of drawing zero labor income. Since the latter gives rise to a self imposed borrowing constraint, asset holdings throughout the life-cycle are always positive. As agents become older, life-cycle motives for saving become more and more relevant and motives for precautionary saving become less strong, also see, e.g., Gourinchas and Parker (2002). Assets are being accumulated in order to finance retirement consumption. In retirement, the only precautionary motive to save is to avoid zero resources in all income states. This motive again becomes more and more relevant as asset holdings converge towards zero when agents get older.

With regard to differences in asset accumulation across types, first focus at the RE type, displayed in the figure as a dotted line. Relative to the data, the dynamically consistent RE model features higher saving and therefore stronger asset accumulation on average until retirement and a faster speed of asset decumulation thereafter. Accordingly, through the lens of the RE model the data are puzzling: households save too little until retirement and have asset holdings in old age that are too high. Any attempt to improve the fit of the RE model by, e.g., decreasing the discount rate would lead to a lower speed of asset decumulation at the cost of even higher saving during the working period and vice versa.

On the contrary, the calibrated naive CEU model gives rise to less saving during the accumulation phase and a much slower speed of asset decumulation than for the RE

Figure 7: Assets-to-Permanent Income, CEU, RE and Data



Notes: Average asset-to-permanent-income ratios from SCF data and from CEU and RE agents using recalibrated preference parameters ρ for all models. The data covers ages 30 and 95, details of how the data is constructed is described in Appendix D.

model, moving it close to the data. The driving force for undersaving (relative to the RE model) is pessimism with regard to survival prospects. The reason for high old-age asset holdings is the strong optimism with regard to surviving into the future, cf. Figure 6. The optimism about future survival prospects gets stronger as the agent ages, eventually becoming the dominating force for trends in assets. This strength of optimism is crucially governed by our (precise) estimate of the relative optimism parameter, λ.

The sophisticated CEU model generates very similar results compared to the naive CEU model: on average saving rates during the working period are almost identical and so is asset accumulation. Old-age asset holdings of sophisticated agents are slightly higher than those of naive agents. The reason is that sophisticates, by foreseeing the optimistic biases of their own future selves, decumulate assets at a lower speed for reasons of consumption smoothing. The close similarities between the two CEU agents only occur because we recalibrate the discount rate. It is almost one percentage point higher for the sophisticated agent, cf. Table 3. We discuss this in detail below in Section 5.2.2, where we also provide additional interpretation for our findings.

Table 4 comprises our results by reporting summary statistics for all three agent types and the data. As a summary statistic for the goodness of fit of the three models we report the R^2s from the non-linear regressions in (20). While R^2 looses its usual interpretation in non-linear models as a measure of the fraction of the overall variation explained by the model, it is still a useful summary statistic of goodness fit. It is bounded from above by 1 and a value closer to 1 indicates better fit. Results on the R^2s confirm the visual impression gained from Figure 7, i.e., the fit of the naive CEU model is best and the one of the RE model is worst.

The average saving rate of both the naive and sophisticated CEU agents during the prime saving years, ages 25-54, is about 9.2%. The corresponding average saving rate in the US is 9.5%.[19] On the contrary, RE agents save on average 13.5%, exceeding the relevant data by 4 percentage points.

Table 4: Summary Statistics

	RE	*CEU*		*Data*[1)]
		Naive	Soph.	
R^2	0.756	0.944	0.935	
Saving rate[2)]	13.5%	9.2%	9.2%	9.5%
Saving rate, planned[2)]	–	14.7%	–	
Assets at age 75 relative to 62[3)]	70.0%	77.9%	79.0%	72.4%
Assets at age 85 relative to 62[3)]	37.0%	57.1%	60.8%	53.0%
Assets at age 95 relative to 62[3)]	*9.1%*	*34.9%*	*42.1%*	*47.9%*

[1)] The data for asset decumulation is calculated from SCF data. Due to small sample sizes, SCF data on average asset holdings at age 95 cannot be measured reliably and are thus reported in italics. The saving rate is the weighted average of ages 25-54 between 1980-85 from the Consumer Expenditure Survey (CES) as reported by Bosworth et al. (1991), Table 3.

[2)] The average saving rate as is defined as the average of individual saving rates between ages 25 and 54. The average planned saving rate is the rate for ages 25-54 planned at age 20.

[3)] Average asset holdings at age 75, 85 and 95 relative to assets at retirement entry at age 62.

Comparing plans and realizations for naive CEU agents we observe that, initially, CEU agents plan to save more and consume less during working life which would result in higher assets. The planned average saving rate of naive CEU agents at age 20 for ages 25-54 is 14.7%, compared to the average realized saving rate for that age bin of 9.2%. The fact that actual saving behavior deviates from plans naturally follows from time inconsistency. That saving is lower than planned means that households moderately overestimate their future survival rates, leading us back to the predictions of the simple 3-period model. If overestimation was stronger, then they would actually save

[19]The SCF does not contain quantitative questions on saving, only qualitative ones such as whether one had positive saving. Furthermore, as the SCF does not have a panel dimension, we cannot compute savings from changes in assets. Thus, we chose CES data as reported by Bosworth et al. (1991).

more than originally planned. These patterns are qualitatively consistent with findings in the literature on undersaving: Barsky et al. (1997) document that agents have a preference for constant or upward sloping consumption paths which cannot be achieved by observed saving rates. Lusardi and Mitchell (2011) present survey results showing that out of those households that made a retirement savings plan, the majority was not able to stick to their plan. Finally, Choi et al. (2006) document that two thirds of respondents in a survey have saving rates below their ideal ones.

Finally, Table 4 also summarizes the sizeable differences in old-age asset holdings between RE and CEU agents. For naive CEU agents, average asset holdings at ages 75, 85 and 95 relative to those at retirement entry are 77.9%, 57.1% and 34.9% compared to 72.4%, 53.0% and 47.9% in the data. Recall that the last data point, i.e., asset holdings at age 95, has to be looked at with care because of few observations. On the contrary, these values are only at 70.0%, 37.0%, and 9.1% for RE agents. Sophisticated CEU agents have even higher assets at old age relative to assets at retirement entry. Also notice that the overall fit of the sophisticated CEU model to the data is worse than for naive agents.

We can therefore conclude that the combination of ambiguous survival beliefs with the assumption of naivity has to be considered as a candidate explanation for the joint occurrence of low retirement savings, time inconsistent saving behavior and high old-age asset holdings.

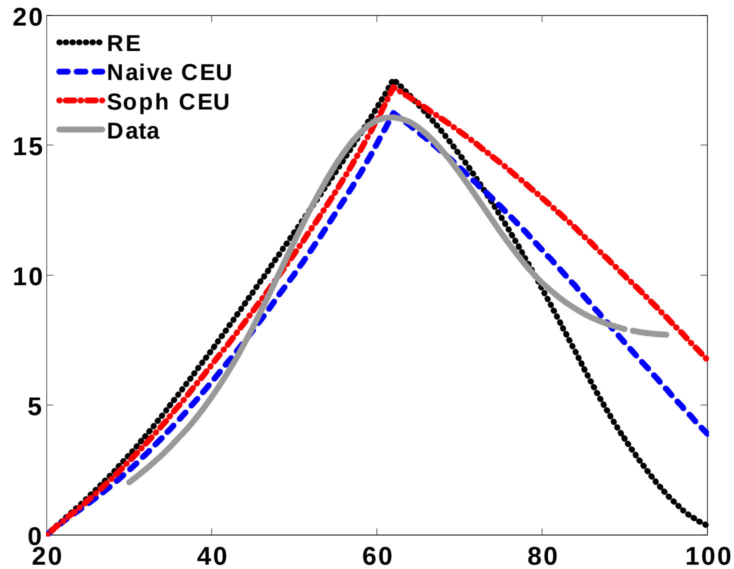
5.2.2 The Effects of Discounting

As described in Section 4, the calibrated discount rate varies across all model variants in our baseline results, cf. Table 3. This section documents how our main findings are affected by this approach to calibration. To this end, we hold constant the value of the discount rate calibrated for the RE model of 3.4% and use it in the two variants of the CEU model. With this strategy we single out the pure effects of ambiguous survival beliefs. Results on asset holdings are displayed in Figure 8 and corresponding summary statistics are provided in Table 5.

As the difference of calibrated discount rates between the RE and the naive CEU models is not large, cf. Table 3, our results do not change much for naive CEU agents. With the lower RE-model discount rate, saving increases slightly and hence the average asset decumulation speed also decreases. The R^2, as a summary statistic for the goodness of fit, decreases very mildly to 0.943.

Significant changes occur for the sophisticated CEU model where the calibrated discount rate is almost one percentage point higher in our baseline calibration. Relative to this, the R^2 strongly decreases to 0.817. The saving rate during the working period

Figure 8: Assets-to-Permanent Income, CEU, RE and Data: Constant ρ



Notes: Average asset-to-permanent-income ratios from SCF data and from CEU and RE agents using $\rho^{RE} = 0.034$ as a preference parameter for all models. The data covers ages 30 and 95, details of how the data is constructed is described in Appendix D.

goes up to 11.4% which is—although still lower than for the RE agent—more than observed in the data.

We complement this picture by displaying life-cycle consumption relative to permanent income across agent types (again holding ρ constant) in Figure 9. At younger ages, both the naive and the sophisticated households consume more than RE agents. In the middle stages of the life-cycle optimism starts to dominate their survival belief formation. This increases the consumption growth rate so that both CEU agents consume less at middle and more at old age than the RE agents. Finally, observe from the figure that sophisticates indeed consume less at young ages than do naifs leading to higher asset holdings over the life-cycle.

Again, the simple three-period model from the Supplementary Appendix provides guidance for understanding these results. Sophisticates foresee the increasing optimism of their own future selves. Given their relatively low inter-temporal elasticity of substitution of 1/3 they therefore place a high value on the marginal utility from saving, give up on consumption when young and build up higher asset positions during the working period than their naive counterparts. In consequence, they also decumulate assets at a lower speed in old age.

Table 5: Summary Statistics: Constant ρ [1)]

	Naive CEU		*Soph. CEU*		*Data*
	Baseline	ρ^{RE}	Baseline	ρ^{RE}	
R^2	0.944	0.943	0.935	0.817	
Saving rate	9.2%	9.0%	9.2%	11.4%	9.5%
Saving rate, planned	14.7%	14.6%	–	–	
Assets at age 75 relative to 62	77.9%	77.8%	79.0%	83.1%	72.4%
Assets at age 85 relative to 62	57.1%	56.8%	60.8%	66.9%	53.0%
Assets at age 95 relative to 62	*34.9%*	*34.6%*	*42.1%*	*48.7%*	*47.9%*

[1)] For a description of how the statistics are constructed see Table 4.

5.2.3 The Effects of Experience

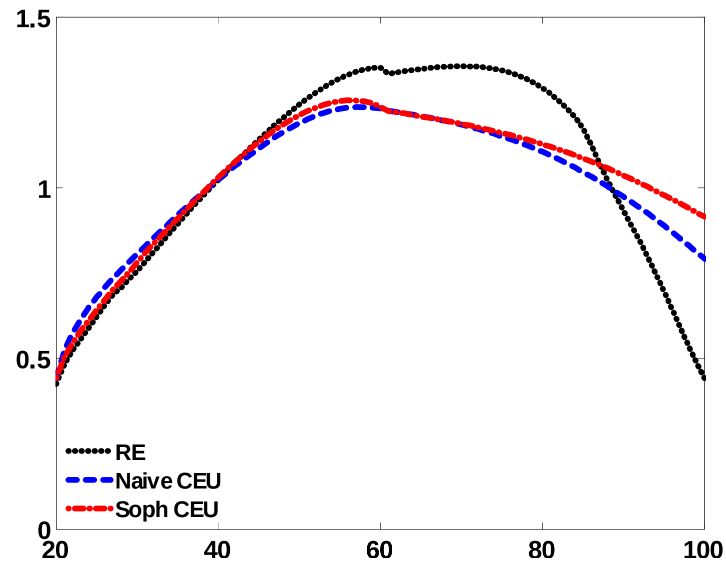
We next shut down experience by setting $e(h) = 0$ for all h so that our notion of ambiguous survival beliefs reduces to the static scenario of an age-independent neo-additive probability weighting function where $\delta_h = \delta$ for all h, cf. Proposition 1. We re-estimate parameters δ and λ which gives $\lambda = 0.424$ and $\delta = 0.565$. The high value of δ roughly corresponds to the mean value of δ_h in our baseline model, cf. Figure 5. It implies that initial biases of subjective survival functions are even stronger than those depicted in Figure 6.

Despite this, the key quantitative implications of our model are little affected. As documented in Table 6, the average saving rate decreases for both CEU types. For naive CEU agents, the difference between planned and realized saving rates increases. The asset decumulation speed during the retirement period, however, is little affected. Finally, notice that the R^2s decrease. This supports our preferred specification with dynamic updating of survival beliefs.

6 Concluding Remarks

This paper studies implications of ambiguous survival beliefs for consumption and saving behavior. Point of departure of our analysis is that people make mistakes in assessing their chances to survive into the future: "young" people tend to underestimate whereas "old" people tend to overestimate their survival probabilities. We adopt and parametrize a model of Bayesian learning of ambiguous survival beliefs which replicates these patterns. The resulting conditional neo-additive survival beliefs are merged into a stochastic life-cycle model with CEU (=Choquet expected utility) agents to study life-cycle consequences compared to agents with rational expectations (RE).

Figure 9: Life-Cycle Consumption of CEU and RE: Constant ρ



Notes: Average consumption relative to permanent income as described in the calibration section over the life-cycle from CEU and RE agents using $\rho^{RE} = 0.034$ as a preference parameter for all models.

We show that agents of our model behave dynamically inconsistent. As a result, CEU agents save less at younger ages than they actually planned to save. Due to underestimation of survival at young age, CEU agents also save less than RE agents. Despite this tendency to undersave, CEU agents eventually have higher asset holdings after retirement because of the overestimation of survival probabilities in old age. Overall, the CEU model provides an accurate quantitative picture of life-cycle asset holdings until about age 85. Furthermore, the assumption of naive CEU agents better fits the data than the assumption of sophisticated CEU agents. Our model of biases in the assessment of survival prospects therefore adds to explanations for three empirical findings: (i) time inconsistency of agents, (ii) undersaving at younger ages and (iii) high asset holdings at old age. Hence, our model hits at—but does not kill—"three birds with one stone".

Our work gives rise to several avenues of future research. First, observe that the ambiguous survival belief functions depicted in Figure 6 closely resemble quasi-hyperbolic time discounting functions, cf., e.g., Laibson (1997). In Groneck et al. (2014) we compare the formal relationship between quasi-hyperbolic time-discounting, on the one hand, and a CPT/CEU model, on the other hand. In contrast to the quantitative analysis of the present paper, which aims to bring the model to the data, the stylized analysis in Groneck et al. (2014) restricts attention to a static CPT/CEU model in which agents do not learn over the life-cycle. As our main finding we show that quasi-hyperbolic

Table 6: Summary Statistics, Constant δ [1)]

	Naive CEU		*Sophisticated CEU*		*Data*
	Baseline	*Const.* δ	*Baseline*	*Const.* δ	
Discount rate ρ	0.0340	0.0323	0.0427	0.0421	
R^2	0.944	0.930	0.935	0.925	
Saving rate	9.2%	8.6%	9.2%	8.9%	9.5%
Saving rate, planned	14.7%	16.6%	–	–	
Assets at age 75 relative to 62	77.9%	79.0%	79.0%	79.9%	72.4%
Assets at age 85 relative to 62	57.1%	58.5%	60.8%	62.1%	53.0%
Assets at age 95 relative to 62	*34.9%*	*36.0%*	*42.1%*	*43.2%*	*47.9%*

[1)] Results for the CEU model without experience and with recalibrated ρ as in the baseline scenario. Forr a description of how the statistics are constructed see Table 4.

time-discounting over the life-cycle is formally equivalent to a static CPT/CEU life cycle model with neo-additive capacities such that (i) the ambiguity parameter is positive whereas (ii) the optimism parameter is zero. Our analysis further implies that a positive optimism parameter rather than Bayesian learning under ambiguity is responsible for the qualitative feature that CPT/CEU agents might—in contrast to quasi-hyperbolic time-discounting agents—oversave at old ages.

Second, we plan to combine our notion of CEU agents with bequest motives in order to cover important aspects of life-cycle decisions. The main challenge for this generalizing approach will be to come up with a parsimonious model in which all calibrated behavioral parameters are identified.

Third, we will extend our framework to address normative questions on the optimal design of the tax and transfer system, similar to Laibson et al. (1998), Imrohoroglu et al. (2003) and, more recently, Pavoni and Yazici (2012, 2013) in the hyperbolic time discounting literature.

References

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

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

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

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.

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

Bebbington, M., C.-D. Lai and R. Zitikis (2011). Modelling deceleration in senescent mortality, Mathematical Population Studies: An International Journal of Mathematical Demography, 18(1), 18–37.

Berk, R. H. (1966). Limiting Behaviour of Posterior Distributions when the Model is Incorrect. The Annals of Mathematical Statistics 31, 51–58.

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.

Bissonnette, L., M. D. Hurd, and P.-C. Michaud (2011). Individual Survival Curves Combining Subjective and Actual Mortality Risks. NetsparWorking Paper.

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

Bloom, D. E., D. Canning, M. Moore, and Y. Song (2006). The Effect of Subjective Survival Probabilities on Retirement and Wealth in the United States. NBER Working Paper 12688.

Bosworth, B., G. Burtless and J. Sabelhaus (1991). The decline in saving: Evidence from household surveys. Brookings Papers on Economic Activity 1991(1), 183–256.

Cagetti, M. (2003), Wealth Accumulation Over the Life Cycle and Precautionary Savings, Journal of Business and Economic Statistics, 21(3), 339–353.

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., 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.

Doob, J. L. (1949). Application of the Theory of Martingales. Le Calcul des Probabilités et ses Applications, Colloques Internationaux du Centre National de la Recherche Scientifique 13, 23–27.

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

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

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. Quarterly Journal of Economics 75, 643–669.

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

Epstein, L. G. and M. Le Breton (1993). Dynamically Consistent Beliefs Must Be Bayesian. Journal of Economic Theory 61(1), 1-22.

Epstein L. G. and M. Schneider (2007). Learning under Ambiguity. Review of Economic Studies 74, 1275-1303.

Fernandez-Villaverde, J. and D. Krueger (2007). Consumption over the Life Cycle: Facts from Consumer Expenditure Survey Data. The Review of Economics and Statistics 89(3), 552–565.

Friedman, B. and M.J. Warshawsky (1990). The Cost of Annuities: Implications for Savings Behavior and Bequests. Quarterly Journal of Economics 104(2), 135-154.

Gampe, J. (2010). Human mortality beyond age 110. In: H. Maier et al. (eds.), Supercentenarians, Demographic Research Monographs, pp. 219–230. Springer 2010.

Gan, L., M. D. Hurd, and D. L. McFadden (2005). Individual subjective survival curves. In Analyses in the economics of aging, pp. 377–412. University of Chicago Press.

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

Gilboa, I. (1987). Expected Utility with Purely Subjective Non-Additive Probabilities. Journal of Mathematical Economics 16(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., Ludwig, A., and A. Zimper (2014). Cumulative Prospect Theory versus Quasi-hyperbolic Discounting. mimeo.

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

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.

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.

Hurd, M. D., D. L. McFadden, and L. Gan (1998). Subjective Survival Curves and Life Cycle Behavior. In Inquiries in the Economics of Aging, pp. 259–309. University of Chicago Press.

Imrohoroglu, A., S. Imrohoroglu, and D. H. Joines (2003). Time Inconsistent Preferences and Social Security. Quarterly Journal of Economics 118 (2), 745-784.

Khwaja, A., F. Sloan, and S. Chung (2007). The Relationship between Individual Expectations and Behaviors: Mortality Expectations and Smoking Decisions. Journal of Risk and Uncertainty 35(2), 179–201.

Laibson, D. I. (1997). Golden Eggs and Hyperbolic Discounting. Quarterly Journal of Economics 112 (2), 443–477.

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. (2013). Incidental Bequests: Bequest Motives and the Choice to Self-Insure Late-Life Risks. Working Paper.

Ludwig, A. and A. Zimper (2006). Investment Behavior under Ambiguity: The Case of Pessimistic Decision Makers. Mathematical Social Sciences 52, 111-130.

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

Lusardi, A. and O. S. Mitchell (2011). Financial Literacy and Planning: Implications for Retirement Wellbeing. NBER Working Papers 17078.

Marinacci, M. (2002). Learning from Ambiguous Urns. Statistical Papers 43, 145-151.

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

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

Nicholls, N., Romm, A. T., and A. Zimper (2014). The Impact of Statistical Learning on Violations of the Sure-Thing Principle. Accepted for publication in Journal of Risk and Uncertainty.

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

Palumbo, M. G. (1999). Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle. Review of Economic Studies 66(2), 395–421.

Pavoni, N. and H. Yzici (2012). Optimal Life-cycle Capital Taxation under Self-Control Problems, Working Paper.

Pavoni, N. and H. Yzici (2013). Present-Biased Preferences and Optimal Taxation of Parental Transfers, Working Paper.

Peracchi, F. and V. Perotti (2010). Subjective Survival Probabilities and Life Tables: Evidence from Europe. EIEF Working Paper 16/10.

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

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

Preston, S. H., P. Heuveline, and M. Guillot (2001). Demography - Measuring and Modeling Population Processes. Population and Development Review 27, 365.

Salm, M. (2010). Subjective Mortality Expectations and Consumption and Saving Behaviours among the Elderly. Canadian Journal of Economics 43(3), 1040–1057.

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.

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

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

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

Weon, B. M. and J. H. 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.

Wu, G. and R. Gonzalez (1999). Nonlinear Decison Weights in Choice under Uncertainty. Management Science 45(1), 74–85.

Wu, S., R. Stevens, and S. Thorp (2013). Die Young or Live Long: Model Subjective Survival Probabilities. Centre for the Study of Choice (CenSoC) Working Paper No. 13-004.

Zimper, A. (2011). Do Bayesians Learn their Way out of Ambiguity? Decision Analysis 8, 269-285.

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.

Zimper, A. and W. Ma (2014). Bayesian Learning with Multiple Priors: The Role of Prior-Revision Rules. Mimeo.

A Appendix: Choquet Decision Theory

A.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):

$$\mathbb{E}[f, \kappa] = \int_{-\infty}^{0} (\kappa(\{\omega \in \Omega \mid f(\omega) \geq z\}) - 1)\, dz + \int_{0}^{+\infty} \kappa(\{\omega \in \Omega \mid f(\omega) \geq z\})\, dz. \tag{21}$$

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 (21) becomes

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

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 1. *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*

$$\nu(A) = \delta \cdot \nu_\lambda(A) + (1 - \delta) \cdot \mu(A) \tag{22}$$

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

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

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 (22) 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 1. *Let $f : \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable function with bounded range. The Choquet expected value (21) of f with respect to a neo-additive capacity (22) is then given by*

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

A.2 The Generalized Bayesian Update Rule

CEU theory has been developed in order to accommodate paradoxes of the Ellsberg (1961) type which show that real-life decision-makers violate Savage's (1954) sure thing principle. Abandoning 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 et al. 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. Epstein and Le Breton 1993; Ghirardato 2002; 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 $A, B \in \mathcal{F}$,

$$\kappa(A \mid B) = \frac{\kappa(A \cap B)}{\kappa(A \cap B) + 1 - \kappa(A \cup \neg B)}. \tag{23}$$

An application of (23) to a neo-additive capacity ν gives rise to the following observation.

Observation 2. *If the Generalized Bayesian update rule (23) is applied to a neo-additive capacity (22), 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

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

B Appendix: Bayesian Learning of Ambiguous Survival Beliefs

This appendix derives Proposition 1 and it constructs the neo-additive probability spaces which we use when we define the life-cycle CEU maximization problem in Section 3.4. To this purpose, we present a simplified version of the learning model of ambiguous survival beliefs as developed 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 with $k < t$. By assumption, these individuals have the same i.i.d. objective survival probability as the agent.

B.1 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

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

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 $\mathbb{E}\left[\tilde{\theta}, \mu\left(\tilde{\theta}\right)\right] = \frac{\alpha}{\alpha+\beta}$. 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.[20]

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

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

[20] The gamma function is defined as $\Gamma(y) = \int\limits_0^{\infty} x^{y-1} e^{-x} dx$ for $y > 0$.

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

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

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 $\mathbb{E}\left[\tilde{\theta}, \mu\left(\tilde{\theta}\right)\right]$, not including any sample information, and the observed sample mean $\frac{j}{n(h)}$.

B.2 Ambiguous Survival Beliefs

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 $\mathbb{E}\left[\tilde{\theta}, \mu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right]$ we now suppose that the agent's estimator for her survival chance is given as the conditional Choquet expected value

$$\mathbb{E}\left[\tilde{\theta}, \nu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right]=\delta_{\tilde{I}_{n(h)}}\left(\lambda \sup \tilde{\theta}+(1-\lambda) \inf \tilde{\theta}\right)+\left(1-\delta_{\tilde{I}_{n(h)}}\right) \mathbb{E}\left[\tilde{\theta}, \mu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right]. \tag{24}$$

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

Observation 3. *The Choquet decision maker's ambiguous survival belief is given as*

$$\mathbb{E}\left[\tilde{\theta}, \nu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right]=\delta_{\tilde{I}_{n(h)}} \cdot \lambda+\left(1-\delta_{\tilde{I}_{n(h)}}\right) \cdot \mathbb{E}\left[\tilde{\theta}, \mu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right], \tag{25}$$

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

$$\begin{aligned}
\mu\left(\tilde{I}_{n(h)}=j\right) &= \binom{n(h)}{j} \frac{(\alpha+j-1) \cdot \ldots \cdot \alpha \cdot(\beta+n(h)-j-1) \cdot \ldots \cdot \beta}{(\alpha+\beta+n(h)-1) \cdot \ldots \cdot(\alpha+\beta)}, \\
\text{for } j &\in \{0, \ldots, n(h)\}.
\end{aligned} \tag{26}$$

In a next step, we employ several simplifying assumptions:

Assumption 1. *The additive part* $\mathbb{E}\left[\tilde{\theta}, \nu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right]$ *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 2. *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 (26) that*

$$\begin{aligned}\mu\left(\tilde{I}_{n(h)} = j\right) &= \binom{n(h)}{k} \frac{k!\,(n(h)-k)!}{(n(h)+1)\cdot n(h)!} \\ &= \frac{1}{1+n(h)}.\end{aligned}$$

Assumption 1 is an extreme version of the rational Bayesian learning part of the model developed in Appendix B.1. It specifies a correct additive prior and hence simplifies upon Ludwig and Zimper (2013).[21] 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 2 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 possibly observed sample information $\tilde{I}_{n(h)}$ if h is fixed.

Observation 4. *Under Assumptions 1-2, the estimator (25) simplifies to*

$$\begin{aligned}\mathbb{E}\left[\tilde{\theta}, \nu\left(\tilde{\theta} \mid \tilde{I}_{n(h)}\right)\right] &= \mathbb{E}\left[\tilde{\theta}, \nu\left(\tilde{\theta} \mid h\right)\right] \\ &= \delta_h \cdot \lambda + (1-\delta_h)\cdot \psi_{k,t}\end{aligned}$$

such that

$$\delta_h = \frac{\delta}{\delta + (1-\delta)\cdot \frac{1}{1+n(h)}}$$

whenever $h \leq k < t \leq T$*, i.e.,* $\psi_{k,t} \in (0,1)$*.*

Finally, identifying the h-old agent's subjective belief to survive from k to t with her estimator (24), i.e., defining $\nu_{k,t}^{h} \equiv \mathbb{E}\left[\tilde{\theta}, \nu\left(\tilde{\theta} \mid h\right)\right]$, and setting $e(h) \equiv n(h)$ for the agent's experience function gives the desired result of Proposition 1.

[21] 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.

B.3 Neo-additive Probability Spaces

It remains to provide a mathematically rigorous translation of the notion of ambiguous survival beliefs $\nu_{k,t}^{h}$ of Proposition 1 into the construction of the conditional neo-additive probability spaces $\left(\Omega, \mathcal{F}^{h}, \nu\left(\cdot \mid h\right)\right)$, $h = 1, ..., T$, that are relevant to the CEU life-cycle maximization problem of Section 3.4.

To this purpose define the finite state space $\Omega = \{0, 1, ..., T\}$ and denote by $\mathcal{F}$ 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.

For each age h, the σ-algebra $\mathcal{F}^{h}$ 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. Observe that our definition of $\mathcal{F}^{h}$ implies a standard information filtration process because of $\mathcal{F}^{1} \subset ... \subset \mathcal{F}^{T} = \mathcal{F}$.

To conclude the construction of $\left(\Omega, \mathcal{F}^{h}, \nu\left(\cdot \mid h\right)\right)$, $h = 1, ..., T$, define $\nu\left(Z_{k,t} \mid h\right) \equiv \nu_{k,t}^{h}$ such that $\nu_{k,t}^{h}$ is given by Proposition 1.

C Appendix: Proof of Propositions

C.1 Proof of Proposition 2

Fix age h and consider the neo-additive probability space $(\Omega, \mathcal{F}, \nu\left(\cdot \mid h\right))$ constructed in Appendix B.3. 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, (8) can be equivalently rewritten as

$$
\begin{aligned}
\mathbb{E}\left[U\left(\mathbf{c}\right), \nu^h\right] = \delta_h &\left(\lambda\left(u(c_h) + \sum_{t=h+1}^{T} \beta^{t-h}\mathbb{E}\left[u\left(c_t\right), \pi\left(\eta_t|\eta_h\right)\right]\right) + (1-\lambda)u(c_h)\right) \\
&+ (1-\delta_h)\left(u(c_h) + \sum_{t=h+1}^{T} \psi^h(D_t) \sum_{s=h+1}^{t} \beta^{s-h}\mathbb{E}\left[u\left(c_s\right), \pi\left(\eta_s|\eta_h\right)\right]\right) \\
= u(c_h) &+ \delta_h\lambda \sum_{t=h+1}^{T} \beta^{t-h}\mathbb{E}\left[u\left(c_t\right), \pi\left(\eta_t|\eta_h\right)\right] \\
&+ (1-\delta_h) \sum_{t=h+1}^{T} \psi_{h,t} \cdot \beta^{t-h}\mathbb{E}\left[u\left(c_t\right), \pi\left(\eta_t|\eta_h\right)\right] \\
= u(c_h) &+ \sum_{t=h+1}^{T} \nu^h_{h,t}\beta^{t-h}\mathbb{E}\left[u\left(c_t\right), \pi\left(\eta_t|\eta_h\right)\right],
\end{aligned}
$$

which, together with Proposition 1, proves the proposition.□

C.2 Proof of Proposition 4

The value functions of self h in periods h and $h+1$ are given by

$$
\begin{aligned}
V^h_h\left(x_h, \eta_h\right) &= \max_{c_h, x_{h+1}} \left\{u\left(c_h\right) + \beta\nu^h_{h,h+1}\mathbb{E}_h\left[V^h_{h+1}\left(x_{h+1}, \eta_{h+1}\right)\right]\right\} \\
V^h_{h+1}\left(x_{h+1}, \eta_{h+1}\right) &= \max_{c_{h+1}, x_{h+2}} \left\{u\left(c_{h+1}\right) + \beta\frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1}}\mathbb{E}_{h+1}\left[V^h_{h+2}\left(x_{h+2}, \eta_{h+2}\right)\right]\right\}.
\end{aligned}
$$

For self $h+1$ we accordingly have

$$
V^{h+1}_{h+1}\left(x_{h+1}, \eta_{h+1}\right) = \max_{c_{h+1}, x_{h+2}} \left\{u\left(c_{h+1}\right) + \beta\nu^{h+1}_{h+1,h+2}\mathbb{E}_{h+1}\left[V^{h+1}_{h+2}\left(x_{h+2}, \eta_{h+2}\right)\right]\right\}.
$$

The first-order conditions with respect to consumption for selves h and $h+1$ are given by

$$
\frac{du}{dc_h} = \beta R\nu^h_{h,h+1}\mathbb{E}_h\left[\frac{\partial V^h_{h+1}(\cdot)}{\partial x_{h+1}}\right] \tag{27a}
$$

$$
\frac{du}{dc_{h+1}} = \beta R\nu^{h+1}_{h+1,h+2}\mathbb{E}_{h+1}\left[\frac{\partial V^{h+1}_{h+2}(\cdot)}{\partial x_{h+2}}\right]. \tag{27b}
$$

To get an expression for the derivative of the value function of self h with respect to cash-on-hand in period $h+1$, appearing on the right-hand-side of (27a) notice that the familiar Envelope condition does not hold. This captures the notion that self h correctly anticipates that future self $h+1$ will deviate from the optimal consumption plan of self h. The respective derivative of the value function writes as

$$\begin{aligned}\frac{\partial V_{h+1}^{h}(\cdot)}{\partial x_{h+1}} &= \frac{du}{dc_{h+1}} m_{h+1} + \beta R \frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h}} \left(1 - m_{h+1}\right) \mathbb{E}_{h+1} \left[\frac{\partial V_{h+2}^{h}(\cdot)}{\partial x_{h+2}} \right] \\ &= m_{h+1} \underbrace{\left(\frac{du}{dc_{h+1}} - \beta R \frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h}} \mathbb{E}_{h+1} \left[\frac{\partial V_{h+2}^{h}(\cdot)}{\partial x_{h+2}} \right] \right)}_{\neq 0 \text{, i.e., the envelope condition does not hold.}} \\ &\quad + \beta R \frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h}} \mathbb{E}_{h+1} \left[\frac{\partial V_{h+2}^{h}(\cdot)}{\partial x_{h+2}} \right]\end{aligned} \tag{28}$$

where $m_{h+1} \equiv \frac{\partial c_{h+1}}{\partial x_{h+1}}$.

Collecting equations, the relevant first-order conditions of self h are (27a) and (28). Condition (27b) is a constraint to the maximization problem of self h, again because self h correctly anticipates optimality of behavior of self $h+1$.

Rewrite (27b) by adding and subtracting terms as

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

to get

$$\beta R E_{h+1} \left[\frac{\partial V_{h+2}^{h}(\cdot)}{\partial x_{h+2}} \right] = \frac{du}{dc_{h+1}} \frac{1}{\nu_{h+1,h+2}^{h+1}} + \beta R E_{h+1} \left[\Delta V_{h+2}^{h,h+1} \right], \tag{29}$$

where $\Delta V_{h+2}^{h,h+1} \equiv \left[\frac{\partial V_{h+2}^{h}(\cdot)}{\partial x_{h+2}} - \frac{\partial V_{h+2}^{h+1}(\cdot)}{\partial x_{h+2}} \right]$.

As a key difference to the quasi-hyperbolic time discounting (QHTD) model, $\Delta V_{h+2}^{h,h+1} \neq 0$ in our model because, in contrast to QHTD, the value functions of selves h and $h+1$ differ in all periods $h+1, h+2, \ldots$.

Remark 1 *Two features of the model cause this difference. First, self h forms survival beliefs with experience $e(h)$ and according ambiguity parameter δ_h whereas self $h+1$ has experience $e(h+1)$ and ambiguity δ_{h+1}, and so forth. Second, the nature of non-additivity beliefs leads to this difference.*

Next, use (29) in (28) to get

$$\begin{aligned}\frac{\partial V_{h+1}^{h}(\cdot)}{\partial x_{h+1}} &= \frac{du}{dc_{h+1}} m_{h+1} + \frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h}} \left(1 - m_{h+1}\right) \left(\frac{du}{dc_{h+1}} \frac{1}{\nu_{h+1,h+2}^{h+1}} + \beta R E_{h+1} \left[\Delta V_{h+2}^{h,h+1} \right] \right) \\ &= \frac{du}{dc_{h+1}} \left(m_{h+1} + \frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h} \nu_{h+1,h+2}^{h+1}} \left(1 - m_{h+1}\right) \right) + \beta R \frac{\nu_{h,h+2}^{h}}{\nu_{h,h+1}^{h}} (1 - m_{h+1}) \mathbb{E}_{h+1} \left[\Delta V_{h+2}^{h,h+1} \right].\end{aligned}$$

Using the above in (27a) we finally get

$$\begin{aligned}
\frac{du}{dc_h} &= \beta R \nu^h_{h,h+1} \mathbb{E}_h \left[\frac{du}{dc_{h+1}} \left(m_{h+1} + \frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1} \nu^{h+1}_{h+1,h+2}} (1 - m_{h+1}) \right) \right. \\
&\quad \left. + \beta R \frac{\nu^h_{h,h+2}}{\nu^h_{h,h+1}} (1 - m_{h+1}) \mathbb{E}_{h+1} \left[\Delta V^{h,h+1}_{h+2} \right] \right] \\
&= \beta R \nu^h_{h,h+1} \mathbb{E}_h \left[\frac{du}{dc_{h+1}} \Theta_{h+1} + \Lambda_{h+1} \right]
\end{aligned}$$

where the last line follows from the law of iterated expectations. □

D Appendix: Details on SCF Data

The Survey of Consumer Finances (SCF) is a representative triennial cross-sectional survey of U.S. families sponsored by the Federal Reserve Board in cooperation with the Department of the Treasury. We merge data from the six waves 1992, 1995, 1998, 2001, 2004 and 2007. We use households whose heads are aged 26-95. Our total sample contains 21.560 respondents.

To construct the average life-cycle profile of the—appropriately smoothed (see below)—asset-to-permanent-income ratio we proceed as follows.[22] Define assets as net worth including housing wealth, but excluding implicit pension and social security wealth. We deflate assets and income to 1992 Dollars. To approximate permanent income we first compute gross labor and social security income by excluding income from capital gains.[23] Using data from Cagetti (2003)—who approximates tax rates for different income percentiles—we next compute after-tax income. Based on the—appropriately smoothed (see below)—age-specific averages of net income we compute the net-present value and convert this to annuities using the calibrated interest rate of $r = 0.042$. This gives our permanent-income approximation. Finally, we compute the asset-to-income ratio from these two time series.

Average age-specific assets and net income are both smoothed over age by applying a cubic spline regression. We use robust fitting by three iterations of weighted least squares. Respective weights are computed from previous residuals.

[22] To construct the data we adopted the approach described in Chris Carroll's lecture notes, cf. http://www.econ2.jhu.edu/people/ccarroll. We thank Chris Carroll for providing us the Stata code.

[23] Our income measure includes 'wages and salaries', 'unemployment or worker's compensation', 'child support or alimony', 'TANF, food stamps, or other forms of welfare or assistance', 'net income from Social Security or other pensions', 'annuities, or other disability or retirement programs' and 'any other sources'. We exclude some few observations with negative income values.

Supplementary Appendix: A Three-Period Model

We provide the intuition for how ambiguous survival beliefs affect consumption and saving behavior in a simple three-period model ($T = 2$) without income risk which can be solved analytically. In this simple model we abstract from borrowing constraints, hence $a_{t+1} < 0$, for $t < T$ is possible. The no-Ponzi condition $a_{T+1} \geq 0$ is of course assumed. To simplify the analysis we assume the discount factor β to be one and an interest rate r of zero.

As shown in Section 3, lifetime utility for $T = 2$ with ambiguous survival beliefs is expressed as

$$U_0^0 = u(c_0) + \nu_{0,1}^0 \left(u(c_1) + \frac{\nu_{0,2}^0}{\nu_{0,1}^0} u(c_2) \right), \tag{30}$$

where $\nu_{k,t}^h$ is the subjective survival belief from Proposition 1. Recall that superscripts denote the respective planning age.

As in Section 3 we normalize the utility from death to zero and assume a $CRRA$ per-period utility function with preference shifter $\Upsilon \geq 0$ for the utility from survival. Since we here ignore income risk the additional Stone-Geary parameter $\bar{c}$ is not required. Also recall that lifetime utility of CEU agents reduces to the standard rational expectations case if and only if there is no initial ambiguity, i.e., iff $\delta = 0$.

We define by $x_t \equiv a_t + y_t$ cash-on-hand as the sum of financial assets a_t and income y_t. In addition, define the present value of future income, $h_t \equiv \sum_{s=t+1}^{T} y_s$, as human wealth. Finally, let total wealth be $w_t \equiv x_t + h_t$. The budget constraint is then given by

$$w_{t+1} = w_t - c_t.$$

In light of the data on subjective beliefs displayed in Figure 1 of the paper we interpret period 0 of the simple model as the period when survival probabilities are underestimated, i.e., up to actual age of about 70. Period 1 then reflects the period when there is overestimation in the data. Correspondingly, we make the following assumption:

Assumption 4. *We assume for some $\delta > 0$ that*

$$\psi_{0,1} > \nu_{0,1}^0 = \delta_0 \lambda + (1 - \delta_0) \psi_{0,1} \tag{31}$$

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

$$\psi_{1,2} < \nu_{1,2}^1 = \delta_1 \lambda + (1 - \delta_1) \psi_{1,2} \tag{32}$$

i.e., that $\lambda > \psi_{1,2}$.[24]

Consumption and Saving Plans

We now turn to the complete inter-temporal household solution to analyze how consumption and saving decisions are altered by biases in subjective survival beliefs.

Rational Expectations

The reference model is the standard solution to the rational expectations model (where $\delta_0 = \delta_1 = 0$). Here, lifetime utility does not depend on the planing period, i.e., $U_1^0 = U_1^1$. Lifetime utility in period 0 is given by $U^0 = u(c_0) + \psi_{0,1}\left(u(c_1) + \psi_{1,2}u(c_2)\right)$.

Observation 5. *Policy functions of the rational expectations solution are linear in total wealth,* $c_t = m_t w_t$, *where*

$$m_t = \begin{cases} \frac{1}{1+\frac{\left(\psi_{t,t+1}\right)^{\frac{1}{\theta}}}{m_{t+1}}} & \text{for } t < T \\ 1 & \text{for } t = T. \end{cases}$$

Hence:

$$m_0 = \frac{1}{1 + \psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}}, \qquad m_1 = \frac{1}{1 + \psi_{1,2}^{\frac{1}{\theta}}}.$$

Proof. See, e.g., Deaton (1992) □.

Naive CEU Households

To draw a distinction between RE and CEU households, we use superscript n to denote policy functions (in terms of marginal propensities to consume) of $n$aive CEU households. Given that the household consumes all outstanding wealth in the final period 2 (i.e., $m_2^n = 1$) the solution of the household's problem for all other periods are as follows:

[24] Notice that, despite equation (32), we may have that the household in period 0 underestimates the probability to survive from period 1 to 2, hence we may have that

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

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

Proposition 5. *For the naive CEU household we get:*

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

$$c_1^n = m_1^n w_1 \quad \textit{where} \quad m_1^n = \frac{1}{1 + \left(\nu_{1,2}^1\right)^{\frac{1}{\theta}}}. \tag{33}$$

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

$$c_1^{0,n} = m_1^{0,n} w_1 \quad \textit{where} \quad m_1^{0,n} = \frac{1}{1 + \left(\frac{\nu_{0,2}^0}{\nu_{0,1}^0}\right)^{\frac{1}{\theta}}}. \tag{34}$$

  where we denote the planning period as a superscript.

- *The solution in period 0 is:*

$$c_0^n = m_0^n w_0 \quad \text{where} \quad m_0^n = \frac{1}{1 + \frac{(\nu_{0,1}^0)^{\frac{1}{\theta}}}{m_1^{0,n}}} = \frac{1}{1 + (\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}}. \tag{35}$$

Proof. The first-order condition in period 1 is:

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

which directly yields (33). Analogously, the first-order condition for period 1 from the perspective of period 0 is given by:

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

which yields (34). Finally, the first-order condition in period 0 is:

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

yielding

$$m_0^n = \frac{1}{1 + \frac{(\nu_{0,1}^0)^{\frac{1}{\theta}}}{m_1^{0,n}}}.$$

Notice that

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

Using this in the above gives the last term in (35).□

Comparing the policy functions of the RE agent (cf. Observation 5) and the naive agent (cf. Proposition 5) yields the following Proposition 6 which highlights the consequences of ambiguous survival beliefs for life-cycle consumption and asset accumulation.

Proposition 6. *Comparing the naive CEU agent to the RE agent we get the following implications:*

- *There is undersaving in a sense that*

$$m_0^n > m_0 \quad \Leftrightarrow \quad c_0^n > c_0 \quad \Leftrightarrow \quad w_1^n < w_1$$

if and only if there is sufficient underestimation (SU) *of survival risk by the naive CEU agent in a sense that*

$$(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}} < \psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}. \tag{SU}$$

- *Naive CEU agents save less in period* 1 *than originally planned, i.e.,*

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

if and only if there is moderate overestimation (MO) *of survival risk in a sense that*

$$\nu_{1,2}^1 < \frac{\nu_{0,2}^0}{\nu_{0,1}^0}. \tag{MO}$$

- *There is oversaving in the sense that*

$$m_1^n < m_1 \quad \Leftrightarrow \quad \frac{c_2^n}{c_1^n} > \frac{c_2}{c_1} \quad \Leftrightarrow \quad \frac{w_2^n}{w_1^n} < \frac{w_2}{w_1},$$

by Assumption 4, equation (32) (i.e., by overestimation). Combined with condition (SU) this implies that

$$c_1^n < c_1.$$

- *Naive CEU agents have higher wealth than RE agents*

$$w_2^n > w_2 \quad \Leftrightarrow \quad c_2^n > c_2,$$

if and only if there is sufficient overestimation (SO) *of survival beliefs in period 1 in a sense that*

$$\nu_{1,2}^1 > \psi_{1,2} \left(\frac{1 + (\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}}{1 + \psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}} \cdot \frac{\psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}}{(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}} \cdot \frac{1 + (\nu_{1,2}^1)^{\frac{1}{\theta}}}{1 + \psi_{1,2}^{\frac{1}{\theta}}} \right)^{\theta} > \psi_{1,2}. \tag{SO1}$$

Proof.

- That $m_0^n > m_0$ under condition (SU) immediately follows from the expressions for the respective marginal propensities in Observation 5 and Proposition 5.
- That $m_1^{0,n} < m_1^n$ under condition (MO) immediately follows from comparing the respective marginal propensities given in Proposition 5.
- That $m_1^n < m_1$ under Assumption 4, equation (32), again immediately follows from comparison of the respective expressions for marginal propensities in Observation 5 and Proposition 5.
- By the respective expressions for marginal propensities in Observation 5 and Proposition 5, the inequality $w_2^n > w_2$ and therefore $c_2^n > c_2$ holds iff

$$\begin{aligned}
& (1-m_0^n)(1-m_1^n) > (1-m_0)(1-m_1) \\
\Leftrightarrow \quad & \frac{(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}}{1+(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}} \cdot \frac{(\nu_{1,2}^1)^{\frac{1}{\theta}}}{1+(\nu_{1,2}^1)^{\frac{1}{\theta}}} > \frac{\psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}}{1+\psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}} \frac{\psi_{1,2}^{\frac{1}{\theta}}}{1+\psi_{1,2}^{\frac{1}{\theta}}} \\
\Leftrightarrow \quad & \left(\frac{\nu_{1,2}^1}{\psi_{1,2}}\right)^{\frac{1}{\theta}} > \underbrace{\frac{1+(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}}{1+\psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}} \cdot \frac{\psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}}{(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}}}_{>1, \text{ by condition (SU)}} \cdot \underbrace{\frac{1+(\nu_{1,2}^1)^{\frac{1}{\theta}}}{1+\psi_{1,2}^{\frac{1}{\theta}}}}_{>1, \text{ by Assumption 4, equation (32)}}
\end{aligned} \tag{36}$$

  from which condition (SO1) readily follows. To see that the first term on the RHS of (36) exceeds one under condition (SU), notice that condition (SU) implies

$$\begin{aligned}
& \frac{\psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}}{(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}} > 1 \\
\Leftrightarrow \quad & 1 + \frac{1}{(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}} > 1 + \frac{1}{\psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}} \\
\Leftrightarrow \quad & \frac{\psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}}{1+\psi_{0,1}^{\frac{1}{\theta}} + \psi_{0,2}^{\frac{1}{\theta}}} \cdot \frac{1+(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}}{(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}} > 1.
\end{aligned}$$

□

As to the first implication, we require condition (SU) in order for the simple model to give rise to undersaving. Observe that condition (SU) has a straightforward interpretation for log-utility ($\theta = 1$). It can then be rewritten as

$$SLE_0 \equiv 1 + \nu_{0,1}^0 + \nu_{0,2}^0 < LE_0 \equiv 1 + \psi_{0,1} + \psi_{0,2},$$

i.e., subjective life-expectancy at birth, SLE_0, is less than the respective objective life-expectancy, LE_0. The term *sufficient underestimation (SU)* means that underestimation of subjective beliefs in period 0, cf. Assumption 4, equation (31), must be sufficiently strong in order to dominate any overestimation of subjective survival beliefs to occur eventually. Otherwise the naive CEU agent would save more than the RE agent, given forward looking behavior and her desire for consumption smoothing.

As to the second implication, we require *moderate overestimation*, cf. condition (MO), of subjective survival beliefs. In contrast to condition SU, condition MO refers to survival beliefs formed in period 1 for the probability to survive to period 2. Accordingly, it restricts $\nu^1_{1,2}$ by an upper bound which is determined by the ratio of subjective beliefs, $\frac{\nu^0_{0,2}}{\nu^0_{0,1}}$. That is, only if overestimation is not too large, we can expect model households to save less than originally planned. Otherwise the naive CEU agent would revise her plan to actually save more than originally planned.

The third implication states that our assumption on overestimation, cf. Assumption 4, equation (32), immediately gives rise to the implication that the speed of asset decumulation of the naive CEU agent is less than the speed of decumulation of the RE agent. This does not, however, imply that period 2 asset holdings of the naive CEU agent exceed those of the RE agent because the effects of sufficient underestimation in period 0 and overestimation in period 1 work in opposite directions as far as asset holdings are concerned.

This observation readily implies that an additional lower bound on the degree of overestimation is required in order to find that asset holdings in period 2 of the naive CEU agent exceed those of the RE agent. This is stated as *sufficient overestimation* in condition (SO1). Optimism has to be sufficiently strong to dominate the initial underestimation of survival beliefs. As an interpretation of the lower bound observe that the lower bound increases if the initial underestimation of survival belief gets stronger, i.e., if the gap between $\psi_{0,1}$ and $\nu^0_{0,1}$ increases.

The analysis so far clarifies that it is a quantitative question whether the calibrated life-cycle model can generate the three empirical regularities on saving behavior: (i) time inconsistent behavior to the effect that people save less than originally planned (under "moderate overestimation"); (ii) undersaving at young age (under "sufficient underestimation"); (iii) too high old-age asset holdings (under "sufficient overestimation").

Sophisticated CEU Agents

Unlike naive agents, sophisticated agents anticipate the correct lifetime utility for all future selves as additional constraints, i.e., they anticipate that their future selves will not be acting in their interest. The only way to influence future selves behavior is via the savings decision of current self 0. Thus, sophisticated agents take (over-) consumption of future selves into account when making their current saving plans.

The solution to the problem of the sophisticated CEU agent is as follows:

Proposition 7. *The solution to the sophisticated CEU agent's problem in period 0 is given by*

$$c_0^s = m_0^s w_0 = \frac{1}{1 + \frac{\left(\Theta_0\left(m_1^s\right)\cdot\nu_{0,1}^0\right)^{\frac{1}{\theta}}}{m_1^s}} w_0 \tag{37}$$

where $m_1^s = m_1^n$ *and* $m_2^s = m_2^n = 1$ *and*

$$\Theta_0\left(m_1^s\right) \equiv \left(m_1^s + \frac{\nu_{0,2}^0}{\nu_{0,1}^0\nu_{1,2}^1}(1 - m_1^s)\right).$$

$\Theta_0\left(m_1^s\right) > 1$ *under condition (MO). The solution for* m_0^s *is given by*

$$m_0^s = \frac{1}{1 + \left(1 + (\nu_{1,2}^1)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}} \left(\nu_{0,1}^0 + \nu_{0,2}^0(\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right)^{\frac{1}{\theta}}}. \tag{38}$$

Proof. In period 2 we obviously have $m_2^s = m_2^n$. The FOC of the sophisticated agent in period 1 is the same as for the period 1 naive agent, cf. the proof of Proposition 5. From this it follows that $m_1^s = m_1^n$. In period 0 the first-order condition is given by

$$\begin{aligned}
& \frac{du}{dc_0^s} = \nu_{0,1}^0\Theta_0\left(m_1^s\right)\frac{du}{dc_1^s} \\
\Leftrightarrow \quad & \frac{c_1^s}{c_0^s} = \left(\nu_{0,1}^0\Theta_0\left(m_1^s\right)\right)^{\frac{1}{\theta}} \\
\Leftrightarrow \quad & c_0^s = \left(\nu_{0,1}^0\right)^{-\frac{1}{\theta}}\left(\Theta_0\left(m_1^s\right)\right)^{-\frac{1}{\theta}}(w_0 - c_0^s)m_1^s \\
\Leftrightarrow \quad & c_0^s = \frac{1}{1 + \frac{1}{\left(\Theta_0\left(m_1^s\right)\cdot\nu_{0,1}^0\right)^{-\frac{1}{\theta}} m_1^s}} w_0
\end{aligned}$$

where

$$\Theta_0\left(m_1^s\right) \equiv m_1^s + \frac{\nu_{0,2}^0}{\nu_{0,1}^0\nu_{1,2}^1}(1 - m_1^s).$$

We obviously get that

$$\begin{aligned}
& \Theta_0\left(m_1^s\right) > 1 \\
\Leftrightarrow \quad & \frac{\nu_{0,2}^0}{\nu_{0,1}^0\nu_{1,2}^1} > 1
\end{aligned}$$

which is condition (MO).

To derive m_0^s start from

$$
\begin{aligned}
\Theta_0\left(m_1^s\right)\nu_{0,1}^0 &= \left(m_1^s + \frac{\nu_{0,2}^0}{\nu_{0,1}^0\nu_{1,2}^1}(1-m_1^s)\right)\nu_{0,1}^0 \\
&= m_1^s\nu_{0,1}^0 + \frac{\nu_{0,2}^0}{\nu_{1,2}^1}(1-m_1^s) \\
&= \frac{1}{1+(\nu_{1,2}^1)^{\frac{1}{\theta}}}\nu_{0,1}^0 + \frac{\nu_{0,2}^0}{\nu_{1,2}^1}\frac{(\nu_{1,2}^1)^{\frac{1}{\theta}}}{1+(\nu_{1,2}^1)^{\frac{1}{\theta}}} \\
&= \frac{\nu_{0,1}^0 + \nu_{0,2}^0(\nu_{1,2}^1)^{\frac{1}{\theta}-1}}{1+(\nu_{1,2}^1)^{\frac{1}{\theta}}},
\end{aligned}
$$

therefore

$$
\begin{aligned}
\left(\Theta_0\left(m_1^s\right)\nu_{0,1}^0\right)^{-\frac{1}{\theta}} m_1^s &= \left(\frac{\nu_{0,1}^0 + \nu_{0,2}^0(\nu_{1,2}^1)^{\frac{1}{\theta}-1}}{1+(\nu_{1,2}^1)^{\frac{1}{\theta}}}\right)^{-\frac{1}{\theta}} \frac{1}{1+(\nu_{1,2}^1)^{\frac{1}{\theta}}} \\
&= \frac{\left(\nu_{0,1}^0 + \nu_{0,2}^0(\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right)^{-\frac{1}{\theta}}}{\left(1+(\nu_{1,2}^1)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}}}
\end{aligned}
$$

and

$$
\begin{aligned}
1 + \frac{1}{\left(\Theta_0\left(m_1^s\right)\nu_{0,1}^0\right)^{-\frac{1}{\theta}} m_1^s} &= 1 + \frac{\left(1+(\nu_{1,2}^1)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}}}{\left(\nu_{0,1}^0 + \nu_{0,2}^0(\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right)^{-\frac{1}{\theta}}} \\
&= \frac{\left(\nu_{0,1}^0 + \nu_{0,2}^0(\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right)^{-\frac{1}{\theta}} + \left(1+(\nu_{1,2}^1)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}}}{\left(\nu_{0,1}^0 + \nu_{0,2}^0(\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right)^{-\frac{1}{\theta}}}.
\end{aligned}
$$

Using the above in (37) gives (38). □

This shows that the solution to the sophisticated agent's problem is identical to the naive agent in periods 1 and 2. This is due to the fact that the marginal propensity in period 2 is known to be $m_2^s = m_2^n = 1$ for all types. Consequently, $\Theta_1(m_2^s) = 1$ and therefore also $m_1^s = m_1^n$.

As $\Theta_0(m_1^s) > 1$ under "moderate overestimation", cf. condition (MO), we find that condition (MO) leads to a higher growth rate of marginal utilities of sophisticates compared to naifs, implying that consumption growth increases. $\Theta_0\left(m_1^s\right) > 1$ reflects the sophisticated agent's high valuation of savings. At the same time, $\Theta_0\left(m_1^s\right)$ depends negatively on the MPC of the future self 1, m_1^s, implying that self 0's propensity to save decreases when future self 1's MPC increases. Yet, these statements refer only to the change of consumption over time. The level of consumption of sophisticates in period 0 of course also depends negatively on m_1^s. The higher is m_1^s the higher will be m_0^s for reasons of consumption smoothing. Hence, whether c_0^s is lower than c_0^n depends on these offsetting forces. The next proposition makes this explicit:

Proposition 8. *Define* Ξ *as*

$$\Xi \equiv \frac{\left(1 + (\nu_{1,2}^1)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}} \left(\nu_{0,1}^0 + \nu_{0,2}^0 (\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right)^{\frac{1}{\theta}}}{(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}}$$

We have

$$\Xi \begin{cases} > 1 & \Leftrightarrow\ m_0^n > m_0^s,\ c_0^n > c_0^s,\ c_1^n < c_1^s,\ c_2^n < c_2^s \\ < 1 & \Leftrightarrow\ m_0^n < m_0^s,\ c_0^n < c_0^s,\ c_1^n > c_1^s,\ c_2^n > c_2^s \\ = 1 \quad \text{if } \theta = 1 & \Leftrightarrow\ m_0^n = m_0^s,\ c_0^n = c_1^s,\ c_1^n = c_1^s,\ c_2^n = c_2^s. \end{cases}$$

Proof. Observe from Propositions 5 and 7 that

$$\begin{aligned} & m_0^n \geq m_0^s \\ \Leftrightarrow \quad & \frac{1}{1 + (\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}} \geq \frac{1}{1 + \left(1 + (\nu_{1,2}^1)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}} \left(\nu_{0,1}^0 + \nu_{0,2}^0 (\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right)^{\frac{1}{\theta}}} \\ \Leftrightarrow \quad & \left(1 + (\nu_{1,2}^1)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}} \left(\nu_{0,1}^0 + \nu_{0,2}^0 (\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right)^{\frac{1}{\theta}} \geq (\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}} \\ \Leftrightarrow \quad & \Xi \equiv \frac{\left(1 + (\nu_{1,2}^1)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}} \left(\nu_{0,1}^0 + \nu_{0,2}^0 (\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right)^{\frac{1}{\theta}}}{(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}} \geq 1. \end{aligned}$$

We readily observe from $\Xi = 1 \Leftrightarrow m_0^s = m_0^n$ for $\theta = 1$ (in which case $\nu_{1,2}^1$ does not influence the consumption decision of sophisticates in period 0). Furthermore, if $m_0^s < m_0^n$, then $w_1^s > m_1^n$. Given that the consumption growth rate between periods 1 and 2 of the two agents is identical and given that they both consume the same present value of life-time resources, w_0, this also implies that $c_1^s > c_1^n$ and $c_2^s > c_2^n$ and vice versa for $m_0^s > m_0^n$. $\square$

It would be desirable to further make statements about how Ξ varies with θ. To approach this, observe from Proposition 5 that an increase of θ decreases m_0^n. This is due to the desire for consumption smoothing: increasing θ (decreasing the IES) increases the consumption growth rate which increases savings and reduces the marginal propensity to consume today. For the sophisticated agent the analogous effects are at work but there is an important (at least partially) offsetting one. This is easiest to see by inspection of equation (37) in Proposition 7. We assume for the remainder of the analysis that condition (MO) holds to the effect that $\Theta_0(m_1^s) > 1$. First, increasing θ reduces m_1^s. As seen from Proposition 7 this reduces m_0^s. Second, increasing θ also increases $(\nu_{0,1}^0)^{\frac{1}{\theta}}$ which also contributes to a reduction of m_0^s. Third, increasing θ also indirectly affects $\Theta_0(m_1^s)$: by decreasing m_1^s, $\Theta_0(m_1^s)$ goes up. As $\Theta_0(m_1^s) > 1$—by condition (MO)—this effect further contributes to a reduction of m_0^s. Finally, however, notice that there is a direct effect of increases in θ via term $\Theta_0(m_1^s)^{\frac{1}{\theta}}$. Under condition (MO) the derivative of this term with respect to θ is negative. It is given by $-\frac{1}{\theta^2}\ln\left(\Theta_0(m_1^s)\right)\Theta_0(m_1^s)^{\frac{1}{\theta}}$. Hence, holding $\Theta_0(m_1^s)$ constant this partially offsetting effect is particularly strong for low values of θ (and weak for high values of θ). The effect is also strong when $\nu_{1,2}^1$ is low because then m_1^s is low. We can therefore expect that $\frac{\partial\Xi}{\partial\theta} > 0$ if $\nu_{1,2}^1$ exceeds some critical value. The next sufficient condition establishes this locally at $\theta = 1$:

Proposition 9. $\left.\frac{\partial\Xi}{\partial\theta}\right|_{\theta=1} > 0$ *if*

$$\nu_{1,2}^1 > \nu_{0,1}^0 + \nu_{0,2}^0 - 1. \tag{SO2}$$

Proof. $\frac{\partial\Xi}{\partial\theta}$ has the same sign as $\frac{\partial\ln\Xi}{\partial\theta}$. We get

$$\begin{aligned}\ln\Xi &= \left(1-\frac{1}{\theta}\right)\ln\left(1+(\nu_{1,2}^1)^{\frac{1}{\theta}}\right) + \frac{1}{\theta}\ln\left(\nu_{0,1}^0 + \nu_{0,2}^0(\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right) - \ln\left((\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}\right)\\ &= \Xi_1 + \Xi_2 + \Xi_3.\end{aligned}$$

We find

$$\frac{\partial\Xi_1}{\partial\theta} = \frac{1}{\theta^2}\ln\left(1+(\nu_{1,2}^1)^{\frac{1}{\theta}}\right) - \frac{1}{\theta^2}\left(1-\frac{1}{\theta}\right)\frac{\ln(\nu_{1,2}^1)(\nu_{1,2}^1)^{\frac{1}{\theta}}}{1+(\nu_{1,2}^1)^{\frac{1}{\theta}}} > 0, \text{ if } \theta \geq 1 \tag{39}$$

$$\frac{\partial\Xi_2}{\partial\theta} = -\frac{1}{\theta^2}\ln\left(\nu_{0,1}^0 + \nu_{0,2}^0(\nu_{1,2}^1)^{\frac{1}{\theta}-1}\right) - \underbrace{\frac{1}{\theta^3}\frac{\nu_{0,2}^0\ln(\nu_{1,2}^1)(\nu_{1,2}^1)^{\frac{1}{\theta}-1}}{\nu_{0,1}^0 + \nu_{0,2}^0(\nu_{1,2}^1)^{\frac{1}{\theta}-1}}}_{<0} \tag{40}$$

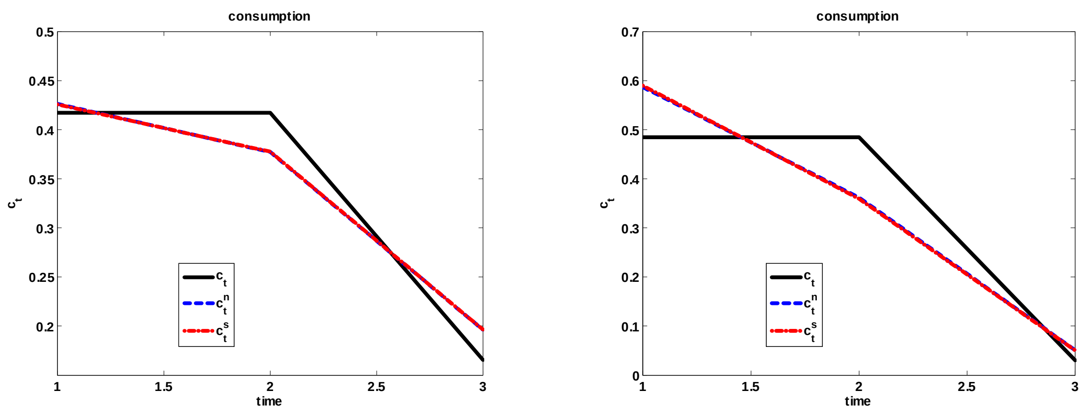
$$\frac{\partial\Xi_3}{\partial\theta} = -\frac{1}{\theta^2}\frac{\ln(\nu_{0,1}^0)(\nu_{0,1}^0)^{\frac{1}{\theta}} + \ln(\nu_{0,2}^0)(\nu_{0,2}^0)^{\frac{1}{\theta}}}{(\nu_{0,1}^0)^{\frac{1}{\theta}} + (\nu_{0,2}^0)^{\frac{1}{\theta}}} > 0. \tag{41}$$

Hence, for $\theta \geq 1$ any ambiguity in the sign of $\frac{\partial\Xi}{\partial\theta}$ can only come from the first term in the derivative (40). For $\theta = 1$, the term is obviously positive (so that the overall derivative is positive) if $\nu_{0,1}^0 + \nu_{0,2}^0 < 1$. For the other case, i.e., for $\nu_{0,1}^0 + \nu_{0,2}^0 \geq 1$, we get by comparing for $\theta = 1$ the first term in (39) with the first term in (40) the sufficient condition. $\square$

Figures 10 and 11 present an illustration for a calibrated version of the simple model. For simplicity, we consider a static model here and choose $\delta_1 = \delta_2 = 0.5$. We also set $\lambda = 0.5$, $\psi_{0,1} = 1$ and $\psi_{1,2} = 0.25$. We consider $\theta \in \{0.5, 1.5\}$. This parametrization is such that Assumption 4 holds. It also gives rise to conditions (MO), (SU), (SO1) and (SO2). The figures confirm the findings from the previous propositions. Importantly, the relative consumption patterns between RE and CEU households displayed in Figure 10 and the differences between sophisticates and naifs shown in Panel (a) of Figure 11 correspond to our findings in the quantitative model (which also features a relatively low IES), cf. Figure 9.

Figure 10: Consumption

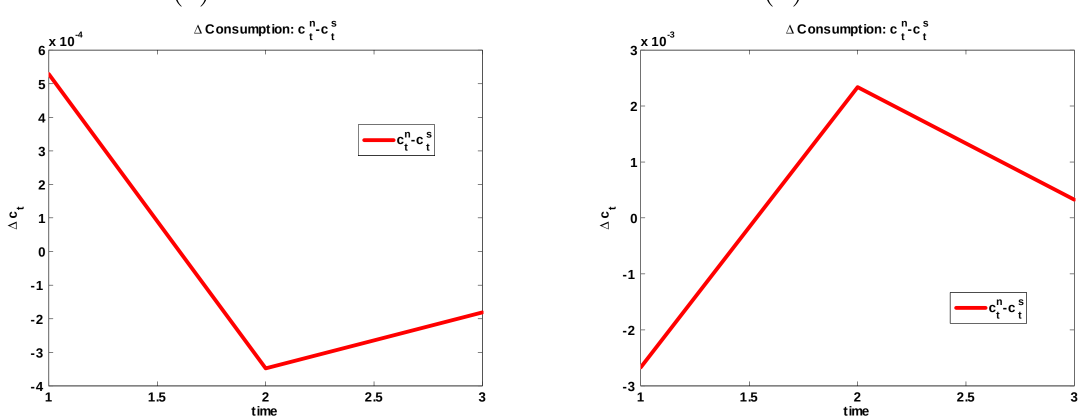
(a) $\theta = 1.5$ (b) $\theta = 0.5$



Notes: Consumption for different values of θ.

Figure 11: Difference in Consumption: Naives & Sophisticates

(a) $\theta = 1.5$ (b) $\theta = 0.5$



Notes: Difference in consumption between naifs and sophisticates for different values of θ.

References

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

HOUSE OF FINANCE

SAFE

WORKING PAPER SERIES

Recent Issues

No. 72	Alexander Ludwig, Matthias Schön	**Endogenous Grids in Higher Dimensions: Delaunay Interpolation and Hybrid Methods**
No. 71	Daniel Harenberg, Alexander Ludwig	**Social Security in an Analytically Tractable Overlapping Generations Model with Aggregate and Idiosyncratic Risk**
No. 70	Deyan Radev	**Assessing Systemic Fragility – a Probabilistic Perspective**
No. 69	Nina Biljanovska, Spyros Palligkinis	**Control Thyself: Self-Control Failure and Household Wealth**
No. 68	Tobias Tröger	**How Special Are They? – Targeting Systemic Risk by Regulating Shadow Banking**
No. 67	Sascha Baghestanian, Paul J. Gortner, Joel van der Weele	**Peer Effects and Risk Sharing in Experimental Asset Markets**
No. 66	Tobias Tröger	**Corporate Groups**
No. 65	Elia Berdin, Helmut Gründl	**The Effects of a Low Interest Rate Environment on Life Insurers**
No. 64	Daniel Herbold	**A Repeated Principal-Agent Model with On-the-Job Search**
No. 63	Nicola Fuchs-Schündeln, Michael Haliassos	**Does Product Familiarity Matter for Participation?**
No. 62	Patrick Behr, Alejandro H. Drexler, Reint Gropp, Andre Guettler	**Financial Incentives and Loan Officers Behavior: Multitasking and Allocation of Effort Under an Incomplete Contract**
No. 61	Iñaki Aldasoro, Mike Seiferling	**Vertical Fiscal Imbalances and the Accumulation of Government Debt**
No. 60	Stefano Colonnello, Giuliano Curatola, Ngoc Giang Hoang	**Executive Compensation Structure and Credit Spreads**