A Parsimonious Model of Subjective Life Expectancy*

Alexander Ludwig[†] Alexander Zimper[‡]

December 6, 2012

Abstract

On average, "young" people underestimate whereas "old" people overestimate their chances to survive into the future. Such subjective survival beliefs violate the rational expectations paradigm and are also not in line with models of rational Bayesian learning. In order to explain these empirical patterns in a parsimonious manner, we assume that self-reported beliefs express likelihood insensitivity and can therefore be modeled as non-additive beliefs. In a next step we introduce a closed form model of Bayesian learning for non-additive beliefs which combines rational learning with psychological attitudes in the interpretation of information. Our model gives a remarkable fit to average subjective survival beliefs reported in the Health and Retirement Study.

JEL classification: C44, D83, D91, I10

Keywords: representative agent; subjective survival expectations; likelihood insensitivity; Choquet decision theory; Bayesian learning

*We thank Axel Börsch-Supan, Andrew Caplin, Alain Chateauneuf, Itzhak Gilboa, Mike Hurd, David Levine, Jürgen Maurer, Dan McFadden, Klaus Ritzberger, Susan Rohwedder, Aylit Romm, Martin Salm, Frank Sloan, Carsten Trenkler, Peter Wakker, David Weir, Joachim Winter and an anonymous referee as well as several seminar participants at the Workshop on Subjective Probabilities and Expectations, Jackson Hole, Wyoming, 2007, at Universität Mannheim and the 2009 SAET Conference for helpful comments and Frank Schilbach for his excellent research assistance. Financial support by the Social Security Administration (SSA) under the 2007 Steven H. Sandell Grant Program is gratefully acknowledged. Alex Ludwig further gratefully acknowledges financial support from the German National Research Foundation (DFG) through SFB 504, the State of Baden-Württemberg and the German Insurers Association (GDV). Alex Zimper is grateful for financial support from Economic Research Southern Africa (ERSA).

†CMR, University of Cologne; Albertus-Magnus-Platz; 50923 Cologne; Germany; Email: ludwig@wiso.uni-koeln.de.

‡Department of Economics; University of Pretoria; Private Bag X20; Hatfield 0028; South Africa; Email: alexander.zimper@up.ac.za.

1 Introduction

Dynamic economic models are based on forward looking behavior of economic agents. In the context of life-cycle models, an individual's consumption and savings decision depends on her subjective beliefs about future interest rates, wage rates and the likelihood of dying. According to these models, individuals have beliefs about such variables and use these beliefs to make decisions today. Until recently, common practice in such studies was to assume rational expectations implying that individuals' beliefs are given as objective probability distributions. The use of objective distributions is by now put into question by numerous researchers who suggest to directly measure subjective expectations. Manski (2004) provides an overview on this literature and McFadden et al. (2005) discuss various survey design issues related to questions about subjective expectations.

This paper focuses on the formation of subjective survival expectations. As point of departure, we present stylized facts on a comparison between average subjective survival expectations from the Health and Retirement Study (HRS) and their objective counterparts. These facts can be summarized as follows: First, on average, individuals of relatively "young" age underestimate survival probabilities. Second, this "pessimistic" bias monotonically decreases with age to zero for respondents of about age 70. Third, "old" respondents overestimate their actual survival probability.

Finally, this "optimistic" bias monotonically increases with age.[1]

We argue that these stylized facts are incompatible with the rational expectations paradigm. Furthermore, the observed age-dependent biases in the data also suggest a violation of the rational Bayesian learning paradigm. Models of subjective belief formation based on rational Bayesian learning generate posterior beliefs that are closer to the true, i.e., objective, distribution the more experienced the agent becomes. Under our maintained assumption that an agent receives more survival-rates related information by getting older, rational Bayesian learning requires the agent to learn with increasing age the true probabilities, cf. Viscusi (1985, 1990, 1991). Under the assumption of rational Bayesian learning any gap between subjective beliefs and objective survival probabilities should therefore decrease in age. This contradicts the data.

Our main contribution is the introduction of a closed-form model of Bayesian learning of survival expectations which extends the standard model rational Bayesian learning by psychological biases. This enables us to match the empirical facts. The model is based on the framework of Choquet decision theory Schmeidler (1986, 1989) and Gilboa (1987). While Choquet decision theory has been originally developed to model ambiguity attitudes as expressed in Ellsberg paradoxes (Ellsberg 1961), our approach demonstrates the usefulness of Choquet decision theory under the assump-

[1]Our findings confirm results documented in a vast literature on subjective survival probabilities, see below.

tion that a representative agent reports survival beliefs which are prone to *likelihood insensitivity*. Wakker (2010) refers to likelihood insensitivity as a new psychological concept which "reflects diminishing sensitivity for a scale bounded from two sides" (p. 227). According to this cognitive interpretation, agents do not sufficiently distinguish between probabilities that lie between zero and one when they report their probabilistic beliefs. Our approach is therefore in line with the psychological literature that documents self-reported probabilities which can be described as non-additive probability measures (cf. Wu and Gonzalez (1996), Wakker (2004, 2010)). An extreme expression of likelihood insensitivity are "fifty-fifty" probability judgements (Bruine de Bruin, Fischhoff, and Halpern-Felsher, Bonnie Millstein 2000) which can be described by the non-additive probability measure that assigns probability 0.5 to every uncertain event.

More precisely, we formally describe self-reported survival beliefs as neo-additive capacities in the sense of Chateauneuf, Eichberger, and Grant (2007). Neo-additive capacities are non-additive probability measures that give rise to a linear transformation of (non-extreme) probabilities such that the probability weighting function for probabilities between zero and one is flatter than the 45-degree line. As a consequence, neo-additive capacities allow for a parsimonious formalization of concepts such as likelihood insensitivity or ambiguity attitudes within Choquet decision theory.

We assume that these neo-additive beliefs are updated in accordance with the generalized Bayesian update rule (Eichberger, Grant, and Kelsey 2007). For the representative agent of our model an initial bias between her subjective beliefs and objective probabilities does thereby not necessarily vanish in the long run. Our formal approach hence accommodates attitudes like "myside bias" or "irrational belief persistence" as documented in the psychological literature (Baron 2008, Ch. 9).

Our theoretical framework provides a parsimonious specification of the representative agent's age-belief pattern with three parameters, reflecting, first, an initial bias in subjective survival probabilities, second, a measure for the agent's likelihood insensitivity with respect to her initial estimator of her subjective survival probability, and, third, an optimism, resp. pessimism, parameter. We estimate these three parameters by pooling HRS data. Despite the low parametrization, our model results in a decent fit to average data on subjective beliefs.

Our work contributes to a vast empirical literature on subjective survival probabilities that was initiated by Hammermesh (1985). In two different data samples from surveys, Hammermesh (1985) found that people do incorporate improvements of life-expectancy into their beliefs about personal longevity and that subjective survival curves are somewhat flatter than objective data. Similar differences between subjective beliefs and objective data have been reported for the HRS by Hurd and McGarry (1995) and Gan, Hurd, and McFadden (2005) and others and, more recently, for the Survey of Health, Ageing and Retirement in Europe (SHARE) data (Hurd, Rohwedder, and Winter 2005) as well as, for direct questions on remaining life expectancy rather than probabilities, for the German SAVE data (Steffen 2009).

How subjective survival expectations respond to new information—health shocks and other candidate predictors of own mortality such as parental death—and whether these expectations serve as predictors of actual mortality is analyzed in the empirical work by, among others, Hurd, McFadden, and Merrill (2001), Hurd and McGarry (2002), Smith, Taylor, and Sloan (2001) and—applying the framework of Viscusi (1985)—Smith, Taylor, Sloan, Johnson, and Desvouges (2001). A general conclusion from this literature is that subjective survival expectations have predictive power for the respondent's own demise and that they are consistently updated with new health information.

However, as summarized by Smith, Taylor, and Sloan (2001), subjective survival expectations "do not appear to reflect all of the information that respondents who subsequently die know about their survival prospects". This finding comes as no surprise as we argue that (updating of) subjective beliefs cannot be fully explained by use of objective data in situations where an individual's learning process may be prone to emotions such as hope or despair. This is particularly true when an individual learns new information about her life expectancy thereby facing the prospect of her own death. Along this line, Kastenbaum (2000) 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." This also suggests that psychological biases in subjective survival expectations may vary with age because cognitive encounters with death are stronger when the risk becomes more relevant, just as in our model.

The remainder of our analysis is structured as follows. Section 2 documents stylized facts in the HRS data. Section 3 develops our parsimonious model of subjective life expectancy. We then present the main results of our empirical analysis in Section 4. Finally, Section 5 concludes. A separate supplementary online appendix[2] contains decision theoretic foundations of our model, provides a detailed description of our data and reports additional results.

2 Stylized Facts

We compare subjective survival beliefs, based on data of the Health and Retirement Study (HRS), with objective survival rates. The data contain information about individuals' expectations to live from age at interview j up to some target age m. Age at interview j and target age m are assigned according to the pattern in Table 1. Objective survival rates are based on cohort life tables for the U.S. population. A detailed description of our data sources and methods is provided in the online appendix.

[2] Available at www.wiso.uni-koeln.de/aspsamp/cmr/alexludwig/downloads/ParsiApp.pdf.

The following section only provides a brief summary.

Table 1: Interview and Target Age

Age at Interview j	Target Age m
$\leq$69	80
70-74	85
75-79	90
80-84	95
85-89	100

Source: RAND HRS Data Documentation, Version F (October 2006).

2.1 HRS Data

In the HRS, respondents of waves 5 through 7 were asked in the respective interview years 2000, 2002 and 2004 about their probability to live from interview age j until a certain target age m, cf. Table 1. In our analysis, we pool the information in these three waves. As we discuss in the online appendix, we do not consider households of age $40-49$ and of age 90 and older. In addition we exclude some observations with inconsistent answering patterns. This selection by age and consistency of answering patterns leaves us with a total sample size of 44671 observations out of which 18341 are male and 26330 are female respondents. While most of our analysis focusses on this "full sample", we further investigate sensitivity of our results with respect to focal point answers at subjective survival probabilities of 0, 50, and 100 percent in Subsection 4.3.

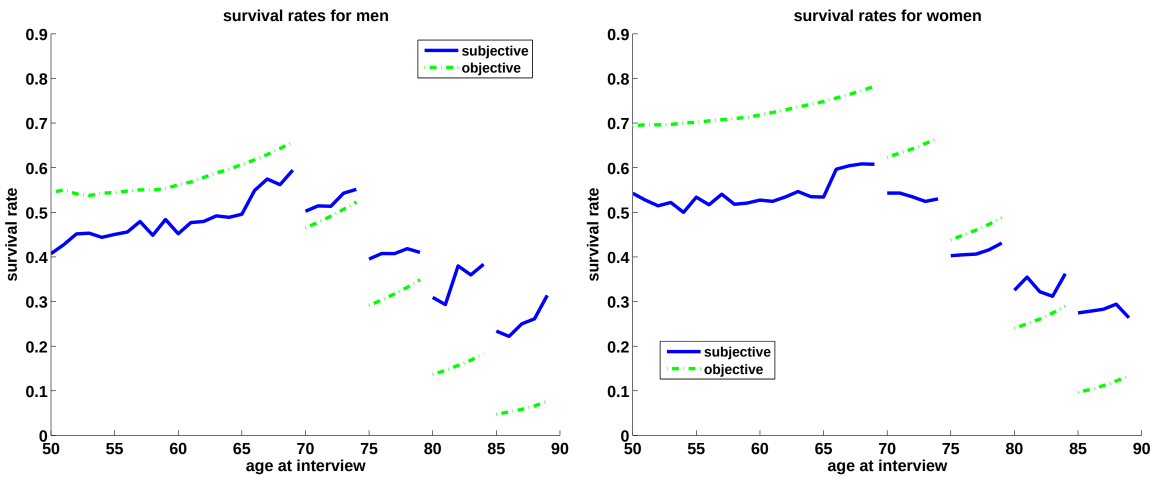
We next construct objective survival rates. In correspondence with our representative agent model that we develop in Section 3, we follow the literature initiated by Hammermesh (1985) and use cohort life tables for the entire U.S. population as objective data. To construct those we predict future survival rates in the population. Estimates are based on data for age-specific survival rates for years 1900 to 2004 taken from the Human Mortality Database (2008) (HMD) and the Social Security Administration (SSA). Since projections from official sources tend to underestimate future increases in survival probabilities, we do not use SSA cohort life tables but rather base the prediction of future survival rates on a Lee-Carter procedure (Lee and Carter 1992). The idea of our approach is that agents in our model base their predictions of their respective objective survival probabilities on past data but it is unobserved to the econometrician which point estimates they use. For this reason we account for uncertainty of objective data in the estimation of standard errors, cf. Section 4. As an additional advantage, our procedure assigns objective information on survival rates in correspondence with the HRS interview years.

2.2 Illustration

Figure 1 summarizes information in our data by displaying average subjective beliefs on survival of HRS respondents against age at interview and the respective objective data for men in panel (a) and women in panel (b). The different line segments are due to changes in target ages, cf. Table 1. Two stylized facts emerge for either gender from the data. First, subjective beliefs on survival are downward biased at younger ages. Second, subjective beliefs on survival are upward biased at older ages whereby the upward bias increases with age. These stylized facts clearly indicate a systematic violation of the rational expectations paradigm of economic theory by which there should be no difference between subjective beliefs and objective survival rates.

Figure 1: Subjective and objective survival probabilities

(a) Men (b) Women



Source: Own calculations based on HRS, HMD and SSA data.

For younger respondents (≤ 69) the data in Figure 1 is compatible with the convergence behavior as predicted by rational Bayesian learning.[3] However, upon inspection of the age-belief pattern of elderly respondents of age 75 and older in Figure 1, the picture changes. In Figure 2 we zoom in from figure 1 average beliefs of male respondents between interview ages 80 to 89 to survive until 95, respectively until 100, against their objective counterparts in panel (a). To illustrate learning behavior in this age

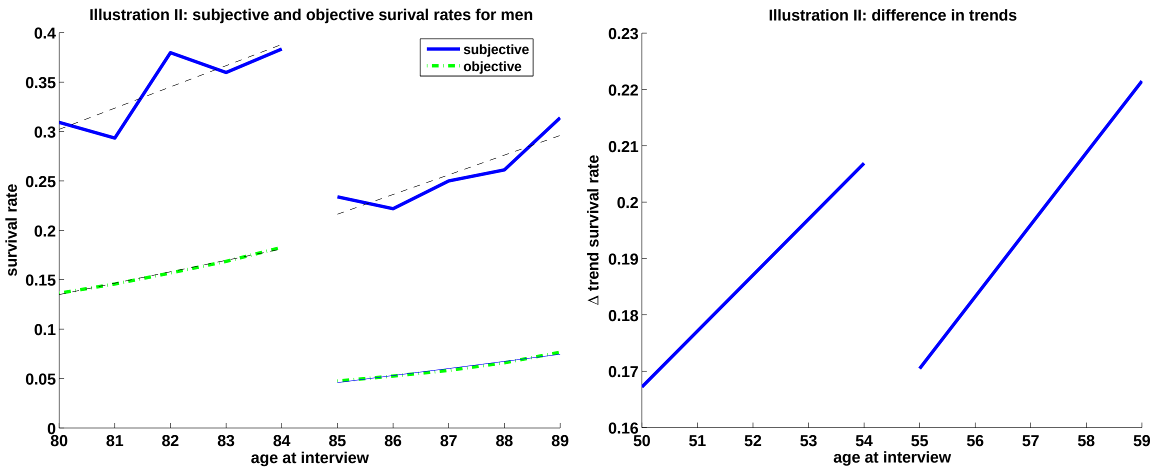
[3]For women we do not observe such a clear convergent pattern even for this age group.

group we estimated simple linear trends for both the subjective and the objective data and display the differences in these trends in panel (b) of the same figure. These graphs indicate divergence with increasing age. This divergent pattern is stronger for the higher interview/target age group. Thus, contrary to the predictions of the rational Bayesian learning model the average bias between subjective beliefs and objective probabilities increases rather than decreases with more experience whereby this effect appears to be stronger for higher target ages.

The patterns shown in Figure 2 illustrate a violation of the rational Bayesian learning paradigm within target age groups. Furthermore, in order to explain the data across target age groups, the rational Bayesian learning hypothesis would require highly implausible prior beliefs. For example, the overestimation of the subjective belief of an 80 year old agent to live until 95 by 17.28 percentage points for men (8.54 percentage points for women), cf. Figure 1, can only be explained with rational Bayesian learning if the same agent expressed a prior belief with a much higher degree of overestimation about her survival at age 50. However, at age 50, we actually observe an average underestimation of the survival belief by −13.70 percentage points for men (−15.07 percentage points for women). Although these are not the same agents, such differences across cohorts appear implausible. We further document in the online appendix that cohort effects are indeed not relevant.

Figure 2: Survival probabilities at age 80 and older for men

(a) Men (b) Older Men



Notes: The dashed lines in panel (a) are predicted values from a simple linear trend estimation. Panel (b) displays the difference in these linear trends.
Source: Own calculations based on HRS, HMD and SSA data.

3 A Parsimonious Model of Subjective Life Expectancy

Our closed form model of Bayesian learning with psychological bias captures the stylized facts of Figure 1 in a very parsimonious way. It also offers a plausible explanation why young people are too pessimistic whereas elderly people are too optimistic about their survival expectations. The formal approach shares similarities with Zimper and Ludwig (2009) and Zimper (2009). In contrast to this earlier work, however, our theoretical model is tailored to the formation of subjective survival beliefs. The statistic and decision theoretic foundations of our theory are provided in the online appendix.

The key building blocks are as follows: we merge a standard rational Bayesian updating model—described in Subsection 3.1—with a Choquet model of updating non-additive beliefs—described in Subsection 3.2. A number of additional plausible assumptions—laid out in Subsection 3.3—then lead to a parsimonious specification of the representative agent's age-belief pattern of survival beliefs with three parameters only. As we discuss in our psychological interpretation of our model—given in Subsection 3.4—, these parameters reflect (i) an initial bias, ϕ, in the additive estimator reflecting overestimation, i.e., $\phi > 1$, or underestimation, i.e., $\phi < 1$, (ii) a measure for likelihood insensitivity, δ, and (iii) the degree of optimism, respectively pessimism, by which the agent resolves her likelihood insensitivity, λ.

3.1 Rational Bayesian Learning

We describe a closed-form learning model with additive beliefs as introduced to the economics literature by Viscusi and O'Connor (1984) and Viscusi (1985). A number of formal definitions of the statistic environment of this model, in particular the additive probability space $(\mu, \Omega, \mathcal{F})$, as well as analytical derivations are given in the online appendix.

Consider the situation of an agent who is uncertain about her likelihood to survive until age m given that she has reached age j. Further, suppose that the agent could observe for n individuals with i.i.d. survival probability whether any individual survived from age j until m or not whereby we allow for the possibility that the sample size n may become arbitrarily large. We denote by $\mathbf{I}_n^k$ the information of the agent, i.e., the event that k out of n individuals have survived until age m. Define the $\mathcal{F}$-measurable random variable $\tilde{\pi}_{j,m} : \Omega \to [0, 1]$ such that $\tilde{\pi}_{j,m}(s, \pi_{j,m}) = \pi_{j,m}$ where s denotes sample-information and $\pi_{j,m} \in [0, 1]$ denotes the true objective survival probability.

For notational convenience, we now drop subscripts and, e.g., simply write $\tilde{\pi}$ instead of $\tilde{\pi}_{j,m}$. Subscripts will be reintroduced below. The agent's (subjective) prior estimate of the true survival probability is then defined as the expected value of $\tilde{\pi}$ with respect to a prior distribution μ, i.e., $E[\tilde{\pi}, \mu]$. Accordingly, the agent's posterior estimate of π conditional on information $\mathbf{I}_n^k$ is defined as the expected value of $\tilde{\pi}$ with respect to the resulting posterior distribution, i.e., $E\left[\tilde{\pi}, \mu\left(\cdot \mid \mathbf{I}_n^k\right)\right]$. We assume that the agent's prior over $\tilde{\pi}$ is given as a Beta distribution with parameters $\alpha, \beta > 0$. We accordingly obtain as prior estimate the expected value of the Beta distribution, implying $E[\tilde{\pi}, \mu] = \frac{\alpha}{\alpha+\beta}$. Furthermore, note that the agent's posterior $\mu\left(\cdot \mid \mathbf{I}_n^k\right)$ is itself a Beta distribution with parameters $\alpha + k, \beta + n - k$.[4] We therefore have $E\left[\tilde{\pi}, \mu\left(\cdot \mid \mathbf{I}_n^k\right)\right] = \frac{\alpha+k}{\alpha+\beta+n}$ as the agent's posterior estimate of π conditional on information $\mathbf{I}_n^k$. Or, equivalently,

$$E\left[\tilde{\pi}, \mu\left(\cdot \mid \mathbf{I}_n^k\right)\right] = \left(\frac{\alpha + \beta}{\alpha + \beta + n}\right) E[\tilde{\pi}, \mu] + \left(\frac{n}{\alpha + \beta + n}\right) \frac{k}{n} \tag{1}$$

where $\frac{k}{n}$ is the sample mean. That is, the agent's posterior estimate of the probability

[4]That is, the prior and the posterior are *conjugate* distributions because both belong to the Beta distribution family.

that an individual of age j survives until age m is a weighted average of her prior estimate and the observed sample mean. The weight attached to the sample mean increases in the sample size, also see below.[5]

As an additional object required below in Proposition 1 we have to specify the unconditional probability of receiving information $\mathbf{I}_n^k$. As shown in the online appendix it is given by

$$\mu\left(\mathbf{I}_n^k\right) = \binom{n}{k} \frac{(\alpha + k - 1) \cdot \ldots \cdot \alpha \cdot (\beta + n - k - 1) \cdot \ldots \cdot \beta}{(\alpha + \beta + n - 1) \cdot \ldots \cdot (\alpha + \beta)}. \tag{2}$$

3.2 Bayesian Learning with Psychological Bias

We develop our concept of Bayesian learning with psychological biases as a generalization of the rational learning model discussed above. We assume that reported beliefs express likelihood insensitivity in the sense of Wakker (2010). Such individuals can be described as Choquet decision makers so that the expectation of a random variable with respect to a non-additive probability measure is given as its Choquet expected value (Schmeidler 1986).[6]

[5]Tonks (1983) introduces a similar model of rational Bayesian learning in which the agent has a normally distributed prior over the mean of a normal distribution and receives normally distributed information.

[6]Besides the formal description of insensitivity to changes in likelihood (Wakker 2004), properties of non-additive beliefs are used in the literature for formal definitions of, e.g., ambiguity and uncertainty attitudes (Schmeidler 1989; Epstein 1999; Ghirardato and Marinacci 2002), as well as pessimism and optimism (Eichberger and Kelsey 1999; Wakker 2001; Chateauneuf, Eichberger, and Grant 2007). See Wakker (2010) for a textbook treatment.

3.2.1 Neo-additive Beliefs

Our approach focuses on non-additive probability measures that are defined as neo-additive capacities in the sense of Chateauneuf, Eichberger, and Grant (2007). According to the standard—motivational—interpretation, neo-additive beliefs stand for a deviation from additive beliefs such that a parameter δ measures the lack of confidence the decision maker has in a subjective additive probability distribution μ. The optimism parameter λ measures how much weight the decision maker puts on the best possible outcome of an alternative. According to an alternative—the cognitive—interpretation, δ is an index of likelihood insensitivity. Factor $(2\delta\lambda + (1-\delta))/2$ is referred to as index of elevation (cf. Equ 7.2.6 in Wakker 2010). In the remainder of our analysis we simply refer to δ as index of likelihood insensitivity and to λ as optimism parameter.

We consider the case of uncertain survival in which the agent is a Choquet decision maker so that her prior survival belief about π is given by a neo-additive capacity ν whose additive part is described by a Beta-distribution μ. Her (prior) estimator is then the Choquet expected value of $\tilde{\pi}$, $\tilde{\pi}(\Omega) \in [0,1]$, with respect to the non-additive prior ν:

$$E[\tilde{\pi}, \nu] = \delta \cdot (\lambda \cdot \max \tilde{\pi} + (1-\lambda) \cdot \min \tilde{\pi}) + (1-\delta) \cdot E[\tilde{\pi}, \mu].$$

As the best, respectively worst, possible outcome is survival, respectively death, we have that $\max \tilde{\pi} = 1$, respectively $\min \tilde{\pi} = 0$. Consequently, the above equation simplifies to

$$E[\tilde{\pi}, \nu] = \delta \cdot \lambda + (1 - \delta) \cdot E[\tilde{\pi}, \mu]. \tag{3}$$

3.2.2 Updating of Neo-additive Beliefs

Under the assumption that the agent updates (3) in light of new information $\mathbf{I}_n^k$, her posterior belief is given by the conditional neo-additive probability measure $\nu\left(\cdot \mid \mathbf{I}_n^k\right)$ so that the (posterior) estimate of π becomes $E\left[\tilde{\pi}, \nu\left(\cdot \mid \mathbf{I}_n^k\right)\right]$.

At this point we have to take a stand on how an agent updates her non-additive beliefs. Several different Bayesian update rules are perceivable for the non-additive beliefs of CEU decision-makers (Gilboa and Schmeidler 1993; Sarin and Wakker 1998; Pires 2002; Eichberger, Grant, and Kelsey 2007; Siniscalchi 2006). In this paper we consider the so-called Generalized Bayesian update rule which is formally defined in the online appendix[7].

Applied to survival beliefs we obtain the following neo-additive (posterior) estimator that the agent of age j will be alive at age m given her information $\mathbf{I}_n^k$

$$E\left[\tilde{\pi}, \nu\left(\cdot \mid \mathbf{I}_n^k\right)\right] = \delta_{\mathbf{I}_n^k} \cdot \lambda + \left(1 - \delta_{\mathbf{I}_n^k}\right) \cdot E\left[\tilde{\pi}, \mu\left(\cdot \mid \mathbf{I}_n^k\right)\right] \tag{4}$$

[7] An axiomatic foundation under the assumption of CEU preferences is provided in Eichberger, Grant, and Kelsey (2007, 2010)

where

$$\delta_{\mathbf{I}_n^k} = \frac{\delta}{\delta + (1-\delta)\cdot\mu\left(\mathbf{I}_n^k\right)}. \tag{5}$$

Finally, we link information received by the agent to her age. We suppose that an agent of age h receives information $\mathbf{I}_{n(h)}^k$. This is equivalent to information gained from a statistical experiment with $n(h)$ trials whereby the *experience* function $n(h)$ satisfies $n(0) = 0$, $n(h) < n(h+1)$ for all h and $n(h) \to \infty$ if $h \to \infty$. That is, our approach associates a higher age with greater experience whereby we do not restrict gaining of experience by any upper bound.[8]

3.3 A Parsimonious Model

Collecting equations (1)–(5), the following proposition summarizes the considerations from above.

Proposition 1. *Under the assumption of Bayesian learning with psychological bias, the posterior belief of an agent of age h to survive from age j to age m conditional on information $\mathbf{I}_{n(h)}^k$ is given by*

$$E\left[\tilde{\pi}, \nu\left(\cdot \mid \mathbf{I}_{n(h)}^k\right)\right] = \delta_{\mathbf{I}_{n(h)}^k}\cdot\lambda + \left(1-\delta_{\mathbf{I}_{n(h)}^k}\right)\cdot E\left[\tilde{\pi}, \mu\left(\cdot \mid \mathbf{I}_{n(h)}^k\right)\right],$$

whereby

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

with

$$\mu\left(\mathbf{I}_{n(h)}^k\right) = \binom{n(h)}{k}\frac{(\alpha+k-1)\cdot\ldots\cdot\alpha\cdot(\beta+n(h)-k-1)\cdot\ldots\cdot\beta}{(\alpha+\beta+n(h)-1)\cdot\ldots\cdot(\alpha+\beta)}$$

[8]Sensitivity analysis with respect to this assumption is presented in the online appendix.

and

$$E\left[\tilde{\pi},\mu\left(\cdot\mid \mathbf{I}^{k}_{n(h)}\right)\right]=\left(\frac{\alpha+\beta}{\alpha+\beta+n\left(h\right)}\right)E\left[\tilde{\pi},\mu\right]+\left(\frac{n\left(h\right)}{\alpha+\beta+n\left(h\right)}\right)\frac{k}{n\left(h\right)},$$

where $E\left[\tilde{\pi},\mu\right]$ is the agent's prior additive estimate of the conditional survival probability. $\frac{k}{n(h)}$ stands for the observed sample mean of individuals who have survived from age j to age m.

Finally, we develop a highly simplified version of our model of Bayesian learning with psychological bias that we bring to the data on survival beliefs in Section 4. To this end, we impose four technically convenient assumptions:

Assumption 1. *The representative agent has a uniform prior distribution over the parameter π. That is, $\alpha=\beta=1$, implying for (2)*

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

Assumption 2. *We suppose that the representative agent observes at every age sample means that actually coincide with objective survival rates. That is, for all h, $\frac{k}{n(h)}$ coincides with the true survival probability $\pi^{*}_{j,m}$ to live from age j to m, whereby we re-introduce the subscript notation that had been dropped in the previous two subsections.*

Assumption 3. *We restrict ourselves to an experience function $n\left(h\right)=h$ whereby we assume that agents start learning at the age of 20 which corresponds to $h=1$ in our model.*

Assumption 4. *We initialize $E\left[\tilde{\pi}_{r,r+1},\mu\right]$ for all ages $r=j,\ldots,m-1$ as*

$$E\left[\tilde{\pi}_{r,r+1},\mu\right]=\phi\pi^{*}_{r,r+1}.$$

Using Assumptions 1 through 4 in Proposition 1 and setting $b_{j,m}^{h} \equiv E\left[\tilde{\pi}_{j,m}, \nu\left(\cdot \mid h\right)\right]$ we can summarize our parameterized Choquet model of subjective life expectancy as follows:

Proposition 2. *Let $h \leq j < m$ and suppose that Assumptions 1-4 hold. Then the posterior belief of an agent of age h to survive from age j to age m is*

$$b_{j,m}^{h} = \delta_h \cdot \lambda + (1 - \delta_h) \cdot \tilde{b}_{j,m}^{h}, \text{ where } \tilde{b}_{j,m}^{h} = \left(\frac{2\phi^{m-j} + h}{2 + h}\right) \pi_{j,m}^{*} \tag{6}$$

whereby

$$\delta_h = \frac{\delta}{\delta + (1 - \delta)\frac{1}{1+h}} \tag{7}$$

and

$$\prod_{r=j}^{m-1} \phi \pi_{r,r+1}^{*} = \phi^{m-j} \pi_{j,m}^{*}. \tag{8}$$

In contrast to Proposition 1, which characterizes our concept of Bayesian learning with respect to any possibly observable information $\mathbf{I}_{n(h)}^{k}$, Proposition 2 presents a model of Bayesian learning for a representative agent of age h that aggregates over all possible observations $\mathbf{I}_{n(h)}^{k}$, with $k = 0, ..., n\left(h\right)$, to the effect that Bayesian learning becomes now conditional on the agent's age only. In moving from the information-conditional learning model of Proposition 1 to the age-conditional learning model of Proposition 2, we have clearly chosen technically convenient over realistic assumptions. Nevertheless, we are going to argue in this paper that the learning model of Proposition 2 is—in spite of its limited realistic appeal—empirically superior compared to the standard model of rational Bayesian learning. In the remainder of this subsection we critically discuss Assumptions 1-4 in more detail.

Under Assumption 1 the impact of received information is independent of the observed k and depends only on the number of observations $n(h)$. The assumption is quite artificial. To see, however, that the uniform distribution assumption is technically very convenient, observe that it implies a constant value of the likelihood insensitivity parameter $\delta_{\mathbf{I}^k_{n(h)}}$ across all possible observations $\mathbf{I}^k_{n(h)}$, with $k = 0, ..., n(h)$. Consequently, by Assumption 1 we trivially obtain for the "aggregated" likelihood insensitivity parameter that

$$\begin{aligned} E\left[\delta_{\mathbf{I}^k_{n(h)}}, \mu\left(\mathbf{I}^k_{n(h)}\right)\right] &= \sum_{k=0}^{n(h)} \frac{\delta}{\delta + (1-\delta) \cdot \mu\left(\mathbf{I}^k_{n(h)}\right)} \cdot \mu\left(\mathbf{I}^k_{n(h)}\right) \\ &= \frac{\delta}{\delta + (1-\delta)\frac{1}{1+h}} \equiv \delta_h. \end{aligned}$$

That is, Assumption 1 greatly simplifies the aggregation problem over different survival information that an agent of age h can possibly observe.

Assumption 2 is, by the law of large numbers, appealing for a large number of observations. We regard it as innocuous. Notice that the assumption also implies that $k = \pi^*_{j,m} n(h)$.

Normalization of initial age in Assumption 3 corresponds with many life-cycle models of consumption and savings where agents are assumed to become economically active at age 20. Our findings are robust with respect to this normalization.[9] The structure of the experience function specified in Assumption 3, however, requires more discussion. We stipulate that the amount of survival information strictly

[9] Results are available upon request.

increases in an individual's age. Our first argument in favor of Assumption 3 is pragmatic: To model an information flow in terms of a filtration process, i.e., no memory losses, is standard in the literature. Second, our assumption captures the notion that older individuals have more information about their own demise because they observe peer groups dying as well as have experienced own health histories. However, our assumption does not account for the empirical fact that memory-losses become increasingly common for older age-groups. To accommodate this, we investigate sensitivity of our estimation results with respect to the experience function, see our online appendix. There we consider faster accumulation of experience in a linear way, a square-root experience function (capturing decreasing marginal experience), and an inverse u-shaped experience function with a peak at real-life age 65 (which represents a reduced form model of decreasing marginal experience combined with an eventually dominating depreciation of memory). Broadly speaking, our findings are robust to these alternative choices.[10]

Assumption 4 implies that belief $E\left[\tilde{\pi}_{j,m},\mu\right]$ for all pairs (j,m) is given by $E\left[\tilde{\pi}_{j,m},\mu\right] = \prod_{r=j}^{m-1}\phi\pi^*_{r,r+1} = \phi^{m-j}\pi^*_{j,m}$. The assumed age-independent difference between the (additive) subjective expectation of the representative agent's survival probability and the true average survival probability keeps the model as mathematically simple as possible.

[10]A qualitative difference is that inverse u-shaped experience translates into inverse u-shaped likelihood sensitivity.

3.4 Psychological Interpretation

If there is no likelihood insensitivity in the agent's beliefs, i.e., $\delta = 0$, our model reduces to a version of rational Bayesian learning such that

$$b_{j,m}^h = \left(\frac{2\phi^{m-j} + h}{2 + h}\right) \pi_{j,m}^*.$$

We occasionally refer to term $\frac{2\phi^{m-j}+h}{2+h}$ as a bias factor. In line with standard results on consistency of additive Bayesian estimators, in particular Doob's consistency theorem (Doob (1949), Breiman, LeCam, and Schwartz (1964), Lijoi, Pruenster, and Walker (2004)), the subjective belief of a rational agent therefore converges to the objective probability $\pi_{j,m}^*$ when her actual age h—and thereby the amount of gathered information—increases. Depending on an initial overestimation ($\phi > 1$), resp. underestimation ($\phi < 1$), subjective beliefs monotonically converge from "above", respectively "below". This convergence behavior is the same for all target ages. Such a model of rational Bayesian learning can obviously not accommodate the stylized facts of Figure 1, showing strong underestimation for a lower target age, e.g., $m = 80$, and strong overestimation for a higher target age, e.g., $m = 95$. In order to accommodate these stylized facts by rational Bayesian learning alone, an adequate model would require target-age specific parameters ϕ_m such that, e.g., $\phi_{85} < 1$ and $\phi_{95} > 1$. Such an extension would come at the cost of losing parsimony without offering a straightforward interpretation of the additional parameters. In our opinion, it is therefore highly implausible that the HRS data may reflect rational Bayesian learning alone.

If, in contrast, there is some likelihood insensitivity involved, i.e., $\delta > 0$, the resulting non-additive Bayesian estimator is no longer consistent but rather biased. Now there exist two different long-run effects in our model of Bayesian learning. On the one hand, the additive part of the agent's belief, i.e., $\left(\frac{2\phi^{m-j}+h}{2+h}\right)\pi^*_{j,m}$, still converges with increasing age h towards the objective probability $\pi^*_{j,m}$. On the other hand, however, the impact of this additive part on the overall belief will decrease with increasing age since, for a given $0 < \delta < 1$, δ_h is strictly increasing in h. That is, the older the agent gets, i.e., the more information she receives, the more is her survival belief determined by the value of the optimism parameter λ. As a consequence, comparatively small objective survival probabilities, i.e., $\pi^*_{j,m} < \lambda$, tend to be overestimated with increasing age h, whereas large objective survival probabilities, i.e., $\pi^*_{j,m} > \lambda$, tend to be underestimated. Since the objective survival probabilities are decreasing with higher age, our formal model is able to capture the stylized facts presented in Section 2 that (i) elderly people appear to be overly optimistic whereas (ii) young people are apparently overly pessimistic with respect to their survival probabilities.

If there is some likelihood insensitivity involved, the agents of our model tend to "flatten out" probabilities with increasing age. We can thus offer the following behavioral interpretation of our formal learning model:

1. With increasing age the agent loses her probabilistic sophistication because any

probabilistic differences become—in her perception—increasingly dominated by her attitudes towards likelihood insensitivity.

2. Nevertheless, the agent also keeps on learning from the data in a rational sense by updating the additive part of her subjective belief in accordance with the standard Bayesian approach.

Alternatively, one could argue that people want to minimize their "cognitive encounters with death" (Kastenbaum 2000) and suppress the notion of death the more, the more relevant the risk of dying becomes, i.e., the older they are. Interpreting such age-increasing minimization of cognitive encounters with death as an increasing deviation from a rational assessment, the increasing optimistic bias of our model might be (heuristically) interpreted as a measure for people's avoidance of "a realistic assessment of their encounter with death".[11]

Regardless of any specific behavioral—either heuristic or decision theoretic—interpretation we may want to attach to our Choquet decision-theoretic model of biased Bayesian learning, we demonstrate in the remainder of this paper that our parsimonious model captures the stylized facts of the HRS data surprisingly well.

[11] We owe this interpretation to an anonymous referee.

4 Empirical Analysis

4.1 Estimation Strategy

According to equations (6) and (7), we have to estimate three parameters, $\Psi = [\phi, \delta, \lambda]$. To estimate these parameters we pool a sample of HRS data of waves $\{2000, 2002, 2004\}$. Except for heterogeneity in sex and age, we ignore all other heterogeneity across individuals. We deliberately choose this strategy in order to focus the analysis on the main message of this paper: Choquet Bayesian learning is a more appropriate model for survival belief formation than rational Bayesian learning.[12] For notational convenience, we again do not display an index for sex.

In each interview age group j we have N_j observations denoted as $i \in \{1, \ldots, N_j\}$ where N_j differs across groups. In our estimation we weigh observations by the inverse of the group sizes, $\frac{1}{N_j}$, so that we down-weigh age groups with many observations relative to age groups with few observations and vice versa.[13] We assume a linearly additive error term and determine parameter values by solving the following non-linear minimization problem

$$\min_{\Psi} \frac{1}{2} \sum_{j=1}^{J} \frac{1}{N_j} \sum_{i=1}^{N_j} \left({}_i b_{j,m}^{j} - b_{j,m}^{j} \right)^2 .$$

[12] In particular, it is not the purpose of this paper to analyze how updating of beliefs differs across particular idiosyncratic health shocks or other events that are regarded as relevant for survival belief formation in the literature, such as parental death. As we further discuss in our concluding remarks in Section 5, a more in depth analysis of survival belief formation based on idiosyncratic events using our framework is left for future research.

[13] Observe that this weighting scheme implies that our point estimates are identical to a regression based on average survival rates of each group. Parameter estimates from an un-weighted regression are similar and are available upon request.

Here, ${}_i b^j_{j,m} = E\left[\tilde{\pi}_{j,m}, \nu_i\left(\cdot \mid j\right)\right]$ denotes individual i's conditional subjective belief to survive from interview age j to target age m as reported in the HRS data. $b^j_{j,m} = E\left[\tilde{\pi}_{j,m}, \nu\left(\cdot \mid j\right)\right]$ is the predicted subjective belief according to our model as defined in the previous section. Recall that target ages are assigned to interview ages according to the pattern in Table 1.

We solve the above non-linear programming problem using a non-linear optimizer. As unique convergence is not guaranteed for such problems, we tried various combinations of starting values, Ψ_0, and alternative optimization routines for all of our scenarios that follow. For all these combinations the numerical routines returned the same solution vector $\hat{\Psi}$. We are therefore confident that the solvers converge to the unique global minimum. We bootstrap standard errors by drawing with replacement from our data on subjective beliefs and from our predicted data on objective survival probabilities in 500 bootstrap iterations.

4.2 Main Results

Our main estimation results are summarized in Table 2. For each estimated parameter, the table contains sex specific information on the point estimates, $\hat{\Psi}$, the respective standard errors, $\hat{\sigma}(\Psi)$, and the 95% confidence intervals of the coefficient estimates, $\widehat{CI}(\Psi)$. We also report the R^2 of the regressions as well as an "average R^2", denoted as $\bar{R}^2$, as a measure of the fraction of the variation in average survival probabilities explained by our model.

Table 2: Parameter estimates

	Men			Women		
	$\hat{\Psi}$	$\hat{\sigma}(\Psi)$	$\widehat{CI}(\Psi)$	$\hat{\Psi}$	$\hat{\sigma}(\Psi)$	$\widehat{CI}(\Psi)$
Initial bias: ϕ	0.891	0.002	[0.887 0.895]	0.900	0.002	[0.896 0.905]
Likelihood Insens.: δ	0.020	0.002	[0.017 0.024]	0.021	0.001	[0.019 0.023]
Degree of optimism: λ	0.454	0.012	[0.431 0.476]	0.394	0.012	[0.371 0.419]
R^2	0.041	0.003	[0.034 0.048]	0.063	0.003	[0.057 0.069]
$\bar{R}^2$	0.803	0.035	[0.691 0.834]	0.943	0.010	[0.905 0.944]

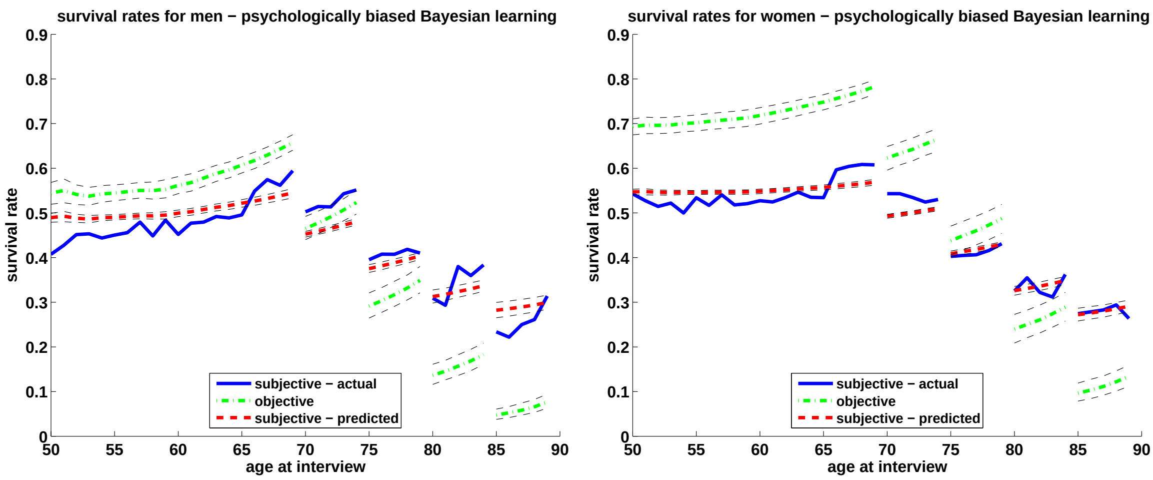
Notes: $\hat{\Psi}$ are point estimates of model parameters, $\hat{\sigma}(\Psi)$ is the respective standard deviation and $\widehat{CI}(\Psi)$ is the respective 95% confidence interval. Standard errors are calculated by bootstrapping the subjective and objective survival probabilities by drawing with replacement in 500 bootstrap iterations.
Source: Own calculations based on HRS, SSA and HMD data.

We have already argued that a model of rational Bayesian learning alone cannot explain the observed patterns in the data because predicted subjective survival rates from such a model would converge to objective data. Quite in contrast, our model of psychologically biased Bayesian learning which considers non-additive beliefs results in a decent fit to average subjective survival expectations, also see Figure 3. For both men and women, predicted subjective beliefs track the average subjective beliefs from the data nicely. The average R^2 is between 0.8 and 0.9 whereby the fit is significantly higher in case of women.[14] Unsurprisingly, our parsimonious specification of average

[14]The value of the two-sided t-test on the difference between the R^2s for men and women is 108.96 with a p-value of 0.0. The values of Jarque-Bera test statistics for normality of the distribution of the bootstrapped R^2s (and their p-values) are at 0.02 (0.99) for men and at 4.49 (0.10) for women so that a standard t-test is applicable.

Figure 3: Actual and predicted survival probabilities for psychologically biased Bayesian learning

(a) Men (b) Women



Notes: Thin dashed lines are 95% confidence intervals obtained from 500 bootstrap iterations.
Source: Own calculations based on HRS, HMD and SSA data.

beliefs results in low R^2s of the regressions—0.041 for men and 0.063 for women—because our representative agent model can only capture some of the variation in answering patterns across individuals.

All parameters are estimated with high precision. Accordingly, parameters δ and λ, which reflect the psychological biases in our model, are key for generating our results and we thereby formally reject the hypothesis of pure rational Bayesian learning. The point estimate of the initial bias, ϕ, is below one and captures the initial pessimism expressed in subjective beliefs documented in Figure 1. Interpretation of the point estimate of $\phi = 0.89$ for men ($\phi = 0.9$ for women) is that a person with one year of experience at age 20 estimates the additive probability to survive from age 50 to age 80 (for which $m - j = 30$) to be $\frac{2\phi^{m-j}+h}{2+h} \cdot 100\% = \frac{2\phi^{30}+1}{3} \cdot 100\% = 35.4\%$ (36.2%) of the objective data. Finally, a person at the age of 50 who has already gathered 31 years of experience ($h = 31$), underestimates the additive probability by factor $\frac{2\phi^{m-j}+h}{2+h} \cdot 100\% = \frac{2\phi^{30}+31}{33} \cdot 100\% = 94.1\%$ (94.2%) only. As we show in the online appendix, however, these biases should not be interpreted too literally because they crucially depend on the speed at which experience is accumulated.

Recall from equation (6) that biases in beliefs are not only governed by the additive model. We find that the measure of optimism under likelihood insensitivity is significantly higher for men, i.e., $\lambda = 0.454$ with a 95% confidence interval of $[0.431, 0.476]$, than for women, i.e., $\lambda = 0.394$ with a 95% confidence interval of $[0.371, 0.419]$.[15] At the same time, the initial degree of likelihood insensitivity is almost identical for both sexes. According to our interpretation of non-additive beliefs, the weight $(1 - \delta_h)$ measures how much evidence gained from rational Bayesian learning is taken into account. Conversely, δ_h corresponds to the weight by which beliefs are affected by some "myside bias," in our model formalized as personal attitudes towards optimism, respectively pessimism, as measured by λ. A literal interpretation of our estimation results therefore suggests that respondents of both sexes are roughly affected by the same degree of likelihood insensitivity, but men resolve their likelihood insensitivity in a more optimistic manner than women. Furthermore, our results indicate that the initial likelihood insensitivity at the age of 19, cf. Assumption 3, is rather

[15] The t-statistic of the two-sided t-test for equality of the point estimates is 3.53.

low; the point estimates are about 0.02. At age 20, with one year of experience, it is $\frac{0.02}{0.02+(1-0.02)\frac{1}{2}} = 0.039$ and at age 50 ($h = 30$) it is $\delta_{30} = \frac{0.02}{0.02+(1-0.02)\frac{1}{31}} = 0.39$, roughly ten times as large. Panel (a) in Figure 4 displays the degree of likelihood insensitivity δ_h for all ages $50 - 89$. Hence, although initial likelihood insensitivity at age 19 is relatively low, it has a strong impact on survival belief formation as individuals get older.

To interpret parameter estimates further, notice that our model represents two maps. The first map is from the objective data $\pi^{\star}_{j,m}$ to a subjective additive measure via bias factor $\frac{2\phi^{m-j}+h}{2+h}$, hence $\tilde{b}^h_{j,m} = \left(\frac{2\phi^{m-j}+h}{2+h}\right)\pi^{*}_{j,m}$. The second transformation maps the resulting object into a subjective non-additive probability measure $b_{j,m}$ with parameters δ, λ, cf. equation (6). As emphasized by Wakker (2010), this second map is a linear approximation to the following inverse S-shaped probability weighting function (cf. Tversky and Kahneman (1992), Wu and Gonzalez (1996))

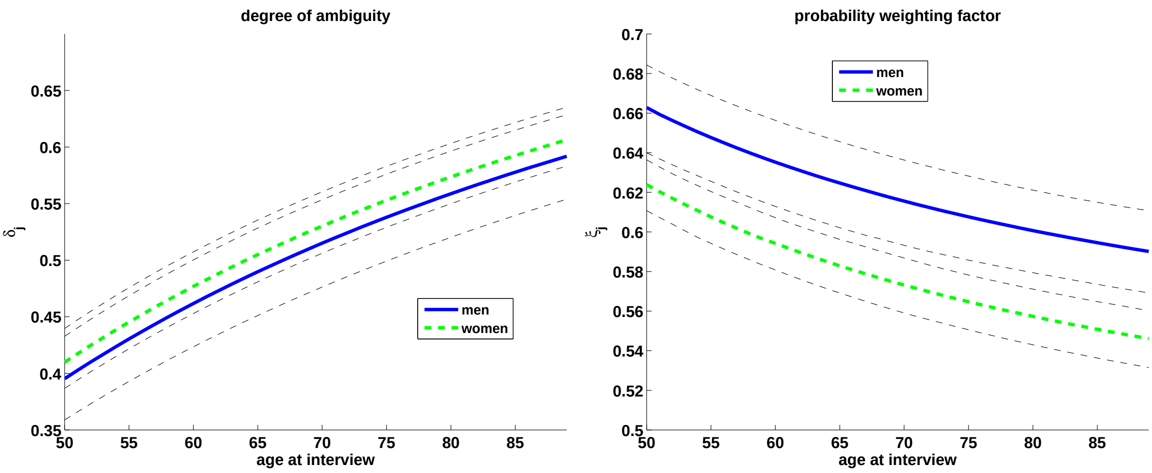
$$\omega(\xi_h, \tilde{b}^h_{j,m}) = \frac{(\tilde{b}^h_{j,m})^{\xi_h}}{\left((\tilde{b}^h_{j,m})^{\xi_h} + (1-\tilde{b}^h_{j,m})^{\xi_h}\right)^{\frac{1}{\xi_h}}} \tag{9}$$

for age-specific parameters ξ_h.

To illustrate this analogy, we back out age-specific parameters ξ_h from our estimates of δ_h and λ by minimizing the Euclidean squared distance between (6) and (9) for $h = 31, \ldots, 71$ (biological ages 50 to 90). Results of the age-specific probability weighting factors ξ_h are shown in Panel (b) of Figure 4. As ambiguity increases (Panel (a)), the linear probability transformation of equation (6) becomes less steep and curvature of the probability weighting function (9) increases, i.e., ξ_h decreases (Panel (b)). Implied values of ξ_h are well within the range of standard estimates as reported by Wu and Gonzalez (1996) which is $[0.5, 0.9]$.

Figure 4: Likelihood insensitivity and implied probability weighting factor

(a) Likelihood insensitivity (δ_h) (b) Implied probability weighting factor (ξ_h)

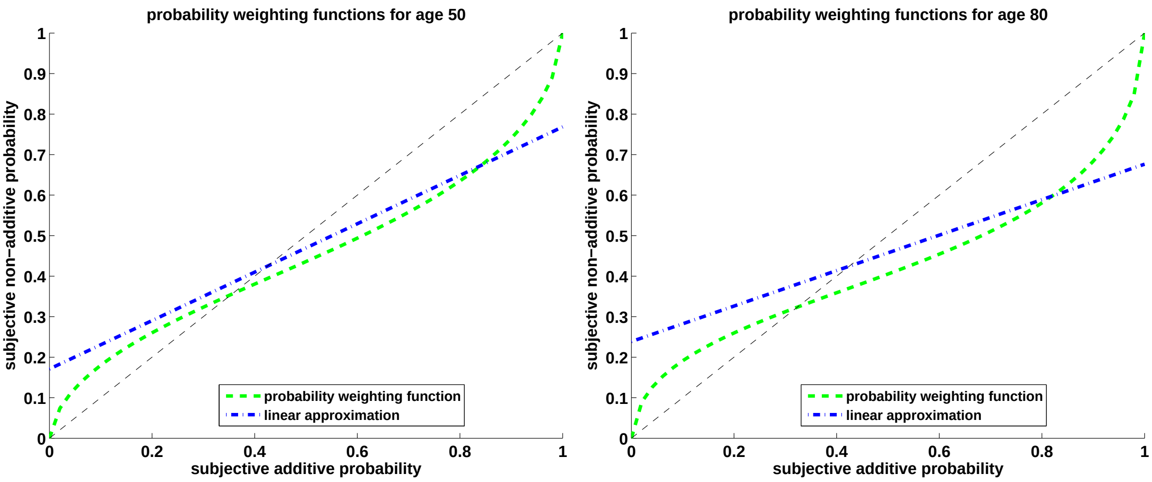


Notes: Thin dashed lines are 95% confidence intervals obtained from 500 bootstrap iterations. *Source:* Own calculations based on HRS, HMD and SSA data.

Figure 5 clarifies this analogy further. Here we plot for biological ages 50 and 80 the probability weighting function (dashed line, equation (9)), and its linear approximation (dashed-dotted line, equation (6)), for our point estimates of δ, λ and the implied values of δ_{31}, ξ_{31}, respectively δ_{61}, ξ_{61}, each for averages over sexes. The graph illustrates how likelihood insensitivity of older households increases in our model, i.e, the linear approximation to the probability weighting function (dashed dotted line) moves closer to a horizontal line where $\tilde{b}^h_{j,m} = 0.5$, a fifty-fifty probability judgement. Intersection of the linear dashed-dotted line representing equation (6) with the 45-degree line (along which $b^h_{j,m} = \tilde{b}^h_{j,m}$) is at $\tilde{b}^h_{j,m} = \lambda$. Hence, $\tilde{b}^h_{j,m} = \lambda$ represents the

Figure 5: Probability weighting functions

(a) Age 50 (b) Age 80



Notes: Implied probability weighting functions are computed for the respective point estimates averaged over sexes.

Source: Own calculations based on HRS, HMD and SSA data.

knife-edge case between an optimistic relative to a pessimistic bias with respect to a subjective additive probability measure. That age 50 (80) households exhibit a pessimistic (optimistic) bias is due to the fact that respective objective probabilities to survive to some target age are more (less) than λ, cf. Figure 3.[16]

4.3 Focal Point Answers

An apparently serious problem in data on subjective survival probabilities is the existence of "focal point answers" at self-reported survival probabilities of 0, 50, and 100 percent (Hurd and McGarry 1995; Gan, Hurd, and McFadden 2005). One interpretation for individuals indicating probabilities of 0 or 100 percent is that they have not fully understood the question.[17] Thus, focal point answers could be regarded as implausible estimates of subjective probabilities. However, as discussed by Smith, Taylor, Sloan, Johnson, and Desvouges (2001) and Khwaja, Sloan, and Chung (2007), these focal point answers at 0% and 100% still have information content regarding the correct subjective belief because smokers provide the answer 0% more frequently than non-smokers. The target age-group specific answer pattern in our data, shown in our online appendix, also illustrates that focal point answers have information content for the true subjective belief because the frequency of focal point answers at 0% increases with target age whereas the frequency of focal point answers at 100% decreases with target age. The overall pattern is the same for male and female respondents.

We investigate sensitivity of our results with respect to focal point answers by

[16] In this argument, we can ignore for age 80 households the difference between the additive objective measure π and the additive subjective measure $\tilde{b}$ because the bias factor $\frac{2\phi^{m-j}+h}{2+h}$ is close to one for that amount of experience.

[17] An alternative interpretation is that focal point answers reflect ambiguity, cf. Hill, Perry, and Willis (2006).

deleting in our sample all observations with focal point answers. This correction leaves us with a sample size of 24225 observations (10188 male and 14037 female respondents). Hence, roughly 46% of interviewees in our full sample have given focal point answers at either 0%, 50% or 100%, respectively. Estimation results for this alternative data set are summarized in Table 3. A comparison with our benchmark results in Table 2 shows that the broad pattern of estimated parameter values does not change. The only discernible difference to our benchmark results is that the estimates of the degree of optimism, λ, do not differ much across sexes. Therefore, our earlier interpretation that men seem to resolve their likelihood insensitivity in a more optimistic manner than women, is not robust with respect to the exclusion of focal point answers.

Table 3: Parameter estimates: Excluding focal point answers

	Men			Women		
	$\hat{\Psi}$	$\hat{\sigma}(\Psi)$	$\widehat{CI}(\Psi)$	$\hat{\Psi}$	$\hat{\sigma}(\Psi)$	$\widehat{CI}(\Psi)$
Initial bias: ϕ	0.894	0.004	[0.886 0.900]	0.909	0.009	[0.901 0.934]
Likelihood insens.: δ	0.023	0.003	[0.019 0.029]	0.028	0.002	[0.025 0.033]
Degree of optimism: λ	0.441	0.012	[0.417 0.467]	0.436	0.011	[0.415 0.455]
R^2	0.043	0.004	[0.035 0.051]	0.051	0.004	[0.043 0.058]
$\bar{R}^2$	0.826	0.055	[0.615 0.833]	0.879	0.033	[0.774 0.901]

Notes: These results are based on a sample which excludes focal point answers at $0\%, 50\%$ and 100%, respectively. $\hat{\Psi}$ are point estimates of model parameters, $\hat{\sigma}(\Psi)$ is the respective standard deviation and $\widehat{CI}(\Psi)$ is the respective 95% confidence interval. Standard errors are calculated by bootstrapping the subjective and objective survival probabilities by drawing with replacement in 500 bootstrap iterations.
Source: Own calculations based on HRS, SSA and HMD data.

Furthermore, our findings are robust to the structure of the experience function, cf. Assumption 3, and our choice of initial age. Results on such sensitivity analyses are available upon request.

4.4 Selectivity

One criticism raised against using population averages as the relevant objective data is that our HRS sample may be prone to selectivity. Reasons for such selection biases are either that households have moved to nursing homes and are not followed by HRS interviewers or that sick people are reachable but may not be able to answer the questionnaire.[18] Such selection effects may explain (some of) the optimism we observe at higher ages in Figure 1.

To address these concerns, we compute the HRS hazard rates between waves 2000 and 2002 and between waves 2002 and 2004, respectively, and compare them to the biannual mortality rates in the population for the respective years. In our online appendix, we present resulting hazard rates for men and women. This shows that HRS hazard rates correspond with the mortality rates in the population.

4.5 Cohort effects

An alternative view at the data might be that answering patterns depicted in Figure 1 reflect cohort rather than age effects. E.g., older households may have been more optimistic when they were young. To accommodate this aspect we plot in the online appendix subjective data for various birth cohorts. There is no apparent indication of relevant cohort effects.

[18] As Mike Hurd pointed out to us, the first selection effect was particularly severe for the early waves of the HRS because people were not followed into nursing homes in the past. Since we use the more recent waves of the HRS where people are in fact followed into nursing homes, selection effects may only play a role for the very old respondents in our sample, if at all.

5 Conclusion

The HRS data on subjective survival beliefs suggest a violation of the rational expectations paradigm as well as of the rational Bayesian learning hypothesis. In a first step we therefore propose a new model of Bayesian learning that encompasses rational Bayesian learning while it additionally allows for the existence of a psychological bias in the interpretation of new information. For this purpose our formal approach combines concepts, such as non-additive beliefs and generalized Bayesian updating, from Choquet decision theory with the standard approach of rational Bayesian learning. The resulting model of psychologically biased belief formation is very parsimonious in that it requires a low parametrization reflecting, first, an initially biased additive estimator of subjective survival probabilities, second, a measure for the agent's likelihood insensitivity with respect to her initial estimator of her subjective survival probability, and, third, a measure for the agent's optimistic versus pessimistic attitudes with respect to this likelihood insensitivity. Besides this parsimonious specification of the formation of subjective survival beliefs, our learning model has the additional advantages that, first, it is axiomatically founded within Choquet decision theory and, second, it is well supported by psychological evidence on diverging learning behavior.

In a second step we estimate parameters of our model by pooling HRS data. Despite the parsimonious parametrization we find that our model explains $80 - 94\%$ of the variation of average subjective survival probabilities in the data. The model's performance is statistically better for women than for men. For both genders we can clearly reject the hypothesis that HRS data on subjective survival probabilities may be explained by rational Bayesian learning. The reason is that the rational Bayesian learning hypothesis implies convergence of the subjective probabilities to the respective objective data at higher ages but we instead observe an increasing degree of optimism in the data. On the contrary, our more sophisticated model of psychologically biased Bayesian learning can match these patterns.

In our theoretical model we condition updating of subjective beliefs on sex and age of individuals only by which we obtain a representative agent interpretation. We deliberately choose this strategy in order to focus the analysis on the main messages of this paper, namely that Choquet Bayesian learning is a more appropriate model for survival belief formation than rational Bayesian learning. The strength of our parsimonious approach is certainly that we can directly map our model into life-cycle models of consumption and savings. Along this line, our current research evaluates the implications of our model for life-cycle models of consumption and savings. We thereby extend the work by, e.g., Bleichrodt and Eeckhoudt (2006) to a multi-period setup.

However, our simple empirical strategy does not allow us to analyze how updating of beliefs differs across a variety of observed idiosyncratic health shocks or other events that are regarded as relevant for survival belief formation in the literature, such as parental death. In our future research, we plan to modify our empirical model in such a way that the objective information is not based on average survival rates in the population but rather on objective information at the individual level. This would enable us to condition updating of beliefs on observed idiosyncratic shocks in between waves of the HRS, similar to Smith, Taylor, Sloan, Johnson, and Desvouges (2001).

References

Baron, J. (2008). *Thinking and Deciding*. New York, Melbourne, Madrid: Cambridge University Press.

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

Breiman, D., L. LeCam, and L. Schwartz (1964). Consistent Estimates and Zero-One Sets. *The Annals of Mathematical Statistics 35*, 157–161.

Bruine de Bruin, W., B. Fischhoff, and S. Halpern-Felsher, Bonnie Millstein (2000). Expressing Epistemic Uncertainty: It's a Fifty-Fifty Chance. *Organizational Behavior and Human Decision Processes 81*, 115–131.

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

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, September). Updating Choquet beliefs. *Journal of Mathematical Economics 43*(7-8), 888–899.

Eichberger, J., S. Grant, and D. Kelsey (2010). Comparing Three Ways to Update Choquet Beliefs. *Economics Letters 107(2)*, 91–94.

Eichberger, J. and D. Kelsey (1999). {E-}Capacities and the Ellsberg Paradox. *Theory and Decision 46*, 107–140.

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

Epstein, L. G. (1999). A Definition of Uncertainty Aversion. *Review of Economic Studies 66*, 579–608.

Gan, L., M. Hurd, and D. McFadden (2005). Individual Subjective Survival Curves. In D. Wise (Ed.), *Analyses in the Economics of Aging*. Cambridge, MA: NBER Books, National Bureau of Economic Research.

Ghirardato, P. and M. Marinacci (2002). Ambiguity Made Precise: A Comparative Foundation. *Journal of Economic Theory 102*, 251–289.

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

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

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

Hill, D., M. Perry, and R. Willis (2006). Estimating Knightian Uncertainty from Survival Probability Questions on the HRS.

Human Mortality Database (2008). No Title.

Hurd, M. D., D. McFadden, and A. Merrill (2001). Predictors of Mortality Among the Elderly. In D. Wise (Ed.), *Themes in the Economics of Aging*. University of Chicago Press.

Hurd, M. D. and K. McGarry (1995). Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study. *Journal of Human Resources 30(0)*, S268–S292.

Hurd, M. D. and K. McGarry (2002). The Predictive Validity of Subjective Probabilities of Survival. *Economic Journal 112*, 966–985.

Hurd, M. D., S. Rohwedder, and J. Winter (2005). Subjective Probabilities of Survival: An International Comparison.

Kastenbaum, R. (2000). *The Psychology of Death.* New York: Springer Publishing Company.

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

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

Lijoi, A., I. Pruenster, and S. O. Walker (2004). Extending Doob's Consistency Theorem to Nonparametric Densities. *Bernoulli 10*, 651–663.

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

McFadden, D. L., A. C. Bemmaor, G. C. Francis, J. Dominitz, J. Byung-Hill, A. Lewbel, R. L. Matzkin, F. Molinari, N. Schwarz, R. J. Willis, and J. K. Winter (2005). Statistical Analysis of Choice Experiments and Surveys. *Marketing Letters 16 (3-4)*, 183–196.

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

Sarin, R. and P. P. Wakker (1998). Revealed Likelihood and Knightian Uncertainty. *Journal of Risk and Uncertainty 16*, 223–250.

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*, 571–587.

Siniscalchi, M. (2006). Dynamic Choice under Ambiguity.

Smith, V. K., D. H. Taylor, and F. A. Sloan (2001). Longevity Expectations and Death: Can People Predict Their Own Demise? *American Economic Review 91(4)*, 1126–1134.

Smith, V. K., D. H. Taylor, F. A. Sloan, F. R. Johnson, and W. H. Desvouges (2001). Do Smokers Respond to Health Shocks? *The Review of Economics and Statistics 83(4)*, 675–687.

Steffen, B. (2009). *Formation and Updating of Subjective Life Expectancy.* MEA Studies 08, MEA, LMU Munich.

Tonks, I. (1983). Bayesian Learning and the Optimal Investment Decision of the Firm. *The Economic Journal 93*, 87–98.

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

Viscusi, W. K. (1985). A Bayesian Perspective on Biases in Risk Perception. *Economics Letters 17*, 59–62.

Viscusi, W. K. (1990). Do Smokers Underestimate Risks? *Journal of Political Economy 98(6)*, 1253–1269.

Viscusi, W. K. (1991). Age Variations in Risk Perceptions and Smoking Decisions. *Review of Economics and Statistics 73(4)*, 577–588.

Viscusi, W. K. and C. J. O'Connor (1984). Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers? *Amercian Economic Review 74*, 942–956.

Wakker, P. (2001). Testing and Characterizing Properties of Nonadditive Measures Through Violations of the Sure-Thing Principle. *Econometrica 69*, 1039–1059.

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

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

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

Zimper, A. (2009). Half Empty, Half Full and Why We Can Agree to Disagree Forever. *Journal of Economic Behavior and Organization 71*, 283–299.

Zimper, A. and A. Ludwig (2009). On Attitude Polarization under Bayesian Learning with Non-Additive Beliefs. *Journal of Risk and Uncertainty 39*, 181–212.