CrossMark

Individual Survival Expectations and Actual Mortality: Evidence from Dutch Survey and Administrative Data

Vesile Kutlu-Koc[1] · Adriaan Kalwij[2]

Abstract Because of the important role that survival expectations play in individual decision making, we investigate the extent to which individual responses to survival probability questions are informative about actual mortality. In contrast to earlier studies, which relied on the Health and Retirement Study (HRS) of US individuals aged 50 and over, we combine household survey data on subjective survival probabilities with administrative data on actual mortality for Dutch respondents aged 25 and over. Our main finding is that in our sample, individual life expectancies (measured as subjective survival probabilities) do predict actual mortality even when we control for a large set of health indicators. Our results further suggest that, on average, women underestimate their remaining life duration, whereas men tend to predict their survival chances more realistically. Both sexes, however, tend to overestimate the age gradient in mortality risk and underestimate the health risks of smoking.

1 Introduction

One reason that mortality risk plays such a key role in individual decision making is that individuals may take life expectancy into account when making economic choices such as when to retire, how much to save for old age, whether to purchase a

---

✉ Vesile Kutlu-Koc
   kutlukoc@mea.mpisoc.mpg.de

   Adriaan Kalwij
   A.S.Kalwij@uu.nl

[1] Munich Center for the Economics of Aging (MEA), Amalienstr. 33, Room. 360, 80799 Munich, Germany

[2] Utrecht University School of Economics, Kriekenpitplein 21-22, Room. 3.02, 3584 EC Utrecht, Netherlands

life insurance policy, and how much money to set aside for bequests. It is as well a consideration in late life decisions like whether or when to move to a nursing home. As a result, theoretical economic models of life cycle behavior have shown the importance of allowing for lifetime uncertainty for economic outcomes (e.g., Hurd 1989) and since the 1990s, several household surveys have begun including probabilistic questions (Manski 2004) that can be used to measure individual lifetime uncertainty. These so-called subjective survival probabilities (SSPs) elicited from survey respondents are promising because unlike standardized tools such as life tables, they provide individual variation in survival chances and use an interpersonally comparable numerical scale of probabilities (Hurd 2009). Nevertheless, the usefulness of SSPs hinges largely on the assumption that respondents are able to answer these questions in a meaningful way.

Our paper thus focuses on the steps that must be taken before developing and estimating economic models that incorporate SSPs. In particular, by comparing average objective mortality in a group with average SSPs (cf. Hurd 2002; Khwaja et al. 2007; Delavande and Rohwedder 2011), we investigate the extent to which individual beliefs about survival relate to actual mortality.[1] There is, however, a general limitation of this validation exercise which is worth mentioning at the beginning. Although SSPs are broadly in line with objective values, some notable discrepancies suggest either that individuals' underlying beliefs are biased or that SSPs do not accurately measure beliefs. If the latter is correct, then either SSPs are inherently useless (at least for some groups) or the presently used question formats are not working efficiently. Hence, before any understanding can be reached of why differences exist between objective mortality and SSPs, these discrepancies must be documented. This paper aims to provide such documentation for the Netherlands.

Earlier research shows that life cycle models that measure mortality based on life tables rather than SSPs are unable to explain several common data anomalies. These include retirees having large amounts of assets even at advanced ages (Poterba et al. 2011; Van Ooijen et al. 2014), inadequate savings before retirement (Hurd and McGarry 2002), a falling retirement age as life expectancy increases (O'Donnell et al. 2008), and the annuitization puzzle (Teppa and Lafourcade 2013). SSPs may shed light on these issues because they contain information about individuals' survival beliefs, which can differ from actuarial survival probabilities.[2] For instance, Teppa and Lafourcade (2013), after finding that annuity demand is higher for those who expect to live longer, argued that individual underestimation of life expectancy may be a reason why only a small fraction of Dutch individuals buy life annuities.

---

[1] Earlier studies in this field have compared SSPs to actuarial survival probabilities when data on actual mortality were unavailable; see, for example, Hamermesh (1985), and Hurd and McGarry (1995) for the USA, O'Donnell et al. (2008) for the UK, and Peracchi and Perotti (2011) for Europe, Teppa and Lafourcade (2013) for the Netherlands, Bucher-Koenen and Kluth (2012) for Germany, and Wu et al. (2015) for Australia.

[2] Alternative explanations exist for why life cycle models with actuarial survival probabilities might fail to explain observed patterns in the data. For instance, in the case of inadequate accumulated assets to finance retirement, the reasons may include lack of financial literacy (Lusardi and Mitchell 2011), incorrect beliefs about future retirement benefits (Rohwedder and van Soest, 2006), and/or hyperbolic discounting (Laibson et al. 1998).

Another related question is whether SSPs predict individuals' economic decisions, especially given US evidence that they explain savings and consumption behaviors better than do life table probabilities (Gan et al. 2004; Salm 2010).

The novelty of our paper is the use of Dutch data and the inclusion of 25- to 49-year-olds in the analysis. To our knowledge, this is the first study which investigates the relation between SSP and actual mortality for a sample that includes individuals under 50. Much previous research (e.g., Smith et al. 2001; Hurd and McGarry 2002; Siegel et al. 2003; Perozek 2008; Elder 2013) has relied on US Health and Retirement Study (HRS) data for those aged 50 and over. The goals of the paper are twofold. First, we test an observation made by several other researchers (Smith et al. 2001; Hurd and McGarry 2002; Siegel et al. 2003) by exploring whether these Dutch respondents' beliefs about their survival chances (SSPs) contain predictive power for their own deaths even when a large set of health indicators are controlled for. Any predictive power for realized mortality in the presence of health indicators would suggest that SSPs convey information not contained in an individual's current health status, such as potential disease risks or parental longevity (Hurd and McGarry 2002).

The overall research evidence on the accuracy of SSPs, however, is inconclusive. On the one hand, Perozek (2008) investigated whether the Social Security Actuary's (SSA) revision to the longevity gender gap could be predicted by the gender gap implied by the subjective cohort life tables based on individuals' SSPs. She found that the latter predicts a smaller gender difference in life expectancy than did the SSA life tables and, thus, claimed that SSPs are better predictors than actuarial forecasts. Elder (2013), however, argued that SSPs are much worse predictors than actuarial forecasts as the implied age profiles are too flat. We contribute to this debate by providing new evidence from another country and a broader age range.

Our second goal is to expand on previous US-based research into whether individuals over- or underestimate their survival chances (e.g., Schoenbaum 1997; Khwaja et al. 2007; Delavande and Rohwedder 2011; Gerking and Khaddaria 2011; Bissonnette et al. 2014), thereby taking into account certain personal characteristics related to socioeconomic status, health, and behavioral risk indicators. Khwaja et al. (2007), for instance, showed that on average SSPs are very close to their objective counterparts, with current smokers being optimistic and those who never smoked being pessimistic in their assessments. Delavande and Rohwedder (2011) subsequently demonstrated that actual mortality and subjective survival expectations are similarly associated with wealth, income, and education.

2 Data

This paper combines Dutch survey data, which include respondents' survival expectations, with individual-level administrative data that contain the respondents' date of death. The survival expectations, expressed as SSPs, are taken from the 1995 and 1996 waves of the DNB Household Survey (DNB-HS), an ongoing longitudinal panel survey formerly known as the CentER Savings Survey. Although initiated in 1993, the survey did not adopt the SSP questions until 1995; hence, our data set

includes only the 1995 and 1996 survey data because these two waves alone are linkable to our administrative data.

The DNB-HS database, compiled using an Internet survey of around 2550 Dutch households, includes detailed information on respondent age, income, health, education, labor market status, assets and liabilities, and psychological state (see Alessie et al. 2002 for a detailed description). Each year, all household members aged 16 or over are interviewed online. Those who do not have a computer and/or Internet access are provided with these tools by the survey agency. The DNB-HS consists of two different panels: a nationwide panel of around 1900 households and a high-income panel of around 650 households representing the top 10% of the income distribution. This sampling method could make our baseline sample non-random. As discussed by Delavande and Rohwedder (2011), non-randomness of the sample does not affect the validity of our empirical results as our analysis compares subjective and objective mortality for a given population. We nevertheless check whether sample non-randomness affects our results by repeating our analysis with a sample restricted to households in the representative panel survey (see Sect. 4).

The DNB-HS measures subjective survival probabilities using the following survey question: 'How big do you think is the chance that you will attain (at least) the age of T?' where T ε {75, 80, 85, 90, 95, 100} is a target age dependent on the respondent's current age. That is, respondents aged 25 through 65 report their survival probability to ages 75 and 80; those aged 65 through 70, to ages 80 and 85; and those aged 70–75, 75–80, and 80–85, to ages 85 and 90, 90 and 95, and 95 and 100, respectively. The responses are measured on a 10-point scale from 0, 'no chance at all,' to 10, 'absolutely certain.' In our analysis, we follow Hurd and McGarry (1995) by assuming that once divided by 10, the responses can be interpreted as probabilities conditional on being alive at a certain age. To construct our main variable of interest, median remaining life duration, we use the survival probabilities for the two different target ages reported by each individual.

The actual mortality data for the survey respondents were obtained from the Dutch causes of death registry (*Doods Oorzaken*, DO), which recorded the death dates of all residents deceased during the 1995–2010 period. These data were provided by medical examiners, who are legally obliged to submit them to Statistics Netherlands. The DO data set also assigns a personal identifier that matches the personal identifier in the DNB-HS, thereby allowing determination of whether individuals in the 1995 or 1996 DNB-HS waves were still alive at the end of the observation period (December 31, 2010) or whether they had died and on which date.

2.1 Sample Selection

If respondents are in both the 1995 and 1996 waves, we use only the earlier response to avoid any potential influence of repeated interviewing on respondent behavior (Lazarsfeld 1940; Sturgis et al. 2009). This method eliminates any possible learning effect such as respondents asked about survival probability in 1995 seeking more information about their survival chances before responding in 1996. As a result, we have essentially a cross-sectional sample comprising one observation per individual

in either 1995 or 1996. We then further exclude individuals under 25 who may still be enrolled in education and have no individual income data available, for a sample size of 5747.

Before starting our analysis, we assess the extent to which DNB-HS respondents were willing and able to respond to SSP questions by reporting the response rates, fraction of consistent answers, and fraction of focal point responses (i.e., answers that cluster around 0, 0.5, or 1). As regards response willingness, earlier research has revealed that the non-response rate to probabilistic questions is as low as that to traditional questions on the same subjects (see Manski 2004 for a review of the US findings). The response rate in our sample is about 86%, considerably lower than the approximately 98 or 90% for the HRS (Hurd and McGarry 1995) and SHARE surveys (Peracchi and Perotti 2011), respectively. Nevertheless, this response rate suggests that a majority of the respondents were willing to answer SSP questions.

It is also well documented in the US literature that most respondents are able to provide SSP responses that are consistent with the laws of probability (e.g., Hurd and McGarry 1995; Perozek 2008). For example, if a respondent indicates a survival probability to age 75 that is less than or equal to the survival probability to age 80, the answer violates the strict monotonicity assumption. In fact, the survival probability to age 75 should be greater than that to age 80 because reaching age 80 requires survival to 75. Moreover, because respondents face a non-negligible mortality risk between the ages of 75 and 80, those who provide the same answer to these two questions are being internally inconsistent, an issue discussed in more detail below.

The fraction of respondents providing answers that satisfied the strict monotonicity assumption was about 67%. Around 32% reported equal survival probabilities for the two target ages, while only about 0.63% indicated a survival chance to the earlier target age that was less than that to the later target age. As noted by Perozek (2008), respondents with equal survival probabilities can still give valuable information about the shape of individual survivor functions. Since the answer to survival probability questions in DNB-HS survey ranges from 0 to 10, respondents who provide equal survival probabilities for the two target ages may be rounding out their true survival probabilities to the nearest tenth. Thus, following Perozek (2008), we retain respondents with equal probabilities in our sample by assuming a 10% shift between the equal responses.[3]

Similarly, the coarseness of the rating scale offered could lead individuals providing focal point answers to SSPs to round their true probabilities to the nearest tenth. Fifty–fifty (50%) answers in particular, rather than reflecting individuals' true assessments of their survival probabilities, could be driven purely by uncertainty about time of death (see Fischhoff and Bruine de Bruin 1999; Hudomiet and Willis 2013). In our sample, the percentage of 50% focal point SSP answers is high, between 23 and 30%, depending on the target age in the question. In light of the

---

[3] For example, the probabilities of an individual who reports P75 = P80 = 0.5 are reassigned so that P75 = 0.55 and P80 = 0.45. Likewise, for the reasons discussed in the Appendix, the probabilities of 0 and 1 are replaced with 0.01 and 0.99, respectively. If the equal probabilities are P75 = P80 = 0.01, then P75 is set to 0.05, and if P75 = P80 = 0.99, then P80 is set to 0.95. We perform robustness checks of this assumption in Sect. 4.

Kleinjans and van Soest (2013) evidence that when rounding and 50% focal point answers are taken into account, the coefficient estimates on the determinants of subjective survival probabilities do not change significantly, we do not treat focal responses any differently.

The baseline sample, then, includes 5747 observations, 1313 (22.9%) of which we are forced to drop because either respondents' reported survival chances to the earlier target age are less than that to the later target age (0.63% of the sample) or information is missing for one of the covariates, most often household income. Our final sample thus includes 4434 individuals aged 25 and over, 463 (10.4%) of whom died during the 15 years of follow-up.

2.2 Covariates

The main variable of interest in the mortality hazard model is subjective remaining life duration, which, as demonstrated in Sect. 3.3, is calculated based on two SSPs. The other variables are gender, birth year, a year dummy for 1996, educational attainment, household income, marital status, whether or not the individual has a chronic illness, self-rated health, and other health-related factors such as smoking, heavy drinking, and being overweight or obese. Smokers are individuals who smoke cigarettes every now and then or every day; heavy drinkers are those who consume at least four drinks a day. Overweight and obese are defined as a body mass index (BMI) greater than or equal to 25 but less than 30, and greater than or equal to 30, respectively.

As Table 1 shows, women, with a 47% share, are slightly underrepresented in our sample. The mean respondent age at the time of interview is about 47 years. About 87.9% of the respondents are married, 31.9% are current smokers, 8% are heavy drinkers, 33.2% are overweight, 6% are obese, 82.9% self-rate their health as good or excellent, with only 23.7% reporting a chronic condition such as a long-term illness, disorder, or disability. The sample consists mostly of individuals with pre-university education or junior/senior vocational training. To measure income, we adopt the Statistics Netherland (SN) definition of standardized household income as the sum of the net annual incomes of all household members divided by SN's equivalence scale (Siermann et al. 2004). The average annual standardized household income for our sample is f43,349 (€19,724). To mitigate the effects of outliers, we follow Delavande and Rohwedder (2011) and divide the income distribution into three parts of equal size, with a tercile dummy indicating to which income tercile the respondent belongs.

The role that each covariate could play in the objective and subjective mortality models as follows: Gender takes into account that women on average live longer than men. To account for a possible time effect, we add a year dummy for 1996. Birth year controls not only for the increase in life expectancy over generations but also for the negative correlation between subjective remaining life duration and baseline age. Controlling for the latter is of particular importance as the scale parameter of the hazard function should be age invariant. We capture the socioeconomic differences in mortality by educational attainment, household income, and marital status (Van Kippersluis et al. 2009; Kalwij et al. 2013). Finally,

Table 1 Variable means, medians, and standard deviations

Variable	Mean	Median	Standard deviation
Female	0.467	0	0.499
Year dummy for 1996	0.167	0	0.373
Smoker	0.319	0	0.466
Good health	0.829	1	0.377
Heavy drinker	0.079	0	0.269
Chronic illness	0.237	0	0.425
Overweight	0.332	0	0.471
Obese	0.058	0	0.234
Low education	0.237	0	0.426
High education	0.371	0	0.483
Standardized household income (in euros)	43,349.24	40,092.46	25,356.68
Married	0.879	1	0.326
Single	0.097	0	0.297
Widowed	0.024	0	0.152
Age at the time of the interview (in years)	46.849	45.333	12.467

$N = 4434$

we include self-rated health, behavioral factors such as smoking, heavy drinking, and BMI, and whether or not the individual has a chronic illness to account for mortality differences across individuals with different current health statuses. In addition to those covariates, we also estimate the shape parameter of the objective and subjective hazard functions, γ, to capture an increase in mortality risk with age.

3 Estimation Methodology

Most previous literature has relied on proportional hazard rate models to estimate the risk of death during a specific time period, with life duration modeled either by parametric distributions like those of Gompertz and Weibull or semiparametric methods like the Cox proportional hazard model (Gompertz 1825; Cox 1972; Wilson 1994; Siegel et al. 2003; Perozek 2008; Bissonnette et al. 2014). In this paper, we adopt a parametric approach to be able to find estimates of a mortality risk model based on both subjective and objective survival information. In contrast to a semiparametric Cox model, in parametric models the shape of the hazard function is determined by the functional form assumptions about the duration dependence. As demonstrated in the next two subsections, choosing a parametric hazard rate function allows comparison between the estimated parameters of the subjective and objective mortality models. Because we observe only two SSPs for each individual, we can estimate the subjective model by assuming a functional form for individual survival functions. The two parameters of, for instance, the Gompertz or Weibull distributions are exactly identified given two points on the survival function. According to Bissonnette et al. (2014), both these distributions lead to quantitatively

similar estimates of objective and subjective mortality models. In our analysis, however, based on its extensive demographic use to model human mortality, we opt for the Gompertz distribution, whose survival function tends to fit the survival data of humans aged 10 to at least 85 better than the Weibull survival function (Wilson 1994).[4]

As to the role of individual characteristics, because they are observed only in the baseline year, our model assumes that, other than age, they are all time-invariant. In other words, we estimate only how objective and subjective mortality risks are associated with current socioeconomic characteristics and health status. The next subsections explain our statistical models and empirical specifications in more detail.

3.1 Objective Mortality Model

In our objective mortality model, respondent i is aged $t_{0,i}$ at the start of the observation period (in 1995 or 1996) and aged t_i at the end of the observation period (December 2010) or at the time of death, whichever comes first. The respondent's characteristics are denoted by $\mathbf{x_i}$ and measured at the start of the observation period (in 1995 or 1996). T is a random variable representing the respondent's age at death (life duration), which is assumed to follow a Gompertz distribution, such that the hazard function can be written as

$$\theta(t_i|\mathbf{x_i}; \boldsymbol{\beta}^{obj}, \gamma^{obj}) = \exp(\gamma^{obj} t_i) \exp(\mathbf{x_i}\boldsymbol{\beta}^{obj}). \tag{1}$$

The maximum likelihood estimates of the model parameters are then given by

$$\left(\hat{\boldsymbol{\beta}}^{obj}, \hat{\gamma}^{obj}\right) = \underset{\boldsymbol{\beta}^{obj}, \gamma^{obj}}{\operatorname{argmax}} \log \prod_{i=1}^{N} \exp\left\{ -\int_{t_{0,i}}^{t_i} \theta(s|\mathbf{x_i}; \boldsymbol{\beta}^{obj}, \gamma^{obj}) ds \right\} \left(\theta(t_i|\mathbf{x_i}; \boldsymbol{\beta}^{obj}, \gamma^{obj})\right)^{d_i}, \tag{2}$$

where N is the number of individuals in our sample, and d_i is a dummy variable equal to 1 if the respondent has died at time t_i and 0 otherwise. The age gradient, hereafter mostly referred to as the shape parameter, is determined by parameter γ.

3.2 Subjective Mortality Model

For this model, we use subjective mortality information to estimate a set of parameters analogous to those of the objective mortality model. As before, we assume a Gompertz hazard function given by

---

[4] It is also worth mentioning that, as Wilson (1994) and Perozek (2008) both noted, in very old age, mortality rates may increase at a lower rate than the Gompertz function predicts. In fact, Perozek (2008) found that a subjective cohort life table derived from the Gompertz distribution predicted about a 2-year shorter life expectancy than that derived from the Weibull distribution, suggesting that use of the Weibull distribution would produce slightly longer life expectancies implied by the objective and subjective mortality models.

$$\theta(t|\mathbf{x_i}; \boldsymbol{\beta}^{subj}, \gamma^{subj}) = \exp(\gamma^{subj}t)\exp(\mathbf{x_i}\boldsymbol{\beta}^{subj}) \qquad (3)$$

Given the resulting survivor function, target ages, and SSPs, we can then estimate the following system of linear equations (see the Appendix for details):

$$\gamma_i^* = \gamma^{subj} + \varepsilon_{1i} \qquad (4)$$

$$\ln(\lambda_i^*) = \mathbf{x}_i^{'}\boldsymbol{\beta}^{subj} + \varepsilon_{2i} \qquad (5)$$

where γ_i^* and λ_i^* are the estimated parameters of the individual survival functions, and the error terms ε_{1i} and ε_{2i} are allowed to be correlated with each other. We obtain the $\hat{\beta}^{subj}$ and $\hat{\gamma}^{subj}$ estimates using seemingly unrelated regression estimator (SUR; Zellner 1962).

3.3 Empirical Specification

The objective of our empirical analysis is to show how well Dutch beliefs about survival relate to actual mortality. To do so, we first check whether SSPs are good predictors of actual mortality even when a large set of health indicators are controlled for, an assessment tested in Sect. 4.2 by estimating the objective mortality model using subjective remaining life duration as a covariate while again controlling for health indicators. This first step, however, although it suggests whether or not SSPs are useful measures, does not identify which groups are relatively better or worse at predicting their life expectancies. Hence, in a second step, we compare the estimated coefficients of subjective and objective mortality models and document the observed discrepancies. This analysis both differentiates groups of people based on socioeconomic status, health, and behavioral risk indicators (e.g., smoking and drinking) and distinguishes those who are relatively correct in predicting their life expectancy from those who over- or underestimate it.

We construct our main variable of interest in the objective mortality model based on the information contained in the respondents' answers to the two SSPs. To do so, we must solve two unknown parameters of the Gompertz distribution using two SSPs for each individual. We then use the derived survival function parameters to compute the subjective median remaining life duration conditional on baseline age for each individual in the sample [Eq. (12), in the Appendix]. We compute median life expectancy rather than expected or average life expectancy because the former has a closed-form solution, whereas the latter requires a discrete approximation of the continuous distribution (see, e.g., van Santen 2013). Likewise, because respondents report their SSPs knowing that they have survived up to their current age, we calculate remaining life duration conditional on baseline age. In addition to subjective median remaining life duration, we include several covariates in the scale parameter of the objective hazard function [Eq. (1) and described in Sect. 2.2] that we also use to estimate the scale parameter of the subjective hazard function (Eq. (3)).

We report our key findings both for the main sample of individuals aged 25+ and separately for an older sample of 50+ individuals. We check the results' sensitivity

to the main sample's non-randomness by excluding the 27% of individuals in the high-income panel. We also conduct additional sensitivity checks to test whether the results are robust to different assumptions about equal probabilities. After first assuming flatter survival functions than before—a 5% shift between equal probabilities rather than a 10% one—we exclude all equal answers from the estimation.

4 Empirical Results

4.1 Descriptive Analysis

A comparison of the subjective and objective remaining life durations across age categories and gender (see Table 2) reveals that for males, the subjective values are marginally lower than the objective values up to age 70 but then slightly exceed their objective counterparts. These differences, although small, are statistically significant at a 5% level for all male age groups except those 65–69. For females, in contrast, the subjective values are substantially lower than their objective counterparts at all ages, and all these differences are significant at the 5% level except for those aged 75–79. On average, the difference between subjective and objective remaining life duration is much smaller for males than for females. Moreover, as the table shows, male and female subjective remaining life durations are quite close to each other at all ages even though the objective values are different. Taken together, these results suggest that males are slightly pessimistic about their survival chances at younger ages but tend to be slightly optimistic at older ages, while females appear overly pessimistic at all ages.

4.2 Predictive Power of SSPs for Actual Mortality

To assess whether SSPs predict actual mortality, we first estimate our objective mortality risk model and then stepwise include a set of covariates (socioeconomic variables and health indicators) in addition to the subjective (median) remaining life duration. After a likelihood ratio test indicates that the pooling of male and female samples at a 5% significance level should not be rejected for any model, we include a gender control variable instead of reporting separate results for men and women. The signs of the estimates are informative about whether mortality hazard increases or decreases with respect to individual characteristics, while the exponential of the coefficients shows the size of the relative change in mortality risk.

The first model in Table 3 explains mortality risk only as a function of birth year, year dummy, age, and subjective remaining life duration. The coefficient of the latter is negative and statistically significant, suggesting that those who expected to live longer in the baseline year experienced a lower mortality risk than those who expected a shorter life. In the second model, which includes additional controls for gender and education, the predictive power of subjective remaining life duration remains unchanged and significant. The model does suggest, however, that, as might be expected, men have a higher mortality risk than women, and the more highly

Table 2 Means of objective and subjective remaining life durations across age categories and gender (in years)

Age	Remaining life duration			Remaining life duration		
	Objective (1) Males	Subjective (2)	(2)–(1)	Objective (1) Females	Subjective (2)	(2)–(1)
25–29	54.35	53.12	−1.22 (0.53)	61.11	54.06	−7.04 (0.40)
30–34	49.74	47.56	−2.19 (0.39)	56.62	48.87	−7.74 (0.33)
35–39	44.78	42.91	−1.87 (0.35)	51.58	42.77	−8.80 (0.31)
40–44	40.06	37.51	−2.55 (0.32)	46.64	38.33	−8.31 (0.32)
45–49	35.06	33.39	−1.67 (0.29)	41.67	34.06	−7.61 (0.31)
50–54	30.35	28.06	−2.29 (0.32)	36.91	28.87	−8.04 (0.38)
55–59	25.33	22.77	−2.56 (0.40)	31.83	23.58	−8.25 (0.40)
60–64	20.82	19.41	−1.41 (0.40)	27.20	19.59	−7.61 (0.47)
65–69	16.62	15.84	−0.78 (0.49)	22.51	16.65	−5.86 (0.52)
70–74	12.76	14.15	1.39 (0.60)	18.37	13.60	−4.77 (0.64)
75–79	9.27	11.19	1.92 (0.76)	14.55	12.64	−1.91 (1.10)
80–84	X	X	X	X	X	X
Sample average	34.95	33.23	−1.73 (0.12)	43.48	35.74	−7.74 (0.12)

$N = 4434$. Standard errors are in parentheses. X denotes a number that Statistics Netherlands cannot release because the underlying number of units is less than 10. The objective remaining life duration is computed based on a Gompertz hazard function estimated as a function of gender and age

educated have, on average, a lower mortality risk than those with medium-level education. The p value of the Wald test for education (bottom of the table) also indicates that the coefficients of high and low education are jointly significant at a 1% level.

In the third model, which adds in standardized household income terciles and marital status as control variables, the predictive power of subjective remaining life duration is the same as before, but household income is only weakly associated with mortality risk and the respondent's marital status has no significant effect on actual mortality. Adding health indicators into the set of covariates in the fourth model slightly reduces the magnitude of the effect, but it is still statistically significant at the 1% level. In fact, the size of the coefficient suggests that a 1-year increase in subjective remaining life duration is associated with a 2.3% lower mortality risk or, on average, a 0.6 years[5] longer realized remaining life duration. Among the health indicators, the statistically significant determinants of mortality risk are smoking, heavy drinking, obesity, having chronic illnesses, and being in good health, with the coefficients of obesity and overweight being jointly significant. The coefficient corresponding to the year dummy for 1996, however, is insignificant. As is to be expected, in all models, mortality risk increases significantly with age.

---

[5] We calculate this number using Eq. (12) for an average individual in our sample who was aged 45 in 2000. This individual's mortality hazard is calculated by taking the average of annual mortality rates of men and women available in Human Mortality Database (HMD 2014).

Table 3 Objective mortality: estimations

Covariates in the scale parameter of the hazard function	(1)	(2)	(3)	(4)
Subjective remaining life duration (in years)	−0.041*** (0.008)	−0.038*** (0.007)	−0.039*** (0.007)	−0.023*** (0.008)
Birth year	0.019 (0.013)	0.022* (0.012)	0.025* (0.013)	0.013 (0.013)
Year dummy for 1996	0.027 (0.149)	0.018 (0.149)	0.015 (0.150)	0.087 (0.151)
Female		−0.785*** (0.105)	−0.793*** (0.107)	−0.772*** (0.109)
Low education		0.062 (0.115)	0.067 (0.116)	0.063 (0.116)
High education		−0.282** (0.110)	−0.232* (0.118)	−0.152 (0.119)
Lowest standardized income tercile			0.198* (0.118)	0.114 (0.119)
Highest standardized income tercile			0.043 (0.119)	0.052 (0.119)
Single			0.244 (0.155)	0.107 (0.156)
Widowed			0.157 (0.177)	0.077 (0.178)
Smoker				0.647*** (0.104)
Good health				−0.372*** (0.118)
Heavy drinker				0.274* (0.146)
Overweight				−0.004 (0.100)
Obese				0.428** (0.167)
Chronic illness				0.236** (0.112)
Constant	−10.873*** (0.860)	−10.851*** (0.863)	−11.056*** (0.870)	−11.279*** (0.877)
Shape parameter (γ)	0.008*** (0.001)	0.009*** (0.001)	0.009*** (0.001)	0.009*** (0.001)
Log likelihood	−537.378	−506.458	−503.341	−464.579
p value Wald test: education		0.009	0.054	0.246

Table 3 continued

Covariates in the scale parameter of the hazard function	(1)	(2)	(3)	(4)
p value Wald test: education and income			0.015	0.311
p value Wald test: education, income, marital status			0.014	0.486
p value Wald test: overweight and obese				0.028

$N = 4434$; no. of failures $= 463$; standard errors are given in parentheses

$* p < 0.10$, $** p < 0.05$, $*** p < 0.01$

Table 4 illustrates the association between subjective remaining life duration and actual mortality when all controls are included under alternative specifications. In the first column, we restrict the sample to 50+ individuals for whom the predictive power of subjective remaining life duration remains unchanged albeit less precisely estimated. The second column reports the estimates when the main sample is restricted to the representative DNB-HS panel, which excludes high-income respondents. In this restricted sample, the magnitude of the coefficient on subjective remaining life duration is slightly reduced, although it remains significant at the 10% level. The third column then lists the outcomes for the main sample under the assumption of a 5% shift between equal probabilities. Again, the magnitude of the effect is quite close to that obtained under an assumption of a 10% shift between equal answers. Finally, the last column shows the results when all individuals who answered the same number for both probability questions are excluded, under which specification the predictive power of subjective remaining life duration completely disappears. It should be noted, however, that discarding equal responses eliminates 32% of the main sample, possibly producing an endogenous sample selection that could influence the estimates. Nevertheless, because equal responses are still informative about the shape of individual survivor functions, we conclude that a correct specification should include these responses.

4.3 Objective and Subjective Mortality Risk Models

As a second step in our validation exercise, we estimate the subjective and objective mortality risk models while including the same socioeconomic variables and health indicators in each model. If the respondents are able to predict their remaining lifetimes correctly, then the signs and magnitudes of the estimates obtained from the objective mortality model should coincide with those obtained from the subjective mortality model. We can then use the estimated coefficients to calculate the predicted median remaining life durations in order to assess which subgroups over- or underestimate their life expectancies [see Eqs. (13) and (14), respectively, in the Appendix].

The first two models reported in Table 5 explain subjective and objective mortality risk as a function of birth year, the year dummy, age effects, and socioeconomic variables. In both models, the shape parameter of the hazard function has the same sign and is statistically significant, although the age gradient is steeper in the subjective than in the objective mortality model. The coefficient on female in the subjective model, however, is insignificant even though the objective model predicts that women have a lower mortality risk than men. This finding suggests that males and females have similar beliefs about survival probabilities, but females do not, on average, expect to live longer than males. In fact, according to the Wald test, the female coefficient is significantly higher in the objective model at the 5% level (not reported in Table 5). We also find that highly educated individuals expect to die sooner than the medium educated even though the objective model suggests the opposite. However, the Wald test (not reported in Table 5) indicates that the equality of the education parameters in the two models cannot be rejected.

Table 4 Objective mortality: sensitivity checks

Covariates in the scale parameter of the hazard function	50+ sample (1)	Representative panel (2)	5% shift in equal prob. (3)	Without equal prob. (4)
Subj. remaining life duration (in years)	−0.020** (0.009)	−0.017* (0.009)	−0.018*** (0.006)	−0.005 (0.009)
Birth year	0.005 (0.014)	0.012 (0.014)	0.009 (0.012)	0.001 (0.154)
Year dummy for 1996	0.003 (0.183)	0.009 (0.155)	0.088 (0.151)	0.062 (0.189)
Female	−0.801*** (0.123)	−0.716*** (0.118)	−0.770*** (0.109)	−0.777*** (0.133)
Low education	0.128 (0.129)	0.010 (0.124)	0.068 (0.116)	0.163 (0.141)
High education	−0.059 (0.132)	−0.070 (0.133)	−0.152 (0.119)	−0.040 (0.139)
Lowest standardized income tercile	0.084 (0.124)	0.141 (0.122)	0.109 (0.119)	0.038 (0.134)
Highest standardized income tercile	−0.126 (0.145)	0.195 (0.137)	0.051 (0.119)	−0.043 (0.143)
Single	−0.065 (0.188)	0.031 (0.166)	0.106 (0.156)	0.096 (0.188)
Widowed	0.034 (0.182)	0.016 (0.186)	0.066 (0.178)	0.104 (0.207)
Smoker	0.538*** (0.120)	0.679*** (0.114)	0.654*** (0.104)	0.726*** (0.125)
Good health	−0.335** (0.130)	−0.278* (0.131)	−0.373*** (0.118)	−0.442*** (0.138)
Heavy drinker	0.242 (0.172)	0.312* (0.165)	0.278* (0.146)	0.297* (0.172)
Overweight	0.054 (0.111)	0.046 (0.109)	−0.004 (0.100)	−0.088 (0.120)
Obese	0.436** (0.185)	0.442** (0.176)	0.427** (0.167)	0.464** (0.191)
Chronic illness	0.238* (0.122)	0.226* (0.121)	0.239** (0.112)	0.260* (0.132)
Constant	−8.678*** (0.684)	−11.576*** (0.964)	−11.268*** (0.877)	−11.502*** (1.043)
Shape parameter (γ)	0.010*** (0.001)	0.009*** (0.001)	0.009*** (0.001)	0.009*** (0.001)
Number of obs.	1601	3229	4434	3059
p value Wald test: education, income, and marital status	0.490	0.838	0.497	0.788
p value Wald test: BMI	0.062	0.042	0.029	0.018

Standard errors are given in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5 Subjective and objective mortality risk: estimations

Covariates in the scale parameter of the hazard function	Subjective (1)	Objective (2)	Subjective (3)	Objective (4)	Subjective (5)	Objective (6)
Birth year	0.005*** (0.001)	−0.004 (0.011)	0.007*** (0.001)	−0.004 (0.011)	0.008*** (0.002)	−0.003 (0.011)
Year dummy for 1996	0.015 (0.045)	0.067 (0.150)	−0.012 (0.040)	0.135 (0.150)	0.006 (0.046)	0.118 (0.151)
Female	−0.045 (0.034)	−0.795*** (0.107)	−0.058* (0.030)	−0.734*** (0.105)	−0.040 (0.036)	−0.778*** (0.109)
Low education	0.021 (0.044)	0.057 (0.116)			0.007 (0.045)	0.052 (0.116)
High education	0.084** (0.040)	−0.242** (0.118)			0.113*** (0.041)	−0.153 (0.119)
Lowest standardized income tercile	−0.036 (0.042)	0.142 (0.118)			−0.046 (0.043)	0.078 (0.119)
Highest standardized income tercile	0.035 (0.041)	0.024 (0.119)			0.053 (0.042)	0.049 (0.119)
Single	0.003 (0.056)	0.280* (0.155)			−0.041 (0.058)	0.112 (0.157)
Widowed	−0.099 (0.114)	0.081 (0.177)			−0.103 (0.116)	0.045 (0.177)
Smoker			0.092*** (0.033)	0.691*** (0.102)	0.113*** (0.038)	0.666*** (0.103)
Good health			−0.356*** (0.045)	−0.447*** (0.116)	−0.370*** (0.052)	−0.436*** (0.116)
Heavy drinker			0.170*** (0.057)	0.316** (0.144)	0.161** (0.065)	0.321** (0.145)
Overweight			0.011 (0.032)	0.018 (0.099)	0.024 (0.037)	0.001 (0.100)
Obese			0.121* (0.065)	0.470*** (0.166)	0.159** (0.075)	0.443*** (0.167)
Chronic illness			0.082** (0.040)	0.276** (0.111)	0.090** (0.046)	0.260** (0.112)
Constant	−13.755*** (0.147)	−10.903*** (0.870)	−13.615*** (0.147)	−11.165*** (0.870)	−13.676*** (0.158)	−11.206*** (0.877)
Shape parameter (γ)	0.013*** (0.000)	0.009*** (0.001)	0.013*** (0.000)	0.009*** (0.001)	0.013*** (0.000)	0.009*** (0.001)
p value Wald test: education	0.107	0.048	–	–	0.014	0.268
p value Wald test: education, income, and marital status	0.081	0.024	–	–	0.002	0.628
p value Wald test: BMI			0.180	0.015	0.101	0.023

$N = 4434$; no. of failures = 463; standard errors are given in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6 Comparison of objective and subjective predicted remaining life durations (means)

Characteristics	Objective (1)	Subjective (2)	Difference (2) − (1)
Reference male[a]	42.60	39.14	−3.46
Female	49.83	39.41	−10.42***
Male smoker	36.50	38.40	1.9*
Obese male	38.42	38.06	−0.36
Male in bad health	38.60	36.70	−1.9
Heavy drinking male	39.61	38.07	−1.54
Male with chronic illnesses	40.17	38.54	−1.63

[a] The reference is a 45-year-old married man, born in 1950, living in a middle-income household, a non-smoker and non-heavy drinker, of normal weight and medium education, and in good health with no chronic illnesses, who reported his SSPs in 1995. After simulating the coefficient estimates of the objective and subjective models with 1000 replications, we then used them to calculate the difference between objective and subjective predicted remaining life durations. To find simulated coefficient estimates, we first drew uniform random variables and then used inverse CDF transformation and Cholesky decomposition to identify normally distributed random variables (see, e.g., Law and Kelton 1982)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The third and the fourth models show the associations between health indicators and subjective and objective mortality risk. Here, smoking, heavy drinking, obesity, having chronic illnesses, and being in good health explain subjective mortality risk, but their associations with subjective survival chances are not as strong as those with objective survival. In addition, a Wald test (not reported in Table 5) indicates that the smoking coefficient in the objective model is statistically higher than that in the subjective model at a 10% level of significance, suggesting that individuals underestimate the health risks of smoking. For the other health indicators, we cannot reject the null hypotheses of equal associations in both the subjective and objective models.

In the last two models in Table 5, which control for both socioeconomic variables and health indicators, the coefficients on smoking, heavy drinking, good health, obesity, chronic illnesses, and the shape parameter of the hazard function have the same sign and are all significant. However, the results of a Wald test (not reported in Table 5) indicate that the shape parameter and the associations between female and smoking in the subjective model are statistically different from those in the objective model. On the other hand, we can find no significant differences between the parameters for the other variables.

Table 6 reports the predicted median remaining life durations by gender and by current health status when all other characteristics are held constant. We restrict the analysis to these characteristics because they have significant associations with actual mortality risk (see column 6, Table 5). As a reference individual (or group), we take a 45-year-old married man, born in 1950, living in a middle-income household, a non-smoker, non-heavy drinker, of normal weight and medium education, and in good health with no chronic illnesses, who reported his SSPs in 1995. According to the table, for this reference male, the difference between subjective and objective predicted remaining life durations is about −3.5 years, a

difference insignificantly different from zero. We then change one characteristic of this reference individual at a time and report the results in the remaining rows. Of particular interest, we find a significant difference of −10.4 years for a 45-year-old woman, suggesting that women substantially underestimate their remaining life duration.

Overall, the differences between subjective and objective remaining lifetimes for obese, unhealthy, heavy drinking men, and men with chronic illnesses are relatively small and statistically insignificant, indicating that these subgroups tend to predict their remaining life durations accurately. Smoking men, on the other hand, significantly overestimate their remaining life duration by about 2 years, which implies that smokers underestimate the health risks of their smoking habit.[6]

5 Discussion

Overall, our use of a sample of Dutch respondents aged 25+ to analyze how well individual beliefs about survival chances predict actual mortality reinforces the findings of the US literature, which has relied primarily on individuals aged 50+. Our study has shown, for example, that in line with HRS-based evidence from Smith et al. (2001), Hurd and McGarry (2002), and Siegel et al. (2003), SSPs in the DNB-HS strongly predict actual mortality even when controls are in place for a large set of health indicators, including self-reported health. This finding suggests that SSPs contain information other than an individuals' current health status. For example, although the age at which a parent died may not affect respondents' current health, it may alter their expectations about the onset of a genetically linked disease (Hurd and McGarry 2002).

The above finding underscores the importance of eliciting SSPs in household surveys since the traditional questions on health status cannot provide this information on individual expectations. Moreover, as argued by Hurd (2009), probabilistic expectations have an advantage over non-probabilistic expectations in that the former is properly scaled, making them more easily comparable across different individuals. Non-probabilistic expectation measures, on the other hand, involve verbal statements such as 'likely' or 'unlikely,' which could be interpreted differently by different individuals. Another advantage of SSPs is that they can be used directly in life cycle models of consumption, savings, and retirement in which individuals' decisions depend on their own mortality expectations. Because life table survival probabilities are homogenous across individuals with different socioeconomic and health status, SSPs are a better instrument for measuring individual mortality risk in these models.

Our analysis has also confirmed the associations between actual mortality and income, education, and marital status found by Van Kippersluis et al. (2009) and Kalwij et al. (2013) for the Netherlands, albeit here less precisely estimated, possibly because of the relatively smaller sample. We have also shown that on

---

[6] When taking women as a reference group, we as well find an underestimation of the health risks of smoking.

average, male survival expectations conform closely to realized mortality, whereas females are overly pessimistic about their survival chances, a finding noted by Perozek (2008) for the US based on fitted survival functions, and Teppa and Lafourcade (2013) for the Netherlands based on actuarial survival probabilities. Additionally, in line with the findings of Khwaja et al. (2007), Hurd (2009), and Bissonnette et al. (2014), smokers in the DNB-HS sample exhibit optimistic survival beliefs, an overestimation of life expectancy that may result from SSPs containing expectations about quitting smoking in the near future (Khwaja et al. 2007). On the other hand, in contrast to Elder (2013), we found that SSPs yield a steeper age profile of mortality risk than actual mortality, suggesting that the old overestimate their mortality risk.

Our analysis also identified discrepancies between SSPs and objective values that may result from individuals' biased survival beliefs which may influence economic decisions. For instance, pessimistic survival expectations may lead to a higher probability of early retirement, a higher incidence of purchasing life insurance[7] or inadequate savings for retirement. They might also have some implications for public health. For example, women with highly pessimistic survival beliefs may underestimate the years they will need health care. In particular, they may underestimate the years they will live without a spouse and/or depend on caretakers outside the home, which would increase health care expenditures and reduce individual well-being at advanced ages.

These findings imply that future research should examine whether incorrect beliefs about the remaining lifetime translate into welfare-reducing behavior in the context of financial decisions, retirement timing, health care needs and similar situations in which lifetime uncertainty is relevant. More investigation is also needed into why women underestimate their life expectancy far more than men. One possible explanation, according to Elder (2013), is that SSPs have a mean-reverting property; that is, probability answers are biased toward 50%. If female probabilities are farther away from 50% than male probabilities, female SSPs will be more biased. Another explanation is a higher prevalence of pain (Banks et al. 2009) and/ or disability (Nusselder and Looman 2004) among women in the Netherlands, which despite our controlling for health indicators to some extent may still affect our results. If women incorrectly believe that their high rates of pain and disability forecast a higher mortality risk, they are likely to be more pessimistic than men about their survival chances. Nonetheless, Elder (2013), like Hudomiet and Willis (2013), calls for caution when interpreting and using SSPs because of the risk of measurement error, which must also be taken into account.

Finally, it would also be worth investigating whether individuals' assessments of their own life expectancy are influenced by their knowledge of actuarial survival probabilities. Some individuals, for example, may have no idea how long people of their own age and sex live on average or may form incorrect beliefs about population life expectancy. If those who underestimate population life expectancy do indeed tend to underestimate their own life expectancy, then a public policy that

---

[7] The holder of a life insurance policy pays a yearly premium while alive for a certain sum that is then inherited by legal heirs, making it an advantageous purchase for someone whose life is short.

aims at informing about population life expectancies should help them give more accurate answers to probabilistic survival questions.

Appendix

Derivation of the Median Remaining Life Duration

For each individual in our sample, we observe two survival function values, $SSP_{1,i}$ and $SSP_{2,i}$, at two different target ages $t_{1,i}$ and $t_{2,i}$ where $t_{1,i} < t_{2,i}$. $t_{0,i}$ is the baseline age at which the respondent reported the SSPs:

$$S_{1,i}(t_{0,i}, t_{1,i}) = \exp\left\{-\int_{t_{0,i}}^{t_{1,i}} \theta(s)ds\right\} = SSP_{1,i} \tag{6}$$

$$S_{2,i}(t_{0,i}, t_{2,i}) = \exp\left\{-\int_{t_{0,i}}^{t_{2,i}} \theta(s)ds\right\} = SSP_{2,i} \tag{7}$$

with $(t_{1,i}, t_{2,i}) \in \{(75, 80), (80, 85), (85, 90), (90, 95), (95, 100)\}$ and $(SSP_{1,i}, SSP_{2,i}) \in \{(P75, P80), (P80, P85), (P85, P90), (P90, P95), (P95, P100)\}$, respectively. $P75$ represents the subjective survival probability (SSP) to age 75, $P80$ that to age 80, and so on. The Gompertz hazard rate function is given by $\theta(t) = \lambda_i \exp\{\gamma_i t\}$.

After substituting the hazard rate function into Eqs. (6) and (7), we evaluate the integral to find

$$S_{1,i}\left(\gamma_i, \lambda_i \middle| t_{0,i}, t_{1,i}\right) = \exp\left\{\frac{\lambda_i}{\gamma_i}\left(\exp\{\gamma_i t_{0,i}\} - \exp\{\gamma_i t_{1,i}\}\right)\right\} = SSP_{1,i} \tag{8}$$

$$S_{2,i}\left(\gamma_i, \lambda_i \middle| t_{0,i}, t_{2,i}\right) = \exp\left\{\frac{\lambda_i}{\gamma_i}\left(\exp\{\gamma_i t_{0,i}\} - \exp\{\gamma_i t_{2,i}\}\right)\right\} = SSP_{2,i} \tag{9}$$

Next, following Perozek (2008), we take logarithms of the survival functions [Eq. (10)] and estimate the parameters γ_i and λ_i for each individual using nonlinear least squares (NLLS). This procedure requires that we replace the survival probabilities of 0 and 1 with slightly different numbers, namely 0.01 and 0.99, respectively.

$$\ln(SSP_{j,i}) = \ln(S_{j,i}(\gamma_i, \lambda_i | t_{0,i}, t_{j,i})) + \varepsilon_{j,i}, \quad j \in \{1, 2\} \tag{10}$$

where $\varepsilon_{j,i}$ is the error term, which is assumed to be independent and identically distributed (i.i.d.) and have a zero mean. We obtain the NLLS estimates of γ_i^* and λ_i^* by minimizing the following expression:

$$\min_{\gamma_i, \lambda_i} \sum_j \left(\ln(SSP_{j,i}) - \ln(S_{j,i}(\gamma_i, \lambda_i | t_{0,i}, t_{j,i})) \right)^2$$

with

$$S_{j,i}(\gamma_i, \lambda_i | t_{0,i}, t_{j,i}) = \exp\left\{ \frac{\lambda_i}{\gamma_i} \left(\exp\{\gamma_i t_{0,i}\} - \exp\{\gamma_i t_{j,i}\} \right) \right\}$$

We then calculate the subjective median remaining life duration conditional on baseline age for each individual denoted by $L_i^{R,S}$, where R and S denote 'remaining' and 'subjective,' respectively, based on the following formula:

$$S(L_i^{R,S} | t_{0,i}) = \exp\left\{ -\int_0^{L_i^{R,S}} \theta(s' + t_{0,i}) ds' \right\} = 0.5 \tag{11}$$

The Gompertz hazard function is then $\theta(s' + t_{0,i}) = \lambda_i \exp\{\gamma_i(s' + t_{0,i})\}$, so evaluating the integral in Eq. (11) and taking the natural logarithm of both sides yields

$$-\ln(0.5) = \frac{\lambda_i}{\gamma_i} \left(\exp\{\gamma_i(t_{0,i} + L_i^{R,S})\} - \exp\{\gamma_i t_{0,i}\} \right)$$

$$L_i^{R,S} = \frac{1}{\gamma_i} \ln\left(\frac{\gamma_i \ln(2)}{\lambda_i \exp(\gamma_i t_{0,i})} + 1 \right) \tag{12}$$

We then replace γ_i and λ_i in Eq. (12) with their estimates γ_i^* and λ_i^*, respectively. Because the variable $L_i^{R,S}$ is created using individual subjective survival probabilities, it represents the subjective median remaining life duration conditional on baseline age for each individual.

Comparison of Objective and Subjective Predicted Remaining Life Durations (Table 6)

The common assumption in the objective and subjective mortality models is that life duration can be modeled using a Gompertz distribution. Under this assumption, the hazard function can be written as

$$\theta(t|\mathbf{x_i}) = \lambda_i \exp\{\gamma_i t\}$$

where $\lambda_i = \lambda^{obj} = \exp\{\mathbf{x_i}\boldsymbol{\beta}^{obj}\}$, $\gamma_i = \gamma^{obj}$ in the objective mortality model, and $\lambda_i = \lambda^{subj} = \exp\{\mathbf{x_i}\boldsymbol{\beta}^{subj}\}$, $\gamma_i = \gamma^{subj}$ in the subjective mortality model. Hence, we

replace λ^{obj}, γ^{obj}, λ^{subj}, and γ^{subj} in Eq. (12) with their estimates, $\hat{\lambda}^{obj}, \hat{\gamma}^{obj}, \hat{\lambda}^{subj}$, and $\hat{\gamma}^{subj}$, respectively:

$$\hat{L}^{R,O} = \frac{1}{\hat{\gamma}^{obj}} \ln \left(\frac{\hat{\gamma}^{obj} \ln(2)}{\exp\left\{ \bar{\mathbf{x}}_i \hat{\beta}^{obj} + \hat{\gamma}^{obj} \bar{t}_{0,i} \right\}} + 1 \right) \tag{13}$$

$$\hat{L}^{R,S} = \frac{1}{\hat{\gamma}^{subj}} \ln \left(\frac{\hat{\gamma}^{subj} \ln(2)}{\exp\left\{ \bar{\mathbf{x}}_i \hat{\beta}^{subj} + \hat{\gamma}^{subj} \bar{t}_{0,i} \right\}} + 1 \right) \tag{14}$$

where $\hat{L}^{R,O}$ and $\hat{L}^{R,S}$ denote objective and subjective predicted remaining life durations, respectively, and $\bar{t}_{0,i}$ is the baseline age, which is equal to 45. The vector $\bar{\mathbf{x}}_i$ contains fixed values of the control variables included in the estimation. For example, in the first row of Table 6, all of the dummy variables in both mortality models are equal to zero.

References

Alessie, R., Hochguertel, S., & van Soest, A. (2002). Household portfolios in the Netherlands. In L. Guiso, M. Haliassos, & T. Jappelli (Eds.), *Household portfolios* (pp. 341–388). Cambridge, MA: MIT Press.

Banks, J., Kapteyn, A., Smith, J. P., & van Soest, A. (2009). Work disability is a pain in the ****, especially in England, the Netherlands, and the United States. In D. Cutler & D. Wise (Eds.), *Health in older ages: The causes and consequences of declining disability among the elderly*. Chicago: Chicago University Press.

Bissonnette, L., Hurd, M. D., & Michaud, P. C. (2014). Individual survival curves comparing subjective and objective mortality risks. IZA Discussion Paper No., 8658.

Bucher-Koenen, T., & Kluth, S. (2012). Subjective life expectancy and private pension. MEA Discussion Paper No. 14–2012.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B, 34*(2), 187–220.

Delavande, A., & Rohwedder, S. (2011). Differential mortality in Europe and the US: Estimates based on subjective probabilities of survival. *Demography, 48,* 1377–1400.

Elder, T. (2013). The predictive validity of subjective mortality expectations: Evidence from the health and retirement study. *Demography, 50*(2), 569–589.

Fischhoff, B., & Bruine de Bruin, W. (1999). Fifty-Fifty = 50%? *Journal of Behavioral Decision Making, 12,* 149–163.

Gan, L., Gong, G., Hurd, M., & McFadden, D. (2004). Subjective mortality risk and bequests. NBER Working Paper No. 10789.

Gerking, S., & Khaddaria, R. (2011). Perceptions of health risk and smoking decisions of young people. *Health Economics, 21,* 865–877.

Gompertz, B. (1825). On the nature of the function expressive of the Law of Human Mortality and on a new mode of determining life contingencies. *Philosophical Transactions of the Royal Society of London, 115,* 513–585.

Hamermesh, D. S. (1985). Expectations, life expectancy, and economic behavior. *Quarterly Journal of Economics, 100,* 389–408.

Hudomiet, P., & Willis, R. J. (2013). Estimating second order probability beliefs from subjective survival data. *Decision Analysis, 10*(2), 152–170.

Lusardi, A., & Mitchell, O. (2011). Financial literacy and planning: Implications for retirement wellbeing, forthcoming. In A. Lusardi & O. Mitchell (Eds.), *Financial literacy: Implications for retirement security and the financial marketplace* (pp. 17–39). Oxford: Oxford University Press.

Human Mortality Database. (2014). http://www.mortality.org/.

Hurd, M. D. (1989). Mortality risk and bequests. *Econometrica, 57*(4), 779–813.

Hurd, M. (2009). Subjective probabilities in household surveys. *Annual Review of Economics, 1,* 543–562.

Hurd, M. D., & McGarry, K. (1995). Evaluating subjective probabilities of survival in the Health and Retirement Study. *Journal of Human Resources, 30*(Supplement), S268–S292.

Hurd, M., & McGarry, K. (2002). The predictive validity of subjective probabilities of survival. *Economic Journal, 112,* 966–985.

Kalwij, A. S., Alessie, R. J. M., & Knoef, M. G. (2013). The association between individual income and remaining life expectancy at the age of 65 in the Netherlands. *Demography, 50,* 181–206.

Khwaja, A., Sloan, F. A., & Chung, S. (2007). The relationship between individual expectations and behaviors: Evidence on mortality expectations and smoking decisions. *Journal of Risk and Uncertainty, 35,* 179–201.

Kleinjans, K. J., & van Soest, A. (2013). Rounding, focal point answers, and nonresponse to subjective probability questions. *Journal of Applied Econometrics, 29*(4), 567–585.

Laibson, D., Repetto, A., Tobacman, J., Hall, R. E., Gale, W. G., & Akerlof, G. A. (1998). Self-control and saving for retirement. *Brookings papers on Economic Activity, 1998*(1), 91–196.

Law, A., & Kelton, W. (1982). *Simulation modelling and analysis*. New York: McGraw-Hill.

Lazarsfeld, P. F. (1940). Panel studies. *Public Opinion Quarterly, 4,* 122–128.

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

Nusselder, W. J., & Looman, C. W. N. (2004). Decomposition of differences in health expectancy by cause. *Demography, 41*(2), 315–334.

O'Donnell, O., Teppa, F., & Van Doorslaer, E. (2008). Can subjective survival expectations explain retirement behavior? DNB Working Paper No. 188.

Peracchi, F., & Perotti, V. (2011). Subjective survival probabilities and life tables: evidence from Europe. EIEF Working Paper No. 1016.

Perozek, M. (2008). Using subjective expectations to forecast longevity: Do survey respondents know something we don't know? *Demography, 45,* 95–113.

Poterba, J. M., Venti, S. F., & Wise, D. A. (2011). Family status transitions, latent health, and the post-retirement evolution of assets. In D. A. Wise (Ed.), *Explorations in the economics of aging* (pp. 23–69). Chicago: University of Chicago Press.

Rohwedder, S., & Van Soest, A. (2006). The impact of misperceptions about social security on saving and well-being, Michigan Retirement Research Center Working Paper No. 2006-11.

Salm, M. (2010). Subjective mortality expectations and consumption and savings behaviours among the elderly. *Canadian Journal of Economics, 43*(3), 1040–1057.

Schoenbaum, M. (1997). Do smokers understand the mortality effects of smoking? Evidence from the Health and Retirement Survey. *American Journal of Public Health, 87*(5), 755–759.

Siegel, M., Bradley, E. H., & Kasl, S. V. (2003). Self-rated life expectancy as a predictor of mortality: Evidence from the HRS and AHEAD surveys. *Gerontology, 49,* 265–271.

Siermann, C., van Teeffelen, P., & Urlings, L. (2004). Equivalentiefactoren 1995–2000: Methode en belangrijkste uitkomsten. *Sociaal-Economische Trends, 3,* 63–66.

Smith, K. V., Taylor, D. H., & Sloan, F. A. (2001). Longevity expectations and death: Can people predict their own demise? *American Economic Review, 91*(4), 1126–1134.

Sturgis, P., Allum, N., & Brunton-Smith, I. (2009). Attitudes over time: The psychology of panel conditioning. In P. Lynn (Ed.), *Methodology of longitudinal Surveys* (pp. 113–126). Chichester: Wiley.

Teppa, F., & Lafourcade, P. (2013). Can the longevity risk alleviate the annuitization puzzle? Empirical evidence from survey data. DNB Working Paper.

Van Kippersluis, H., O'Donnell, O., & van Doorslaer, E. (2009). Long run returns to education: does schooling lead to an extended old age? *Journal of Human Resources, 4,* 1–33.

Van Ooijen, R., Alessie, R., & Kalwij, A. (2014). Saving behavior and portfolio choice after retirement. Netspar Panel Paper No. 42, Network for Studies on Pensions, Aging and Retirement (www.netspar.nl).

Van Santen, P. (2013). Precautionary saving, wealth accumulation and pensions: an empirical microeconomic perspective. Ph.D. Thesis, *Netspar Theses*, 2013-001.

Wilson, D. L. (1994). The analysis of survival (mortality) data: Fitting Gompertz, Weibull, and logistic functions. *Mechanisms of Aging and Development, 74,* 15–33.

Wu, S., Stevens, R., & Thorp, S. (2015). Cohort and target age effects on subjective survival probabilities: Implications for models of the retirement phase. *Journal of Economic Dynamics and Control, 55*(June), 39–56.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association, 57,* 348–368.