Comparing observed and perceived joint survival probabilities of married couples

Lkhagvaa Erdenesuren*[1], Jochem De Bresser[1], and Arthur Van Soest[1,2]

[1] *Tilburg University, the Netherlands*

[2] *Netspar*

January 8, 2021

Abstract

We develop and estimate models for actual and perceived survival probabilities of married couples. In the model, the potential lifetime dependence between spouses in a couple is generated by correlated (un)observables that affect both spouses' survival probabilities, and by bereavement effects - the impact of one spouse's bereavement on the surviving spouse's survival rate. We use married couples' actual mortality experience and their survey reports of the probability to survive until given target ages from the Health and Retirement Study. We account for possible reporting biases in the elicited probabilities, jointly modelling couples' genuine subjective survival probabilities and their reporting behaviour. We find that actual and perceived lifetimes of the two spouses in a couple are strongly positively correlated. The bereavement effect explains around 60-98% of this correlation, with correlated observables and unobservables explaining the remainder. On average, non-widowed people younger than 75 (older than 75) tend to underestimate (overestimate) their actual expected remaining life time. Widowers substantially underestimate their expected life time, while widows tend to underestimate their expected life-years if they are younger than 72.

Keywords— Joint survival probability, Subjective survival expectation

JEL Codes— J11, J12, J19

*This work was funded by Instituut GAK through Netspar. If you have any comments regarding this work, please send an email to l.erdenesuren@tilburguniversity.edu

1 Introduction

In a life-cycle model with two-person households, forward-looking married couples' economic decisions - such as how much to save and consume, when to retire and leave a bequest, and what type of insurance product to purchase - hinge upon their perceived survival expectations (Hurd, 1989; Browning, 2000; Van der Klaauw & Wolpin, 2008; Michaud, Van Soest, & Bissonnette, 2019). Studying these life-cycle models empirically requires modeling of the couples' perceived joint survival probabilities. Traditionally, researchers used observed mortality data as a proxy for perceived survival probabilities, assuming that couples correctly assess their mortality risks (Casanova, 2010; Michaud et al., 2019). This assumption is controversial because couples' perceptions regarding their survival probabilities can systematically deviate from their actual survival probabilities. Indeed, acknowledging this controversy and with the increasing availability of data on subjective expectations, e.g. in the Health and Retirement Study (HRS) representing the 50+ population in the US, an increasing number of studies using subjective survival expectations have appeared in the past two decades (see, e.g., Smith, Taylor, and Sloan (2001); Van der Klaauw and Wolpin (2008); Iannario and Piccolo (2010); Ludwig and Zimper (2013); Bissonnette, Hurd, and Michaud (2017); Heimer, Myrseth, and Schoenle (2019)). The focus of these studies has been limited to develop models for individuals' marginal perceived survival probabilities, ignoring the potential remaining lifetime dependence between spouses in a couple.

In this paper, we develop an econometric model for actual as well as perceived joint survival probabilities of married couples. We estimate the model using 13 waves (from 1992 to 2016) of the HRS data.[1] The HRS is ideal in our context as it has accurate information on the respondents' actual mortality experience as well as their perceived survival expectations. The longitudinal nature of the data allows us to observe respondents' survival expectations before and after widowhood. Therefore, we can identify the actual and perceived survival rates of respondents conditional on their partners' vital status. To our knowledge, our paper is the first to analyze the joint perceived survival probabilities of married couples using subjective expectations held by survey respondents.

Our model accounts for two sources of dependencies between remaining lifetimes of married couples, both for actual and perceived survival models. The first is *the common-lifestyle effect*: spouses tend to have similar lifestyles because they are coupled based on their (observed or unobserved) similar characteristics that affect their actual and perceived survival probabilities, such as their habits, ethnicity, and educational backgrounds (Hollingshead, 1950; Burgess & Wallin, 1943). For example, smokers may attract each other, and smoker-couples tend to have lower actual survival rates than non-smoker-couples. The perceived survival rates of spouses in a smoker-couple can also be correlated because smokers tend to underestimate the negative impact of smoking on their mortality rates (Khwaja et al., 2007; Wang, 2014). The second source is *the bereavement effect*, also known as the

[1] Numerous studies used the reported survival expectation data of the HRS,e.g., Bissonnette et al. (2017); Khwaja, Sloan, and Chung (2007); Salm (2010); Gan, Gong, Hurd, and McFadden (2015).

broken-heart syndrome: the (negative) impact of one spouse's death on the surviving spouses' survival rates. Surviving spouses tend to encounter elevated stress, loneliness, depression, and anxiety after their spouses have passed away (Williams Jr, 2005; Wittstein et al., 2005). These mental and physical health problems reduce the actual survival rates of surviving spouses (Van den Berg, Lindeboom, & Portrait, 2011; Spreeuw & Owadally, 2013; Sanders & Melenberg, 2016). Moreover, surviving spouses' perceptions regarding their remaining lifetimes may change after spousal bereavement if they update their survival expectations against health shocks (Smith et al., 2001).

To model the common-lifestyle and bereavement effects, we follow the Markov-type models proposed by Freund (1961) and Gourieroux and Lu (2015). Freund's model captures the bereavement effect by allowing jumps in mortality hazards at the time of death of one spouse. Gourieroux and Lu (2015) extend Freund's model by allowing asymmetric reactions of the mortality hazard rates for male and female surviving spouses at the moment the first spouse passes away. They also mix the model with common unobservable factors to capture common lifestyle effects. We extend the model of Gourieroux and Lu (2015) in three ways. First, we allow that spouses share correlated but not necessarily perfectly correlated (or common) frailty terms. Second, we model that spouses' mortality hazards depend on certain observed characteristics that are known to be good predictors of people's actual and perceived survival probabilities. Third, we impose the same structure on both the actual and the perceived mortality hazard rates. Thus, our approach enables us to directly test whether people with the same characteristics, on average, correctly perceive their mortality risks.

We do not utilize the reported survival probabilities to directly infer couples' true perceived survival probabilities because of the following challenges. First, people tend to round their subjective responses to reflect their uncertainty regarding the underlying process (Manski, 2004; Manski & Molinari, 2010). People who are more uncertain regarding their survival rates are more likely to report probabilities that are multiples of larger integers, for example, 0%, 50%, and 100%. The second challenge is that if spouses in a couple are interviewed together, they might want to please their partners by mimicking each other's subjective responses (Aquilino, 1993; Aquilino, Wright, & Supple, 2000), thus influencing each other's reported expectations or each other's reporting behaviour. Spouses' reporting behaviour may also be correlated due to other reasons, such as assortative matching.

To overcome the impediments of using subjective reports, we model couples' reporting behaviour explicitly. Rounding is accounted for following the models proposed by De Bresser and Van Soest (2013); Kleinjans and Van Soest (2014), and De Bresser (2019). To control for spouses' influence on each others' responses, we allow for correlated observed and unobserved factors that affect the reporting behaviour of both spouses.

The main results of the paper are the following. In line with previous literature, we find that younger spouses underestimate their remaining lifetime relative to within-sample actual survival (Bissonnette et al., 2017; Heimer et al., 2019). In contrast, old couples overestimate their actual expected remaining life years on average. We find both actual and perceived remaining lifetimes of spouses in a couple are positively related. Keeping observed and unobserved characteristics constant, depending on their age, males (females) tend to live 0.6-5 additional years (0.5-4.4 years) longer if their partners are alive than if their partners are dead (the bereavement effects on actual mortality). Individuals whose partners are alive also perceive they would live longer than those whose partners are dead. Depending on people's age, males (females) perceive to live for 1.4-4.8 years (1.6-2.3 years) more if their partners are alive than if their partner is dead (the bereavement effects on perceived mortality). The bereavement effects explain 83%-98% of the actual remaining lifetime dependence, and 60%-90% of the perceived remaining lifetime dependence. The rest is explained by the correlations between the (un)observed characteristics between spouses.

The actual and perceived survival curves are predicted to have substantial dispersion in survival curves around their corresponding medians at a given age, gender, and partner's vital status. Most of the dispersion in actual survival curves is explained by the variations in observed characteristics, such as their ethnicity, education level, cohort, income, and health state. The unobserved frailties explain only a small part of the dispersion in actual survival curves. In contrast, both observed and unobserved factors explain substantial dispersion in perceived survival curves. Thus, at given observed characteristics, there is a substantial heterogeneity in people's survival expectations compared to their actual survival rates.

The estimation results of the reporting model suggest that people give more precise answers if they have a high cognitive ability. Moreover, the estimated correlation between the unobserved factors that affect the rounding behaviour of spouses in a couple is positive and significant. This evidence may imply that spouses in a couple have similar rounding behaviours due to assortative mating, or they influence each other's subjective responses during the interview.

The paper is structured as follows. Section 2 describes the data and discusses the survival expectation question in the HRS. Section 3 presents the joint survival probability model of married couples and the reporting model. Section 4 presents the estimation results. The final section concludes.

2 Data

2.1 Sample selection and descriptive statistics

We use biennial data of the Health and Retirement Study (HRS) from 1992 to 2016. The HRS is a national panel survey of individuals over age 50 and their spouses, and it is representative of the elderly population of the U.S. The HRS contains comprehensive information on each respondent's marital history, vital status, survival expectations, socioeconomic characteristics, and health indicators. Throughout our analysis, we define a respondent as *married* at a given survey wave if the partner is alive, and as *widowed* if the partner is dead. If both spouses in a couple responded that they are officially married or unmarried partners, we consider them as *married*.

Table 1 shows our sample selection procedure. We restrict our sample to couples in which both spouses were alive when they entered the survey. This selection excludes the sample of respondents who entered as a widow(er) or those who never married. To simplify, we drop the observations of respondents who were in a same-sex relationship. To avoid complications that arise due to marital dissolution, we only keep respondents who were neither divorced nor separated during the period of interest. Some respondents lost their first spouse and then married again during the sample period. We treat these respondents' survival dates are right-censored at the time they re-married, and discard their second spouse's observations.[2] We also exclude the observations of 36 couples in which both spouses died on the same day. We suspect that the cause of death of these couples is some unnatural cause, such as an accident. Finally, we keep couples in which both spouses participated for at least two waves and both responded to the survival expectation questions at least once.

Table 1: Sample selection

	number of households	number of individuals
Initial sample	26,598	42,053
Dropping respondents who were in the same-sex relationship	26,468	41,797
Dropping if divorced or separated	19,966	33,221
Keeping if entered the survey as married	13,540	26,578
Dropping observations of second, third, and fourth spouses	13,540	26,195
Keeping if participated in the survey for at least two waves	11,838	22,994
Dropping if a couple consists of only one active respondent	11,156	22,312
Keeping if provided a valid response to the survival expectation questions at least once	11,079	22,158
Dropping if spouses died on the same day	11,043	22,086

2.2 Observed mortality experience of couples

In our final sample, during the period of interest (1992-2016):

- 48.4% of couples did not encounter a spousal bereavement

[2] In our sample, people re-married at most four times. For those who re-married more than twice, we dropped the observations of their second, third, and fourth spouses.

- In 14.8% of couples, the male spouse died first, and the female spouse died second
- In 20.9% of couples, the male spouse died first, and the female spouse did not die
- In 8.3% of couples, the female spouse died first, and the male spouse died second
- In 7.6% of couples, the female spouse died first, and the male spouse did not die

More than 51% of couples have experienced a spousal bereavement between 1992 and 2016. In these couples, male spouses died first in around 69%[3] of all cases, due in part to the higher average entry age for males than for females (61 versus 58, on average). This also leads to higher mortality rates for male than for female spouses.

2.3 Survival questions in the HRS

In each wave, the HRS asks the following question:

...Pick an integer between 0 and 100 where "0" meaning no chance and "100" meaning absolutely sure. "What is the percent chance that you will live to be *ta* or more?"

Here *ta* is the target age, linked to the respondent's age. The difference between the target age and the age of respondents is at least 10 years, and the target age is always a multiple of 5. In each survey wave, respondents answered at most two survival expectation questions with different target ages. In 1992, the questions were asked with a scale of 0 to 10; in the other waves, the responses are scaled from 0 to 100. De Bresser (2019) finds that using scales from 0 to 10 or from 0 to 100 gives similar measurements of the subjective survival expectations. We multiply the responses in wave one by 10 to make them comparable to the responses in other waves.

The number of observations (respondent $\times$ wave $\times$ number of survival expectation questions at given wave) in our sample is 73,487 for males and 87,961 for females. Besides choosing responses between 0 and 100, respondents can also choose to answer "Don't know" or "Refuse" (DK/R). Since rather few responses are DK/R (roughly 4.5% of males and 5.8% of females of our sample), we do not consider the sample selection issue of ignoring these non-responses.

Table 2 shows the heterogeneity in respondents' reported survival probabilities for each target age and a given age-interval. The average reported survival probabilities are shown separately for married and widowed males and females. For each age interval, the mean for married respondents is higher than that of widow(er)s. The large magnitude of the standard deviations compared to their means imply substantial heterogeneity in people's perceptions regarding their subjective survival probabilities (or their reporting biases).

Column 2 of Table 2 shows the mean life-table probabilities as a measure of objective survival

[3] $\frac{20.9+14.8}{20.9+14.8+8.3+7.6} \approx 69.1$

probabilities. The HRS provides the annual life tables, controlling for age and sex, retrieved from the National Center for Health Statistics and the Berkeley Mortality Database. Average reported probabilities to target age 75 and 80 are lower than the corresponding mean life-table probabilities, for males and females and for married and widowed respondents. On the other hand, for target age 85, the average reported survival probabilities are slightly larger than those in the life-table. Hence, at least compared to the mean life-table probabilities, reported survival probabilities substantially differ from actual survival probabilities.

Figure 1 shows the empirical distribution of the reported probabilities to the subjective survival expectation questions, separately for males and females. We present histograms for target ages 80 and 85 as illustrating examples. The figure shows that most responses are concentrated at multiples of 10%, and the highest concentration is at 50%; thus, the reported probabilities suffer from rounding.[4] Substantial number of respondents chose 0% and 100% although these responses are unrealistic.

Table 2: Descriptive statistics of reported survival probabilities and mean life-table probabilities

		Mean Life-table	Married			Widowed		
	Current age		N	Mean	S.D.	N	Mean	S.D.
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Males								
Target age 75	64-66	73.45	2,995	65.12	27.9	132	64.05	29.76
Target age 80	68-70	63.53	2,598	58.53	28.33	234	55.41	29.98
Target age 85	73-75	47.22	2,933	53.43	29.59	301	47.79	32.21
Females								
Target age 75	64-66	82.32	2697	67.66	27.53	456	64.41	27.7
Target age 80	68-70	73.11	2970	61.63	28.3	791	58.22	29.84
Target age 85	73-75	59.73	2524	54.27	29	917	52.79	30.17

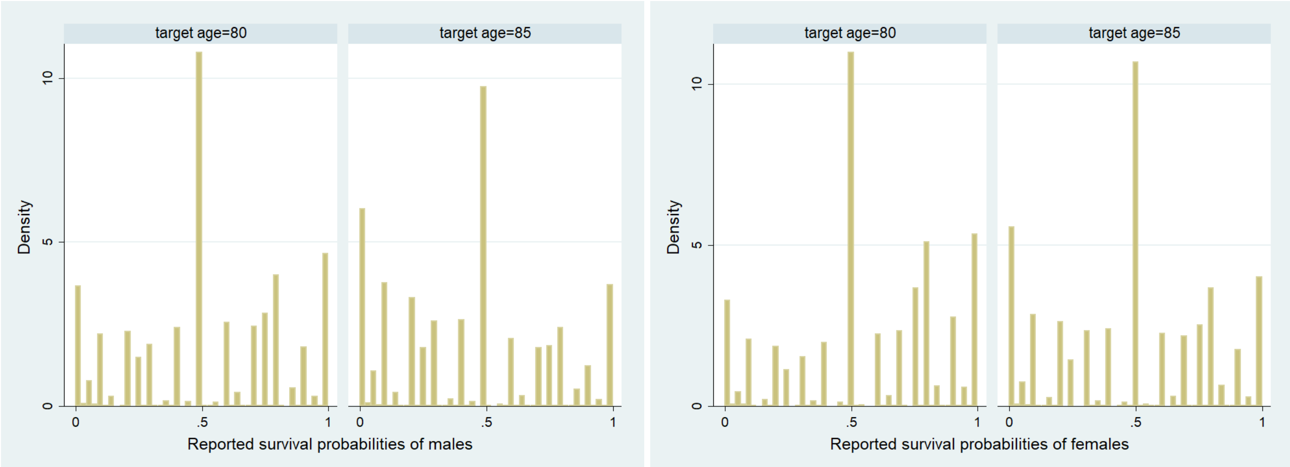
Note: Column (2) shows the sample average of the life-table probability of reaching a certain target age conditional on survived until the current age. The life-table probabilities are only available conditional on age and sex. Columns (4) and (7) show the sample average of the reported survival probabilities of reaching a certain target age conditional on survived until the current age for married and widowed respondents, respectively. The descriptive statistics of responses to target age 90 are not shown since very few people responded to the survival expectation question with target age 90.

2.4 Covariates

Definitions of the covariates used in the actual and perceived survival probability models are given in Panel A of Table 3. Table 4 shows some summary statistics. Since the survival probabilities are likely to vary with respondents' health state and income, we use the following covariates: functional

[4] Bruine de Bruin, Fischbeck, Stiber, and Fischhoff (2002) consider that people might report 50% to express their inability to reason in terms of probabilities, also known as epistemic uncertainty, rather than expressing their uncertainty regarding the underlying process. De Bresser and Van Soest (2013) and Kleinjans and Van Soest (2014) model the possibility that respondents answer 50% to reflect their epistemic uncertainty. Bissonnette et al. (2017) and De Bresser (2019) do not find any evidence of the existence of the epistemic uncertainty in subjective survival responses once the rounding behaviors are controlled. Therefore, in our model, we assume that the reported probability of 50% only expresses respondents' uncertainty regarding their true survival expectations, not their epistemic uncertainty.

Figure 1: Self-reported probabilities



(a) Males (b) Females

limitations in daily living, ever recorded chronic conditions, and household-size adjusted log income at the time they entered the survey. The advantage of using health and income indicator variables at the time of entry is that we can treat these variables as time-invariant, simplifying their use in our survival models (Bissonnette et al., 2017).We also control for basic demographics and socio-economic status, including dummies for whether a respondent is Black or Hispanic, whether a respondent's highest obtained degree is high-school or college, and cohort dummies.

As already shown in Figure 1, the reported probabilities suffer from rounding. Section 3.3.2 discusses how we model rounding behaviour in detail. In the reporting model, we use the covariates shown in Panels A and B of Table 3. As suggested by Manski and Molinari (2010), we use the reported probabilities of other subjective expectation questions than the survival expectation questions asked in the HRS in a given wave.[5] We expect that if respondents choose mostly 0%, 50%, or 100% in other subjective expectation questions, they are more likely to report coarse answers to the survival expectation questions also.

Moreover, Hurd, McFadden, and Gan (1998) and Lillard and Willis (2001) find that whether respondents report a coarse responses, such as 0%, 50%, or 100%, or more precise responses, such as multiples of 1 but not of 5, is correlated with their education and cognitive abilities. Thus, we use respondents' education degree, and their score on the immediate word recall test as proxy measures for cognitive capacity.[6]

[5] The number of expectation questions asked in the RAND version of the HRS varies from wave to wave. Besides survival expectation questions, the most common questions in the survey are (1) probability of receiving any inheritance, (2) probability of leaving any bequest, (3) probability of leaving bequest that values more than 10,000 \$, (4) probability of leaving bequest that values more than 100,000 \$, (5) probability of working full-time after age 62, (6) probability of working full-time after age 65, (7) probability of moving to nursing home in next 5 years, and (8) probability of having a work limiting health problem in next 10 years. Some probability questions were not asked depending on respondents' work, marital and health status, and age. Since the total number of subjective expectation questions vary across waves and across individuals, we construct the proportions of choosing $\{0\%, 50\%, 100\%\}$ to make this variables comparable across individuals and waves.

[6] The HRS asks respondents to memorize a particular list of words to measure their cognitive ability. It provides measures for immediate and delayed word recalls, which are counts of the number of correctly recalled words from a 10 or 20-word list. We divide the number of correctly recalled words by the total number of words in a list to

Table 3: Variable descriptions

Variable		Description
Panel A.		
Log of income	time-invariant	Log of real total household income in 2016 U.S. dollars divided by square root of household size[a]
Ever smoked	time-invariant	=1 if ever smoked or current smoker
Functional limitations	time-invariant	Number of functional limitations in daily living[b]
Chronic conditions	time-invariant	Number of chronic conditions ever recorded[c]
High school	time-invariant	=1 if highest obtained degree is high school or GED
College	time-invariant	=1 if having at least college or graduate degrees
Hispanic	time-invariant	=1 if Hispanic
Black	time-invariant	=1 if Black
Cohort before 1930	time-invariant	=1 if born before 1930
Cohort 1931-1945	time-invariant	=1 if born between 1931 and 1945
Panel B.		
Prop. multiples 50	time-varying	proportion of choosing responses of 0, 50 and 100 in other subjective expectation questions
Immediate word recall	time-varying	proportion of number of immediate word-recalls

[a] : We use the imputed version of the household income available in the RAND version of the HRS. The household size counts the number of living spouses plus the number of children living together.
[b] : Respondents were asked whether they have difficulty engaging with the following five activities: (1) walking across a room, (2) getting in and out of bed, (3) dressing, (4) bathing, and (5) eating. Depending on the number of activities respondents find to have difficulties with, the variable takes values from 0 to 5.
[c] : The number of chronic conditions respondents ever had. The following chronic conditions include (1) diabetes, (2) cancer, (3) lung disease, (4) heart disease, (5) stroke, and (6) arthritis, (7) high blood pressure and (8) psychiatric problems. Depending on the number of chronic conditions that respondents ever had, the variable takes values from 0 to 8.

Table 4: Summary statistics

	Male				Female			
	mean	st.dev	min	max	mean	st.dev	min	max
Panel A.								
Log of income	10.252	1.065	0	14.297	10.247	1.056	0	14.297
Ever smoked	0.697	0.459	0	1	0.461	0.499	0	1
Functional limitations	0.195	0.691	0	5	0.185	0.6632	0	5
Chronic conditions	1.193	1.175	0	7	1.109	1.127	0	8
High school	0.316	0.465	0	1	0.378	0.485	0	1
College	0.426	0.495	0	1	0.407	0.491	0	1
Hispanic	0.113	0.317	0	1	0.117	0.322	0	1
Black	0.131	0.338	0	1	0.128	0.334	0	1
Cohort before 1930	0.312	0.463	0	1	0.233	0.423	0	1
Cohort 1931 - 1945	0.383	0.486	0	1	0.389	0.487	0	1
Panel B.								
Prop. multiples 50	0.648	0.336	0	1	0.647	0.335	0	1
Immediate word recall	0.492	0.189	0	1	0.554	0.19	0	1

Note: We use one observation from each household to calculate descriptive statistics of the variables in Panel A since they are time-invariant. We use all available wave×respondent observations to calculate the descriptive statistics of they variables in Panel B since the are time-varying.

make these variables comparable across waves. In couple-households, each spouse was assigned a different list. The words in the list are different in each wave.

3 Econometric model and estimation strategy

In this section, we discuss our econometric model for the actual and perceived survival probabilities. We use *the objective model* to refer to the part of the model explaining actual mortality, and *the subjective model* for the part explaining perceived survival probabilities.

Section 3.1 discusses the general structure of the joint remaining lifetime distribution of spouses in a couple. Section 3.2 shows the likelihood contribution of objective model. Section 3.3 consists of two subsections. In the first part, we link the perceived survival probability structure in Section 3.1 to the conditional survival probabilities in the survey. In the second part, we discuss the reporting model. This describes our strategy to disentangle the true subjective survival from the reported survival probabilities by controlling for both spouses' (correlated) rounding behaviours. Section 3.4 links the likelihoods of the objective and the subjective models, allowing the unobserved factors of these to be correlated. In the second part of Section 3.4, we discuss our estimation strategy.

3.1 Joint distribution of the remaining lifetimes

Consider a couple with two spouses: spouse 1 (male), and spouse 2 (female). In our sample, spouses are first observed when the older spouse is aged 50 or older. Thus, we set the first time the spouses are at risk of mortality to the point in time when the older spouse reaches age 50. Moreover, we assume that spouses are a couple since before the older spouse reaches age 50.

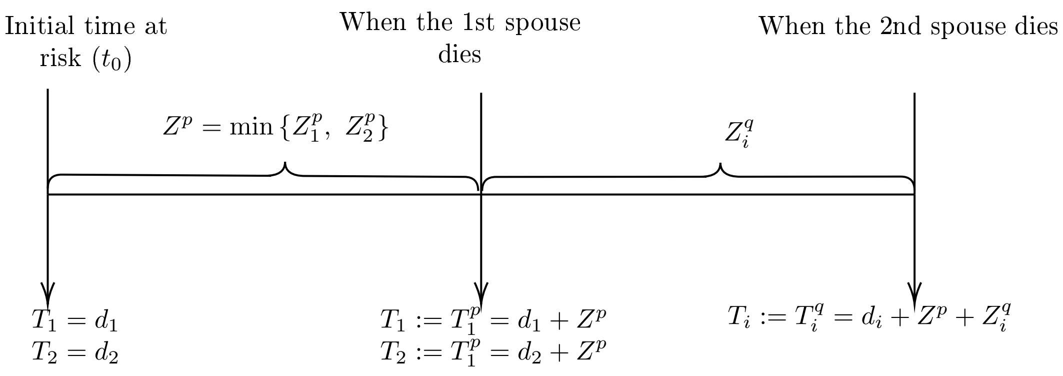


Figure 2: Potential lifetimes of spouses

Note: After the death of the first spouse, spouse i is a surviving spouse. $i = 1$ (resp. $i = 2$) if the male (resp. the female) spouse is a surviving spouse.

Figure 2 illustrates the potential lifetimes of the two spouses. d_1 and d_2 are their ages at the initial time of risk (t_0); if spouse 1 is older than spouse 2, then $d_1 = 50$ and $d_2 \leq 50$. Z_1^p and Z_2^p are random variables representing the lifetimes of spouses 1 and 2 as of t_0, when their spouse is alive. The first spouse dies at time $t_0 + Z^p = t_0 + \min(Z_1^p; Z_2^p)$. After one spouse's death, there can be a change in the surviving spouse's residual lifetime distribution due to the bereavement effect. Let the remaining lifetime of spouse 1 (spouse 2) after the death of other spouse be Z_1^q (Z_2^q). Then the total lifetimes after t_0 of both spouses are:

$$T_1 = d_1 + Z^p + Z_1^q \cdot 1\{Z^p = Z_2^p\} \tag{1}$$

$$T_2 = d_2 + Z^p + Z_2^q \cdot 1\{Z^p = Z_1^p\} \tag{2}$$

Here $1\{\cdot\}$ is the indicator function (1 if the argument is true and 0 otherwise). Our model does not explicitly capture the possibility that both spouses die at the same time.

Let $x = (x_1, x_2)$ be a vector of observed initial conditions or other time-invariant regressors (not including an intercept) and let $\eta = (\eta_1, \eta_2)$ be time-invariant unobserved frailty terms driving the distributions of $(Z_1^p, Z_2^p, Z_1^q, Z_2^q)$. Spouses' remaining lifetimes can be correlated via the observed regressors and the frailty terms. We assume that at t_0, the population distribution of unobserved frailty terms $\eta = [\eta_1, \eta_2]$ is bivariate normal with mean 0 and arbitrary covariance matrix Σ. We assume Z_1^p, Z_2^p, Z_1^q and Z_2^q are mutually independent conditional on observed regressors and unobserved frailty terms.

Following the conditional independence assumption, the joint distribution of $T_1, T_2|(d, x, \eta)$ can be fully characterized by the conditional hazard rates of each spouse at each state. We impose the mixed-proportional hazards (MPH) structure on the conditional hazard rates.[7] Moreover, we assume that the baseline hazard rates follow a Gompertz specification. Under this specification, the shape of the baseline hazard depends on two unknown parameters: an intercept and a duration dependence parameter. Previous studies have shown that the Gompertz specification approximates the mortality rates of the elderly reasonably well (Frees, Carriere, & Valdez, 1996; Carriere, 2000; Luciano, Spreeuw, & Vigna, 2008).

The hazard rates driving $Z_i^p, i = 1, 2$ at age t_i^p, conditional on (x_i, η_i), are:

$$\lambda^p(t_i^p|x_i, \eta_i) = \underbrace{\exp(\alpha_i^p t_i^p + \beta_i^p)}_{\text{baseline hazard rate}} \cdot \exp(x_i\beta_i + \eta_i), \text{ where } d_i \leq t_i^p < d_i + z^p \tag{3}$$

These are the relevant hazard rates as of t_0 as long as both partners are alive. Before t_0, i.e., for $t_i^p < d_i$, the hazard rate is set to 0. Let the corresponding integrated hazards from age a_i to age t_i^p be $\Lambda^p(t_i|a_i, x_i, \eta_i)$.

Suppose that spouse i lives longer than the partner, so $Z^p = Z_{3-i}$. Then the hazard rate of widow(er) i at age $t_i \geq d_i + Z^p$, conditional on (x_i, η_i) is:

$$\lambda^q(t_i|x_i, \eta_i) = \underbrace{\exp(\alpha_i^q t_i + \beta_i^q)}_{\text{baseline hazard rate}} \cdot \exp(x_i\beta_i + \eta_i), \text{ where } d_i + Z^p \leq t_i \tag{4}$$

The corresponding integrated hazard rate from age a_i to t_i, where $d_i + Z^p \leq a_i \leq t_i$, is $\Lambda^q(t_i|a_i, x_i, \eta_i)$. In other words, bereavement changes α_i^p and β_i^p into α_i^q and β_i^q. Thus, our specification captures the

[7] The MPH models are one of the most widely used duration models in econometrics (Van den Berg, 2001).

bereavement effect by allowing slope and level shifts in the surviving spouse's log baseline hazard rates.

Together with the conditional independence assumptions, the four hazard rates determine the joint distribution of the lifetimes of both partners given (x_i, η_i). For example, the joint survival probability of reaching ages $t_1 = d_1 + z$ and $t_2 = d_2 + z$ conditional on (d, x, η) is:

$$S_0(t_1, t_2|d, x, \eta) = P(Z_1^p > z, Z_2^p > z|d, x, \eta) \tag{5}$$

$$= P(Z_1^p > z|d_1, x_1, \eta_1) \cdot P(Z_2^p > z|d_2, x_2, \eta_2) \tag{6}$$

$$= P(T_1^p > t_1|d_1, x_1, \eta_1) \cdot P(T_2^p > t_2|d_2, x_2, \eta_2) \tag{7}$$

$$= \exp(-[\Lambda^p(t_1|d_1, x_1, \eta_1) + \Lambda^p(t_2|d_2, x_2, \eta_2)]) \tag{8}$$

The probability that spouse i survives until at least age $t_i > t_i^p = d_i + z^p$ as a widow, given that spouse $3-i$ died at age $t_{3-i}^p = d_{3-i} + z^p$ and was the first who died, and given (d_i, x_i, η_i), is:

$$S_{i|3-i}(t_i|t_i^p, x_i, \eta_i) = P(T_i^q > t_i|T_i^q > t_i^p, x_i, \eta_i) \tag{9}$$

$$= \exp(-\Lambda^q(t_i|t_i^p, x_i, \eta_i)) \tag{10}$$

In the next two subsections, we will discuss our strategies to construct objective and subjective likelihoods. We use η^o and η^s to denote the bivariate unobserved frailty terms of the objective and subjective hazard rates.

3.2 Objective joint survival model

The objective survival probability model is estimated using the observed mortality data. The spouses in a couple entered the survey at ages $a^* = (a_1^*, a_2^*)$, and they are last observed at ages $t^* = [t_1^*, t_2^*]$. If spouse i is dead at the last observed age, then $t_i^* = T_i^o$, otherwise his/her remaining lifetime is right-censored. Left-truncation occurs in our data-set because couples' entry ages (a_1^*, a_2^*) can be higher than their ages at the initial time of risk (d_1, d_2). (t_1^p, t_2^p) are ages of spouses when the first death occurs in a couple.

At the last observed period each couple is in one of the following states: (1) husband died first, and wife died second, (2) husband died and wife survived, (3) husband died second, and wife died first, (4) wife died and husband survived, or (5) both spouses are alive. Depending on a couple's state at the last observed period, the conditional likelihood is defined as follows:

$L^o(\gamma^o|d,x,\eta^o) =$

$$\begin{cases} \lambda_1^p(t_1^*|x_1,\eta_1^o)\lambda_2^q(t_2^*|x_2,\eta_2^o)S_0(t_1^*,t_2^p|d,x,\eta^o)S_{2|1}(t_2^*|t_2^p,x_2,\eta_2^o) & \text{if husband died 1st, wife 2nd,} \\ \lambda_1^p(t_1^*|x_1\eta_1^o)S_0(t_1^*,t_2^p|d,x,\eta^o)S_{2|1}(t_2^*|t_2^p,x_2,\eta_2^o) & \text{if husband died and wife alive,} \\ \lambda_1^q(t_1^*|x_1,\eta_1^o)\lambda_2^p(t_2^*|x_2,\eta_2^o)S_0(t_1^p,t_2^*|d,x,\eta^o)S_{1|2}(t_1^*|t_1^p,x_1,\eta_1^o) & \text{if wife died 1st, husband 2nd,} \\ \lambda_2^p(t_2^*|x_2,\eta_2^o)S_0(t_1^p,t_2^*|d,x,\eta^o)S_{1|2}(t_1^*|t_1^p,x_1,\eta_1^o) & \text{if wife died and husband alive,} \\ S_0(t_1^*,t_2^*|d,x,\eta^o) & \text{if both survived} \end{cases} \quad (11)$$

at the last observed period. Here γ^o is a vector of unknown parameters that includes α_i^p, α_i^q, β_i^p, β_i^q, and β_i, where $i = 1, 2$.

3.3 Subjective joint survival model and reporting model

3.3.1 True subjective survival probability model

This section models how married and widowed respondents formulate their survival probability to reach a certain target age conditional on their age at a given survey wave. As the distribution of the remaining lifetimes of married and widowed can be different due to the bereavement effect, married and widowed respondents follow different structures to formulate their survival expectations.

Married respondents

Let $a = (a_1, a_2)$ be the ages of spouses at the time of the survey, and ta_i be target age of the survival question asked from spouse i in that wave. Without loss of generality, we consider the case where spouse i formulates his/her survival expectation of reaching age ta_i. Let spouse j be the partner of spouse i such that $j = 3 - i$. The age of spouse j at the time spouse i reaches age ta_i is $ta_j = ta_i + d_j - d_i$.

Married respondents encounter a risk of losing their partners; thus, they take into account the potential consequences of losing their partners on their survival chances when they formulate their survival expectations. We assume that spouses are aware that if both spouses are alive at ages (ta_1, ta_2) the conditional hazard rate of spouse i is $\lambda_i^p(\tau_i|x_i,\eta_i^s)$ when $a_i \leq \tau_i \leq ta_i$. However, if spouse j dies when i had age t_i^p and spouse i has survived at age ta_i, spouse i's conditional hazard rate is $\lambda_i^p(\tau_i|x_i,\eta_i^s)$ when $a_i \leq \tau_i < t_i^p$, and $\lambda_i^q(\tau_i|x_i,\eta_i^s)$ when $t_i^p \leq \tau_i \leq ta_i$.[8]

Following this structure, spouse i's survival probability of reaching age ta_i conditional on $(T_1^s >$

[8] These assumptions are necessary to identify the model and cannot be relaxed without further information on respondents' survival expectations. First, to relax these assumptions, we need information on how married respondents perceive their survival chances conditional on their partner is alive or dead in all future periods. Next, we need information on how respondents perceive their partners' survival chances (aka. cross-expectations). Kapteyn and Kooreman (1992) and Michaud et al. (2019) show how to measure cross-expectations and identify the joint expectations of married couples from each spouse's perspective.

$a_1, T_2^s > a_2, x, \eta^s)$ is defined by the following equation:

$$
\begin{aligned}
&P(T_i^s \geq ta_i | T_1^s \geq a_1, T_2^s \geq a_2, d, x, \eta^s) \\
&\quad = \frac{P(T_i^s \geq ta_i \cap T_j^s \geq a_j | d, x, \eta^s)}{P(T_1^s \geq a_1 \cap T_2^s \geq a_2 | d, x, \eta^s)} = \frac{\int_{a_j}^{\infty} P(T_i^s \geq ta_i \cap T_j^s = \tau_j | d, x, \eta^s) d\tau_j}{P(T_1^s \geq a_1 \cap T_2^s \geq a_2 | d, x, \eta^s)} \\
&\quad = \frac{\int_{a_j}^{ta_j} P(T_i^s \geq ta_i \cap T_j^s = \tau_j | d, x, \eta^s) d\tau_j + P(T_1^s \geq ta_1 \cap T_2^s \geq ta_2 | d, x, \eta^s)}{P(T_1^s \geq a_1 \cap T_2^s \geq a_2 | d, x, \eta^s)} \\
&\quad = \frac{\int_{a_j}^{ta_j} \lambda_j^p(\tau_j | x_j, \eta_j^s) S_0(\tau_1, \tau_2 | d, x, \eta^s) S_{i|j}(ta_i | \tau_i, x_i, \eta_i^s) d\tau_j + S_0(ta_1, ta_2 | d, x, \eta^s)}{S_0(a_1, a_2 | d, x, \eta^s)}
\end{aligned} \tag{12}
$$

where $\tau_i = \tau_j + a_i - a_j$. The first line of Eq. 12 is obtained using the Bayes rule. In the second line, the integral in the numerator is separated into two additive terms: the first is integral where τ_j goes from a_j to ta_j, and the second one is integral where τ_j goes from ta_j to infinity. The second additive term has a closed-form solution. In the third line, the probabilities are replaced by their definitions in Eq. 3, 4, 8 and 10.

Widowed respondents

We assume that widowed individuals expect to remain widowed until they pass away.[9] Thus, if spouse i is a widow(er) at age a_i, his or her survival probability of reaching age ta_i conditional on $(T_i^s \geq a_i, x_i, \eta_i)$ is:

$$
P(T_i^s \geq ta_i | T_i^s \geq a_i, x_i, \eta_i) = \exp\Big(- \Lambda_i^q(ta_i | a_i, x_i, \eta_i) \Big) \tag{13}
$$

As shown in Eq. 13, the subjective integrated hazard rate of widowed spouse i follows $\Lambda^q(\cdot)$. This implies that the perceived remaining lifetime distribution of spouse i has changed from the moment he/she enters the widowhood state.

3.3.2 Reporting model

Since reported survival probabilities do not directly reflect respondents' true survival expectations, we follow the approaches proposed by De Bresser and Van Soest (2013); Kleinjans and Van Soest (2014); Bissonnette et al. (2017) and De Bresser (2019) to control for respondents' rounding behavior. We incorporate a dependence between spouses' reporting behaviors in a couple as spouses can affect each other's reporting decisions.

In each survey wave, respondents answered at most two subjective survival expectation questions with different target ages. We use subscript $k \in \{1, 2\}$ to denote the question order, and subscript $w \in \{1, 2, .., 13\}$ to denote the survey wave. Let P_{iwk} be the observed response of spouse i of kth question in survey wave w. Let S_{iwk} be the true survival probability that spouse i reaches target age

[9] In our sample, the total number of widowed people is 5,699, and 386 of them (around 6.7%) married again between 1992 and 2016.

ta_{iwk}, that is kth question in survey wave w. If spouse i's partner is alive at given survey wave, then S_{iwk} follows Eq. 12. If spouse i is a widow(er), then S_{iwk} equals Eq. 13.

We do not directly observe S_{iwk}. The reported probabilities can be noisy:

$$P^*_{iwk} = S_{iwk} + \varepsilon^*_{iwk} \tag{14}$$

$$\text{where } \varepsilon^*_{iwk} \sim N(0, \sigma^{2*}_{iw}), \text{ and } \sigma^*_{iw} = \exp(z_{iw}\beta^\sigma_i). \tag{15}$$

Here ε^*_{iwk} is a measurement error that is independent of the true survival probability of spouse i. Since some groups of people with the same observed characteristics can make more errors than the other groups, we allow the variance of ε^*_{iwk} to vary with z_{iw}, a vector of observed covariates of spouse i (including a constant term).

Reported survival probabilities are subject to rounding. We model the rounding behavior of respondents such that they choose a certain degree of rounding (or rounding rule) which defines how precise the reported probabilities are. If respondents decide to report their survival probabilities precisely, they round their answer to the nearest integer that is a multiple of 1. On the other hand, if respondents prefer reporting a coarser answer, they choose a closest integer that is a multiple of some number larger than 1. For example, let the perceived survival probability of one respondent be 65.3%. If that respondent chooses to provide a precise report, he/she would report 65%, the nearest multiple of 1. Instead, if the respondent reports a less precise answer by rounding to a nearest multiple of 10, he/she will report 70%. As an extreme example, if the respondent chooses to round to a nearest integer that is a multiple of 50, he/she will report 50%.

Let R_{iwk} be the rounding decision of spouse i for question k in wave w. The realization of R_{iwk} is r. We assume that there are five possible rounding rules, $r \in \{1,2,3,4,5\}$, and each rule corresponds to a set of numbers Ω_r that have a common divisor δ_r:

- $R_{iwk} = 1$: Multiples of $\delta_1 = 1 \Longleftrightarrow \Omega_1 = \{0, 1, ..., 99, 10\}$
- $R_{iwk} = 2$: Multiples of $\delta_2 = 5 \Longleftrightarrow \Omega_2 = \{0, 5, ..., 95, 100\}$
- $R_{iwk} = 3$: Multiples of $\delta_3 = 10 \Longleftrightarrow \Omega_3 = \{0, 10, ..., 90, 100\}$
- $R_{iwk} = 4$: Multiples of $\delta_4 = 25 \Longleftrightarrow \Omega_4 = \{0, 25, 50, 75, 100\}$
- $R_{iwk} = 5$: Multiples of $\delta_5 = 50 \Longleftrightarrow \Omega_5 = \{0, 50, 100\}$

Since the rounding rules can be ordered in terms of precision, we model R_{iwk} with an ordered response equation, assuming spouse i chooses rounding rule r if:

$$R_{iwk} = r, \text{ if } m_{i,r-1} < z_{iw}\beta^R_i + \eta^R_i + \varepsilon^R_{iwk} \leq m_{i,r}. \tag{16}$$

Here $m_{i,0}, m_{i,1}, ..., m_{i,5}$, $i = 1, 2$ are the threshold parameters. We set $m_{i,0} = -\infty$, $m_{i,5} = \infty$, and $m_{i,1}$ is normalized to zero as z_{iw} includes a constant. The random component η_i^R is independent of z_{1w} and z_{2w}, and it captures the unobserved time-invariant effects that determine spouse i's rounding decision. The idiosyncratic rounding shock ε_{iwk}^R follows a standard normal distribution and is independent of observed covariates and all other errors.

Following the distribution assumption on ε_{iwk}^R, the probability of spouse i chooses rounding rule r conditional on (z_{iw}, η_i^R) is:

$$P(R_{iwk} = r | z_{iw}, \eta_i^R) = \Phi(m_{i,r} - z_{iw}\beta_i^R - \eta_i^R) - \Phi(m_{i,r-1} - z_{iw}\beta_i^R - \eta_i^R), \text{ for } r = 1, 2, 3, 4, 5. \quad (17)$$

Choosing a specific rounding rule implies an interval for the perturbed probability $P_{iwk}^* = [LB_{iwk}, UB_{iwk})$. Since $P_{iwk}^* \sim N(S_{iwk}, \sigma_{iw}^{2*})$, the probability that P_{iwk}^* is in the interval $[LB_{iwk}, UB_{iwk})$ is given by

$$\begin{aligned} &P(LB_{iw1} \leq P_{iw1}^* < UB_{iw1} | d_i, x_i, z_{iw}, \eta_i^s) \\ &\qquad = \Phi\Big(\frac{UB_{iw1} - S_{iw1}}{\sigma_{iw}^*} | d_i, x_i, z_{iw}, \eta_i^s\Big) - \Phi\Big(\frac{LB_{iw1} - S_{iw1}}{\sigma_{iw}^*} | d_i, x_i, z_{iw}, \eta_i^s\Big). \end{aligned} \quad (18)$$

Respondents answer at most two survival expectation questions with different target ages in each survey wave. Given that the first reported probability is P_{iw1}, if a respondent chooses rounding rule $R_{iw1} = r$, the upper and lower boundaries in Eq. 18 are equal to

$$P(LB_{iw1} \leq P_{iw1}^* < UB_{iw1} | d_i, x_i, z_{iw}, \eta_i^s) = \begin{cases} P(100\% - 0.5\delta_r \leq P_{iw1}^* | d_i, x_i, z_{iw}, \eta_i^s) & \text{if } P_{iw1} = 100\% \\ P(P_{iw1} - 0.5\delta_r \leq P_{iw1}^* \leq P_{iw1} + 0.5\delta_r | d_i, x_i, z_{iw}, \eta_i^s) & \text{if } 0\% < P_{iw1} < 100\% \\ P(P_{iw1}^* \leq 0.5\delta_r | d_i, x_i, z_{iw}, \eta_i^s) & \text{if } P_{iw1} = 0\% \end{cases} \quad (19)$$

Intuitively speaking, in the first question, if a perturbed probability 10% is rounded to a multiple of 5%, then the interval is $[7.5\%, 12.5\%)$. Since survival probabilities are bounded between 0% and 100%, we take account of censoring as shown in Eq. 19.

Following De Bresser (2019), whether a perturbed probability of the second question P_{iw2}^* is censored or not-censored depends on the degree of rounding rule and on the reported probability to the first question. Since the second question's target age is higher than that of in the first question, the perturbed probability of the second question should be lower than that of in the first question. Thus, for a given response of the second question P_{iw2} and at given rounding rule $R_{iw2} = r$, the lower and upper boundaries are defined as:

$$P(LB_{iw2} \leq P^*_{iw2} < UB_{iw2}|d_i, x_i, z_{iw}, \eta^s_i) = \\ \begin{cases} P(P_{iw1} - 0.5\delta_r \leq P^*_{iw2}|d_i, x_i, z_{iw}, \eta^s_i) & \text{if } P_{iw2} \geq P_{iw1} - 0.5\delta_r \\ P(P_{iw2} - 0.5\delta_r \leq P^*_{iw2} \leq P_{iw2} + 0.5\delta_r|d_i, x_i, z_{iw}, \eta^s_i) & \text{if } 0.5\delta_r \leq P_{iw2} < P_{iw1} - 0.5\delta_r \\ P(P^*_{iw2} \leq 0.5\delta_r|d_i, x_i, z_{iw}, \eta^s_i) & \text{if } P_{iw2} < 0.5\delta_r. \end{cases} \tag{20}$$

Since we do not directly observe which rounding rules respondents choose, a reported probability P_{iwk} may result different degrees of rounding. For example, observed response of 10% can be rounded to a multiple of 1 (interval: $[9.5\%, 10.5\%)$), a multiple of 5 (interval: $[7.5\%, 12.5\%)$), or a multiple of 10 (interval: $[5\%, 15\%)$). Thus, the probability of observing P_{iwk} conditional on $(d_i, x_i, z_{iw}, \eta^R_i, \eta^s_i)$ is:

$$P(P_{iwk}|d_i, x_{iw}, z_{iw}, \eta^R_i, \eta^s_i) = \\ \sum_{r=1}^{5} 1\{P_{iwk} \in \Omega_r\} P(R_{iwk} = r|z_{iw}, \eta^R_i) P(LB_{iwk} \leq P^*_{iwk} < UB_{iwk}|d_i, x_i, z_{iw}, \eta^s_i) \tag{21}$$

The conditional likelihood contribution of the subjective model given d_i, x_i, z_{iw} and η^s_i, η^R_i equals the probability of observing the reported probability P_{iwk} and is given by:

$$L^s_{iwk}(\gamma^s, \gamma^R|d_i, x_i, z_{iw}, \eta^R_i, \eta^s_i) = \begin{cases} P(P_{iwk}|d_i, x_i, z_{iw}, \eta^R_i, \eta^s_i), & \text{if a response is valid} \\ 1, & \text{otherwise} \end{cases} \tag{22}$$

As shown in Eq. 22, we set the subjective likelihood contribution to 1 if a response is invalid (answer "Don't know" or "Refused to answer") or if a respondent did not participate in the survey wave.

The conditional likelihood contribution of a couple, combining all waves and questions, is:

$$L^s(\gamma^s, \gamma^R|d, x, z, \eta^s, \eta^R) = \prod_{w=1}^{13} \prod_{i=1}^{2} \prod_{k=1}^{2} L^s_{iwk}(\gamma^s, \gamma^R|d_i, x_i, z_{iw}, \eta^R_i, \eta^s_i) \tag{23}$$

Here γ^R is a vector of unknown parameters of the reporting model, which includes $\beta^\sigma_i, \beta^R_i$, $m_{i,2}$, $m_{i,3}$, and $m_{i,4}$ for $i = \{1, 2\}$.

3.4 Likelihood contribution and estimation strategy

For the complete model, the conditional likelihood contribution of couple n is given by:

$$L_n(\gamma^o, \gamma^s, \gamma^R|d, x, z, \eta^o, \eta^s, \eta^R) = L^o_n(\gamma^o|d, x, \eta^o) \cdot L^s_n(\gamma^s, \gamma^R|d, x, z, \eta^s, \eta^R) \tag{24}$$

We allow the random components of the objective and subjective models to be correlated and assume that $[\eta_1^o, \eta_2^o, \eta_1^s, \eta_2^s, \eta_1^R, \eta_2^R]^T$ follows a six-dimensional normal distribution with zero means and arbitrary covariance matrix,Σ.

The economic interpretations of some elements of Σ are worth noting. Following the common lifestyle argument, we expect positive signs in $\sigma_{\eta_1^o,\eta_2^o}$ (e.g., the covariance between η_1^o and η_2^o) and $\sigma_{\eta_1^s,\eta_2^s}$. If respondents' subjective and objective survival expectations are positively correlated after controlling observed factors, as noted by Smith et al. (2001), then we expect $\sigma_{\eta_1^o,\eta_1^s} \geq 0$ and $\sigma_{\eta_2^o,\eta_2^s} \geq 0$. The unobserved factors of subjective survival expectations and those of reporting models can also be correlated. For example, suppose that cognitive capacity of people is partially captured through unobserved factors. If people with high cognitive ability perceive they are likely to survive to very old ages and tend to elicit precise subjective responses, we expect $\sigma_{R_1,\eta_1^s} \geq 0$ and $\sigma_{R_2,\eta_2^s} \geq 0$.[10] Moreover, in a couple's household, if two spouses tend to elicit their subjective responses with a similar degree of precision, we expect $\sigma_{R_1,R_2} > 0$.

Since the random components are unobserved, we need each couple's unconditional likelihood contribution. Moreover, the average of the unobserved frailty terms of the sample decreases as the average age of the population increases because people with higher values of η^o die sooner than people with low values of η^o. This implies that the distribution of unobserved frailty terms depends on the entry ages of couples (Lancaster, 1990). Accounting for this, the (unconditional) likelihood contribution of couple n is given by:

$$
\begin{aligned}
&L_n(\gamma^o, \gamma^s, \gamma^R, \Sigma | d, x, z) \\
&\quad = \int_{R^6} L_n(\gamma^o, \gamma^s, \gamma^R | T_1^o > a_1^*, T_2^o > a_2^*, d, x, z, \eta^o, \eta^s, \eta^R) g(\eta | T_1^o > a_1^*, T_2^o > a_2^*, d, x, z) d\eta && (25) \\
&\quad = \int_{R^6} \frac{L_n(\gamma^o, \gamma^s, \gamma^R | d, x, z, \eta^o, \eta^s, \eta^R)}{P(T_1^o > a_1^*, T_2^o > a_2^* | d, x, \eta^o)} \frac{g(\eta) P(T_1^o > a_1^*, T_2^o > a_2^* | d, x, \eta^o)}{P(T_1^o > a_1^*, T_2^o > a_2^* | d, x)} d\eta && (26) \\
&\quad = \frac{\int_{R^6} L_n(\gamma^o, \gamma^s, \gamma^R | d, x, z, \eta^o, \eta^s, \eta^R) g(\eta) d\eta^o}{P(T_1^o > a_1^*, T_2^o > a_2^* | d, x)} && (27) \\
&\quad = \frac{\int_{R^6} L_n(\gamma^o, \gamma^s, \gamma^R | d, x, z, \eta^o, \eta^s, \eta^R) g(\eta) d\eta}{\int_{R^2} S_0(a_1^*, a_2^* | d, x, \eta^o) g(\eta^o) d\eta^o} && (28)
\end{aligned}
$$

Here $g(\eta)$ is a six-variate probability density function of η, and $g(\eta^o)$ is the bivariate probability density function of η^o. In the first line of Eq. 28, both conditional likelihood function, and the density function of η are conditioned on spouses surviving until ages (a_1^*, a_2^*). The second line is of the equation is obtained using Bayes rule. Note that the conditional objective probability that spouses survive until ages (a_1^*, a_2^*), $P(T_1 > a_1^*, T_2 > a_2^* | d, x, z, \eta^o)$, does not depend on (z, η^s, η^R). Because of the common multipliers in denominator and numerator, we obtain the third line. Finally, the joint probability in the denominator is replaced by its definition in Eq. 8.

[10] As shown in Eq. 17, a higher value of η_i^R implies coarser rounding behavior. Eq. 3 and 4 show that a higher value of η_i^s implies higher hazard rates; thus, lower subjective survival probability. Thus, following our prior intuition, we expect positive signs in σ_{R_1,η_1^s} and σ_{R_2,η_2^s}.

Eq. 28 does not have a closed-form expression; thus, we rely on numerical methods to approximate the equation. First, to calculate Eq. 28, we need to approximate the integral in the numerator of Eq. 12. We follow the Mid-point approximation rule to numerically approximate the integral in the numerator of Eq. 12. For details, see Appendix A.

Second, we approximate the integral over unobserved random components using a simulation method, rewriting the unobserved factors as follows:

$$[\eta_{1n}^{o}, \eta_{2n}^{o}, \eta_{1n}^{s}, \eta_{2n}^{s}, \eta_{1n}^{R}, \eta_{2n}^{R}]^{T} = \Lambda u_n \tag{29}$$

where u_n is a column vector of six independent standard normal random variables. Here Λ is a Cholesky lower-triangular matrix with $\Sigma = \Lambda(\Lambda)^T$.

We can now approximate Eq. 28 with:

$$L_n^s(\gamma^o, \gamma^s, \gamma^R, \Lambda | d, x, z, P^R) \approx \frac{\frac{1}{M}\sum_{m=1}^{M} L_n(\gamma^o, \gamma^s, \gamma^R | \Lambda u_n^m, d, x, z, P^R)}{\frac{1}{M}\sum_{m=1}^{M} S_0(a_1^*, a_2^* | \Lambda^o u_n^{o,m}, d, x)} \tag{30}$$

Here M is the number of simulations drawn for each couple n. For couple n, u_n^m is mth simulation draw with six elements, and $u^{o,m}$ is the first two elements of u_n^m. In the denominator, $\Lambda^o u_n^{o,m}$ is the first two elements of vector Λu_n^m. We use Halton draws to generate random simulated draws to reduce the variance induced from the simulation procedure.[11] The model is estimated by the Maximum Simulated Likelihood method. If the number of simulation draws for each observation -M goes to infinity, the Maximum Simulated Likelihood estimator is asymptotically equivalent to the Maximum Likelihood estimator (Train, 2009).

[11] The Halton sequences are generated with prime bases 3, 5, 7, 11, 13, and 17. The elements of the first 17 six-pairs of the sequence are discarded since the early elements appear to be correlated (Train, 2009).

4 Results

4.1 Objective hazard rates

Columns (1) and (4) of Table 5 show the estimation results of the objective hazard rates of males and females, respectively. Using these estimates, we predict log objective hazard rates shown in subfigures (a) and (c) of Figure 3. The log hazard rates in the figure are predicted for an average individual, with observed regressors equal to the sample averages and unobserved frailties equal to zero. The solid and dashed lines depict the predicted log hazard rates if respondents' partners are alive, and if their partners died at a given age. The observed characteristics and frailties are the same when partners are alive or dead; the only difference is in the baseline hazards. The dotted lines are the 95% confidence intervals, calculated using a parametric bootstrap method.

According to subfigures (a) and (c) of Figure 3, the objective log hazard rates of people whose partners are alive are significantly lower than those of widowed people wth the same average characteristics. Since the confidence intervals of married and widowed people never overlap, we conclude that there exists a statistically significant and negative bereavement effect on the objective remaining lifetimes of surviving spouses. These results are in line with the results of the Van den Berg et al. (2011) and Spreeuw and Owadally (2013).

For the objective model, a positive parameter implies higher mortality risk and a lower survival probability.[12] According to the results in Columns (1) and (4) of Table 5, observed mortality risks covary in expected ways with demographic characteristics. For example, if household-size adjusted income increases by 1%, males' (females') mortality hazards decrease by 9.2% (10,9%). Those who ever smoked males (females) have 36% (39%) higher mortality rates than non-smoker males (females). If males' (females') number of functional limitations increases by one, their mortality hazards increase by 18.9% (24.4%). If males' (females') number of ever-recorded chronic conditions increases by one, their mortality hazards increase by 27.6% (27.3%).

Compared to less than high-school educated (fe)males, high-school educated males (females) have 9.1% (10.8%) lower mortality hazards. The magnitude of the effect is even larger for college-educated people. For example, compared to less than high-school educated (fe)males, college-educated males (females) have 22.8% (25.1%) lower objective hazards. Hispanic males (females) have 18.7% (21.2%) lower mortality hazards than non-Hispanic (fe)males. Black males have 11.1% higher mortality hazards than non-Black males; however, there is no significant difference between Black and non-Black females' mortality hazards.

[12] The parameters of the hazard rates can be interpreted as relative effects.Following the MPH specification:

$$\frac{\partial\lambda(t|x,\eta)}{\partial x} = \frac{\partial\lambda(t)\exp(x\beta+\eta)}{\partial x} = \beta\lambda(t|x,\eta) \to \beta = \frac{\partial\lambda(t|x,\eta)/\partial x}{\lambda(t|x,\eta)}$$

where $\lambda(t)$ is the baseline hazard.

4.2 Subjective hazard rates

Columns (2) and (5) of Table 5 show the estimation results of males' and females' subjective hazards rates, respectively. Sub-figures (b) and (d) of Figure 3 show the log subjective hazard rates for an average individual male and female.

The results suggest that both males' and females' perception regarding their mortality hazards increases significantly once they encounter a spousal bereavement. However, the magnitude of the bereavement effect differs between males and females. Males' perceived mortality risk increases irrespective of their age. For females the bereavement effect is small and insignificant at age below 55, but larger and significant at older ages.

To test whether people over- or underestimate the impact of certain observed factors on their actual mortality hazards, we report the estimated differences between the parameters of objective and subjective hazard rates in Columns (3) and (6) of Table 5. If the difference is positive, people overestimate the impact of the covariate on their actual mortality hazard rates. The differences between the objective and subjective duration dependence parameters of the baseline hazard rates are estimated to be significant and positive. These results imply that compared to the log objective baseline hazard rates, the log subjective baseline hazard rates start at a higher level at the initial time of risk (i.e., when the older spouse of a couple reaches age 50). As people grow older, the average subjective hazard rates increase slower than the average objective hazard rates. In other words, young people have more pessimistic views regarding their survival chances. These findings are in line with the results of Bissonnette et al. (2017) and Heimer et al. (2019).

We find that males correctly perceive the impact of income on their mortality hazards. In contrast, females significantly undervalue the impact of income. College or high-school educated people value the impact of their education on their actual mortality hazards correctly. Respondents significantly underestimate the impact of being Hispanic, whereas they significantly overestimate the impact of being Black. Respondents undervalue the impact of smoking, and the positive impact of functional limitations and chronic conditions on their actual mortality hazards.

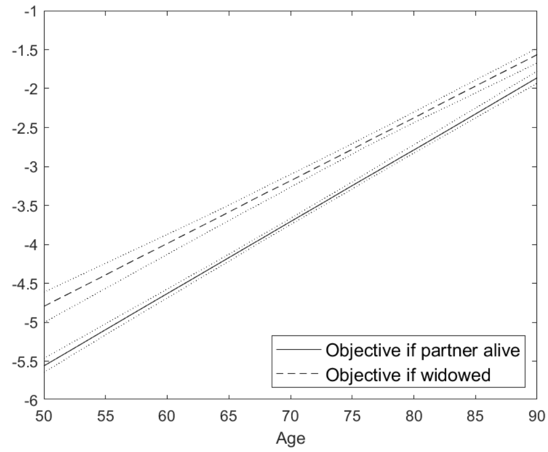
4.3 Reporting model

The estimation results of the reporting model are shown in Columns (7) to (10) of Table 5. Columns (7) and (9) show the estimation results of the log of the variance of measurement errors. The positive parameters in Columns (7) and (9) imply that people with certain characteristics report their perturbed probabilities with more measurement error than their counterparts.[13]

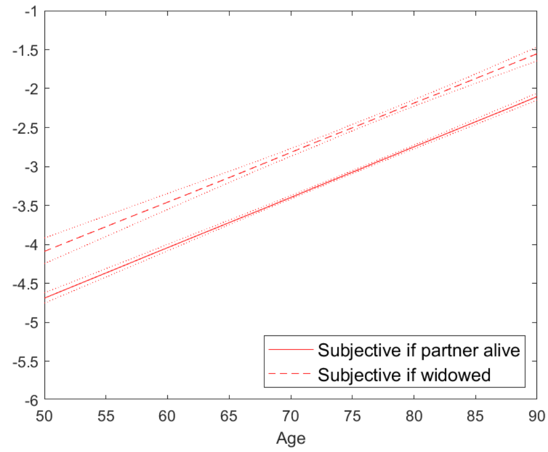
The results show that people who correctly recall a higher number of words tend to report with

[13] Recall that we define measurement error as the difference between people's perturbed probabilities, that are not polluted with a rounding bias, and their true perceived survival probabilities

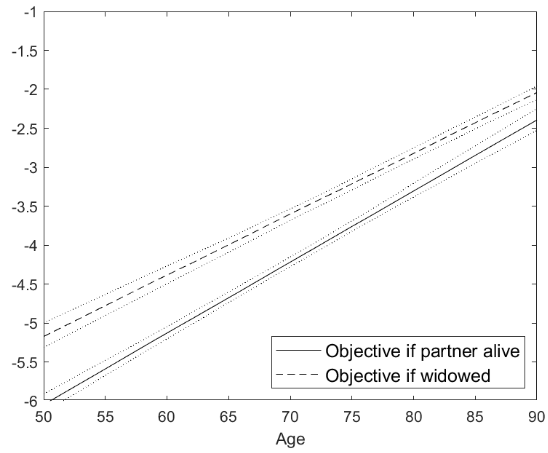
Figure 3: Predicted log hazard rates at an average individual level, conditional on partner's vital status



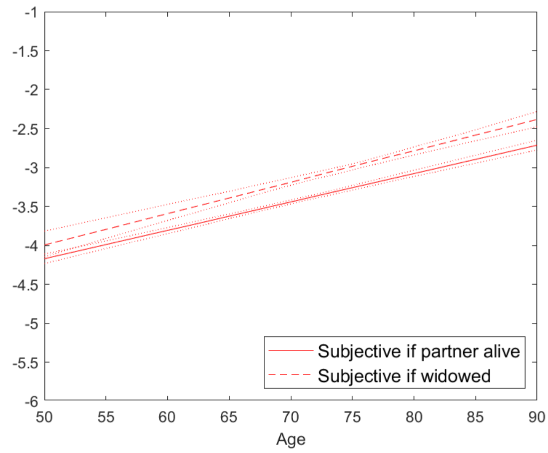
(a) Male, objective



(b) Male, subjective



(c) Female, objective



(d) Female, subjective

Note: The log hazard rates are linear, following the assumption that the baseline hazards follow Gompertz's specification. The top two and the bottom two sub-figures are the predicted log hazard rates of males and females, respectively. The left- and right-hand sides are the corresponding predictions of the objective and the subjective models, respectively. The dashed and the solid lines are the spouses' log hazards if their partners are alive and if their partners died at a given age. The dotted lines around the dashed and solid lines represent the 95% confidence intervals that reflect the estimation uncertainty.

less measurement error than those who recall fewer words. In contrast, if people's proportion of choosing multiples of 50 in other subjective expectation questions increases, their chance of reporting responses with more measurement error in survival expectation questions increases. High-school or college-educated people elicit subjective survival responses with less measurement error than less-than high-school educated people do. Blacks and Hispanics, and those born before 1945, are found to elicit their subjective responses with more measurement errors than their respective counterparts.

Columns (8) and (10) of Table 5 show the estimation results of the rounding equation. A positive parameter implies coarser rounding if the covariate increases. If respondents' proportions of choosing multiples of 50 in other expectation questions are high, those respondents' probability of choosing imprecise rounding rules is also high, in line with the findings of Manski and Molinari (2010). If the number of correctly recalled words increases, the probability to report more precisely increases.

Table 5: Estimation results, part 1

Column	(1)	(2)	(3)	(4)	(5)	(6)
		MALE			**FEMALE**	
	Objective	**Subjective**	**Difference (obj-subj)**	**Objective**	**Subjective**	**Difference (obj-subj)**
constant-married (β^p)	-9.757	-7.224	-2.533	-9.739	-5.488	-4.25
	(0.215)	(0.124)	(0.248)	(0.277)	(0.126)	(0.308)
constant-widowed (β^q)	-8.389	-6.52	-1.87	-8.218	-5.515	-2.704
	(0.330)	(0.356)	(0.487)	(0.299)	(0.263)	(0.402)
alpha-married (α^p)	0.092	0.065	0.027	0.091	0.036	0.055
	(0.002)	(0.001)	(0.003)	(0.003)	(0.002)	(0.004)
alpha-widowed (α^q)	0.08	0.063	0.017	0.078	0.04	0.038
	(0.004)	(0.005)	(0.006)	(0.003)	(0.003)	(0.005)
Log of income	-0.092	-0.073	-0.018	-0.109	-0.057	-0.052
	(0.012)	(0.006)	(0.014)	(0.018)	(0.005)	(0.019)
Ever smoked	0.358	0.091	0.267	0.394	0.04	0.354
	(0.036)	(0.011)	(0.038)	(0.034)	(0.009)	(0.036)
Functional limitations	0.189	0.066	0.123	0.244	0.073	0.171
	(0.016)	(0.009)	(0.018)	(0.018)	(0.009)	(0.021)
Chronic conditions	0.276	0.151	0.125	0.273	0.11	0.163
	(0.012)	(0.005)	(0.013)	(0.013)	(0.005)	(0.014)
High-school	-0.091	-0.028	-0.063	-0.108	-0.094	-0.014
	(0.038)	(0.018)	(0.042)	(0.041)	(0.019)	(0.046)
College	-0.228	-0.164	-0.064	-0.251	-0.268	0.017
	(0.04)	(0.018)	(0.044)	(0.046)	(0.019)	(0.051)
Hispanic	-0.187	-0.027	-0.16	-0.212	0.211	-0.423
	(0.06)	(0.023)	(0.065)	(0.069)	(0.018)	(0.071)
Black	0.111	-0.356	0.467	0.031	-0.164	0.195
	(0.044)	(0.017)	(0.047)	(0.053)	(0.016)	(0.056)
Cohort before 1930	0.015	-0.193	0.208	-0.106	0.213	-0.319
	(0.085)	(0.026)	(0.089)	(0.094)	(0.036)	(0.102)
Cohort 1931-1945	0.106	-0.077	0.183	-0.179	0.048	-0.227
	(0.073)	(0.011)	(0.074)	(0.078)	(0.011)	(0.079)

Note: The standard errors, in the brackets, are calculated using the outer-product gradient estimates. We use 30 Halton draws for each observation. The estimated negative log-likelihood is 640,730. The total number of parameters is 119. The sample consist of 11,043 couples' observations. The total number of observations (spouses × waves × valid responses) is 126,250.

Furthermore, high-school or college-educated people appear to answer more precisely than less educated people. Hispanics, Black, and those born before 1945 are more likely to give imprecise answers than their counterparts.

4.4 Covariance matrix of unobserved factors

Columns (11) to (16) of Table 5 show the estimates of the Cholesky lower triangular matrix. In the bottom half of Columns (11) to (16), we show the covariance and correlation coefficients of unobserved factors. The estimated covariance and correlation coefficients are in the lower and upper triangular parts of the matrix, respectively. The estimated variances of the unobserved factors are on the diagonal. The standard errors of the parameters in Part B of Columns (11) to (16) are calculated using a parametric bootstrap method.

We find significant variation in the unobserved factors of both objective and subjective hazard rates, both for males and females ($\widehat{Var}(\eta_i^o)$ and $\widehat{Var}(\eta_i^s)$ for $i = 1, 2$). The estimated variances of

Table 5 continued, part 2.

Column	(7)	(8)	(9)	(10)
	MALE		**FEMALE**	
	log of variance of measurement error	**rounding model**	**log of variance of measurement error**	**rounding model**
constant	-1.24	2.012	-0.977	2.201
	(0.006)	(0.024)	(0.006)	(0.023)
Prop. multiples of 50	0.104	0.717	0.109	0.696
	(0.004)	(0.014)	(0.004)	(0.013)
Immediate word recall	-0.139	-0.223	-0.095	-0.196
	(0.008)	(0.028)	(0.008)	(0.025)
High-school	-0.094	-0.069	-0.141	-0.117
	(0.003)	(0.015)	(0.003)	0.015
College	-0.164	-0.280	-0.294	-0.314
	(0.003)	(0.015)	(0.003)	(0.015)
Hispanic	0.186	0.066	0.131	0.037
	(0.004)	(0.018)	(0.004)	(0.017)
Black	0.1	0.211	0.055	0.206
	(0.003)	(0.015)	(0.003)	(0.015)
Cohort before 1930	0.341	0.078	0.406	0.145
	(0.004)	(0.017)	(0.006)	(0.019)
Cohort 1931-1945	0.065	0.098	0.117	0.078
	(0.002)	(0.011)	(0.002)	(0.011)
m_1		1.180		1.217
		(0.012)		(0.011)
$m_2 - m_1$		1.346		1.405
		(0.008)		(0.007)
$m_3 - m_2$		0.389		0.431
		(0.005)		(0.004)

the unobserved frailty terms are much larger for subjective hazards than for objective hazards. This suggests a higher variation among couples' perceived mortality compared to their actual mortality rates. The correlation between the objective frailty terms of males and females is positive but insignificant ($\widehat{Corr}(\eta_1^o, \eta_2^o) > 0$). The correlation between the subjective frailty terms of males and females is positive and significant ($\widehat{Corr}(\eta_1^s, \eta_2^s) > 0$). Hence, once we control observed regressors, spouses' subjective remaining lifetimes are positively correlated.

The correlations between objective and subjective frailty terms are positive and significant for males ($\widehat{Corr}(\eta_1^o, \eta_1^s) > 0$) but negative and insignificant for females ($\widehat{Corr}(\eta_2^o, \eta_2^s) > 0$). These results are in line with the results of Smith et al. (2001), who conclude that even after controlling for the main observed mortality predictors, there is a correlation between the subjective survival responses and the objective survival rates. The reason could be that males are aware of some factors that affect their mortality rates that are unobserved to the econometrician and take account of those factors when they formulate their survival expectations.

There is significant variation in unobserved factors of the rounding decisions, both for males and females ($\widehat{Var}(\eta_i^R)$ for $i = 1, 2$). The positive correlation between the unobserved factors of males' and females' rounding decisions implies that spouses in a couple tend to choose a similar degree of rounding, controlling for the observed characteristics ($\widehat{Corr}(\eta_1^R, \eta_2^R) > 0$). The intuition can be that spouses in a couple have intrinsically similar rounding behaviors and/or spouses interfere with each other's reporting decisions during an interview; the latter is plausible if couples are interviewed together.

Table 5 continued. part 3

Column	(11)	(12)	(13)	(14)	(15)	(16)
Panel A. Cholesky lower triangular						
	η_1^o	η_2^o	η_1^s	η_2^s	η_1^R	η_2^R
η_1^o	0.172					
	(0.023)					
η_2^o	0.025	0.077				
	(0.032)	(0.026)				
η_1^s	0.355	0.209	0.514			
	(0.006)	(0.005)	(0.006)			
η_2^s	-0.304	0.002	0.358	0.001		
	(0.006)	(0.005)	(0.006)	(0.005)		
η_1^R	-0.09	0.15	-0.09	0.01	-0.398	
	(0.006)	(0.006)	(0.005)	(0.007)	(0.006)	
η_2^R	-0.179	0.438	-0.047	0.041	0.155	0.013
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
Panel B. Estimated covariance - correlation matrix						
	η_1^o	η_2^o	η_1^s	η_2^s	η_1^R	η_2^R
η_1^o	0.029	0.302	0.539	-0.647	-0.203	-0.357
	(0.008)	(0.361)	(0.008)	(0.010)	(0.014)	(0.011)
η_2^o	0.004	0.007	0.465	-0.192	0.26	0.723
	(0.006)	(0.005)	(0.182)	(0.234)	(0.102)	(0.21)
η_1^s	0.061	0.024	0.434	0.247	-0.16	0.011
	(0.008)	(0.013)	(0.008)	(0.014)	(0.011)	(0.012)
η_2^s	-0.052	-0.007	0.076	0.22	-0.021	0.162
	(0.007)	(0.01)	(0.005)	(0.005)	(0.015)	(0.013)
η_1^R	-0.016	0.009	-0.047	-0.004	0.197	0.112
	(0.002)	(0.005)	(0.003)	(0.003)	(0.005)	(0.017)
η_2^R	-0.031	0.029	0.004	0.038	0.025	0.252
	(0.004)	(0.012)	(0.004)	(0.003)	(0.004)	(0.006)

Notes: In Part B, the diagonal elements are the estimates of the variance of the corresponding elements. The elements in the lower-triangular diagonals are the estimated covariances of the corresponding row and column elements. The elements in the upper-triangular diagonals are the estimated correlation coefficients of the corresponding row and column elements. In Part B, the standard errors are calculated using a parametric bootstrap method.

4.5 Comparing predicted objective and subjective survival probability curves

In applied work, researchers often need a proxy for people's perceived survival probabilities (French, 2005; De Nardi, French, & Jones, 2009; Casanova, 2010; Michaud et al., 2019). Using the estimation results of the objective and subjective hazard rates, we show predicted survival probabilities in Figure

4, presenting median survival curves but also illustrating heterogeneity.

To predict these survival curves, we first predict unobserved frailties of each couple, see Appendix B for details. We predict these survival curves for individuals who survived to age 70. The black dashed lines and solid red lines are the median predictions of the objective and subjective survival curves.

We predict the curves for the cases the partner is alive at respondent's age 70 and the partner died at respondent's age 70. The figure shows that if an average male's partner is alive at respondent's age 70, his median objective and subjective survival curves are very close to each other. However, if the partner died at respondent's age 70, his subjective survival curves are depicted above the objective survival curve. Thus, an average male tends to undervalue the impact of encountering a spousal bereavement on his actual mortality. In contrast, an average female at age 70 predicts her subjective survival curves very close to her objective survival curves, both if her partner is alive and if her partner is dead. Thus, it seems that at the median, the average female with age 70 correctly perceives the impact of encountering a spousal bereavement on mortality.

The blue and red areas reflect the 10th and 90th percentiles of the objective and subjective curves that reflect the dispersion in predictions. In the top two sub-figures, they reflect the dispersion induced through unobserved frailties. They show that people's perceived survival curves are more dispersed than their objective survival curves in terms of the unobserved factors, as expected from the estimated variances of the unobserved frailties in the subjective and objective models.

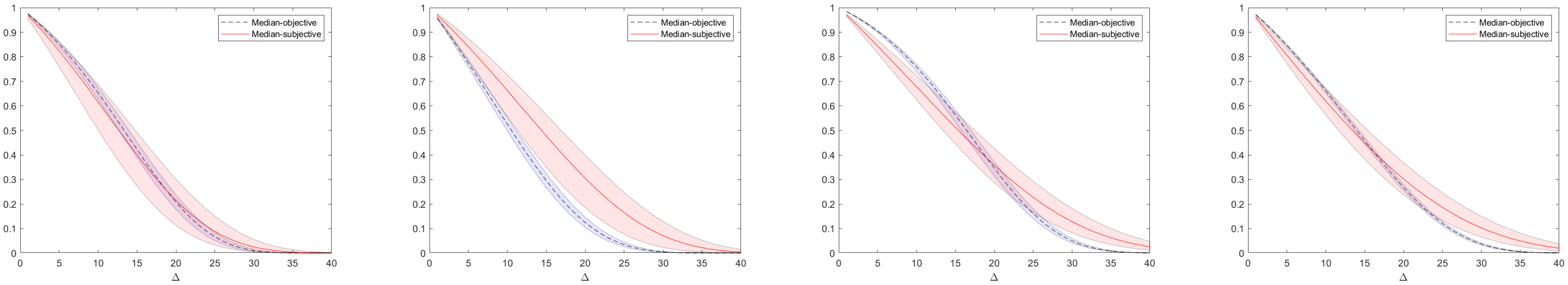
In the bottom part of Figure 4, the blue and the red areas reflect the 10th and 90th percentiles accounting for both unobserved frailties and observed characteristics. Here the objective survival curves have much more dispersion than the subjective survival curves. This is explained by the fact that the estimated impact of most observed characteristics is much larger for objective than for subjective hazards. For example, the ever-smoking males' objective hazards rates are, on average, 35.8% higher than the objective hazards rates of never-smoking males. In contrast, the ever-smoking males' subjective hazards rates are, on average, only 9.1% higher than the subjective hazards rates of never-smoking males. Although the opposite is true for Blacks and those born before 1945, the overall impact of observed characteristics makes the predictions of objective survival curves much more dispersed than those of subjective survival curves.

4.6 Dependent remaining lifetimes

The remaining lifetimes of spouses in a couple at a given age can be positively related through two different channels: (1) bereavement effects and (2) correlation between observed and unobserved factors that affect spouses' mortality hazards. Gourieroux and Lu (2015) propose an index to measure the remaining lifetime dependence of spouses in a couple when the bereavement effect is present. They compare the hazard rates of spouses whose partners are alive with those whose partners died at a cer-

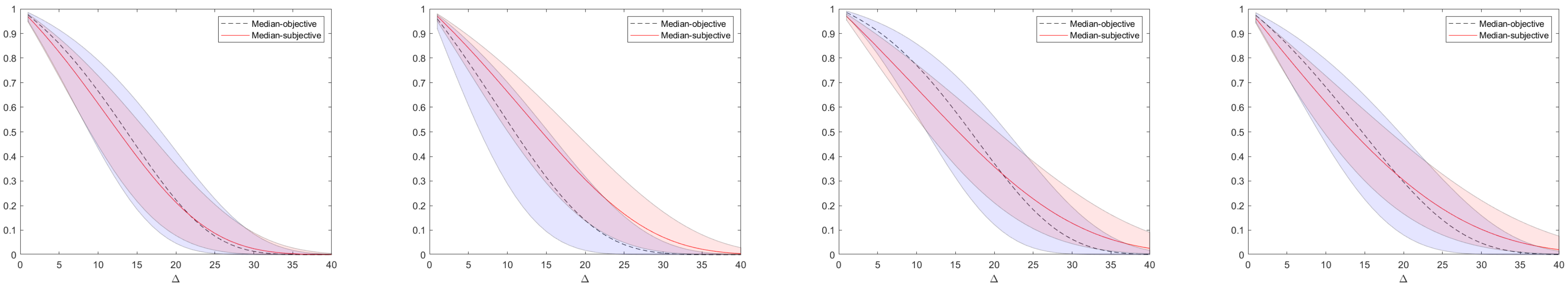
Figure 4: Predicted survival probabilities of 70-year-old people to survive for a certain number of years (Δ), conditional on their partners' vital status

A. Dispersion in unobserved frailties



Males whose partners alive | Males whose partners dead | Females whose partners alive | Females whose partners dead

B. Dispersion in unobserved frailties and observed factors



Males whose partners alive | Males whose partners dead | Females whose partners alive | Females whose partners dead

Note: The left- and right-hands sides of the figure illustrate the predicted values of $P(T_i \geq 70 + \Delta | T_i \geq 70, T_{3-i} \geq 70 + d_{3-i} - d_i, x, \hat{\eta})$ if partners are alive, and $P(T_i \geq 70 + \Delta | T_i \geq 70, T_{3-i} = 70 + d_{3-i} - d_i, x, \hat{\eta})$ if partners died. Here $d_{3-i} - d_i$ is the age difference between spouse $3 - i$ and i, and $\hat{\eta}$ is the predicted values of unobserved frailties. $i = 1$ for males and $i = 2$ for females. The black and red areas reflect the prediction heterogeneity induced via unobserved frailties in the top sub-figures, and via unobserved frailties and observed characteristics in the bottom sub-figures. The survival curves are predicted at estimated values of parameters; thus, the blue and red areas do not reflect the estimation uncertainty.

tain age, conditional on both spouses surviving until that age. In other words, the index measures the immediate jump in the mortality hazard rates of surviving spouses at a time of spousal bereavement. This dependence index can be decomposed into a part explained by the bereavement effect, and a part that is explained by correlations between (un)observed factors. Details on the dependence index are presented in Appendix C.

The dependence index is calculated for objective and subjective remaining lifetimes, and the results are illustrated in Figure 5. The orange and blue areas show the parts of the dependence index explained by bereavement effects and the correlation between (un)observed factors, respectively. Since the bereavement effect can have asymmetric effects on males' and females' hazard rates, the dependence index is calculated separately for males and females. Note that the blue areas are the same for males and females as this reflects the part explained by the correlated (un)observed factors.

The figure shows that spouses' objective remaining lifetimes are more related than spouses' subjective remaining lifetimes. For example, conditional on both spouses have reached age 50; if a male (female) spouse gets widowed at age 50, his (her) objective log hazard rates increases by 1 (1.2). In contrast, conditional on both spouses have reached age 50; if a male (female) spouse gets widowed at age 50, his (her) subjective hazard rate increases by 0.8 (0.35). Moreover, depending on age and gender, the bereavement effect explains 87-99% of the immediate jump in the objective hazard rates. The rest is explained by the correlations between (un)observed factors. The bereavement effect explains 85-92% of the immediate jump in males' subjective hazard rates and 60-88% of those of females.

The dependence index of Gourieroux and Lu (2015) has the limitation that it focuses on the immediate changes in the hazard rates at the time of spousal bereavement. An alternative is to compare expected remaining life-years of people whose partners are alive with those whose partners died at a certain age. These can be computed using the complete predicted survival curves.

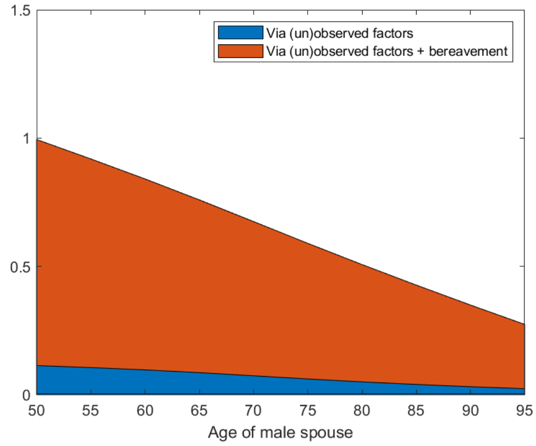
Figure 6 shows respondents' expected remaining life-years conditional on surviving to a certain age, separately for males and females. The black and red lines are the objective and subjective expectations, respectively. The solid lines are if spouses' partners are alive, and the dashed lines if their partners died.

To decompose into bereavement effect and correlation of (un)observed factors, we run several counterfactual scenarios in which some dependence channels are shut down. In the left-hand sides of Figure 6, we depict the expected remaining life-years of average individuals if spouses' remaining lifetimes only depend on (un)observed factors. We set the baseline hazards of widowed people equal to those of married people. In that way, we shut down the bereavement effect channel.[14] The results show that if partners are alive at a given age, spouses' expected remaining life years are slightly larger than the expected remaining lifetimes of spouses whose partners died.

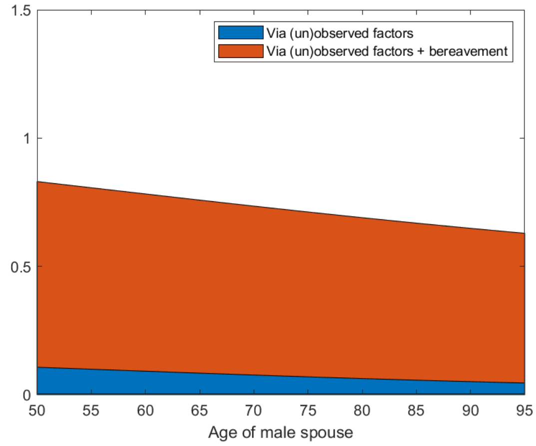
[14] We have analyzed the scenarios in which either the channel via observed characteristics or the channel via unobserved frailties is present, and the other one is shut down. However, the impact of unobserved characteristics on the expected remaining life-years was marginal, both for males and females and for objective and subjective models. Thus, we do not show the analysis results in which either observed characteristics or unobserved frailties are shut down.

Figure 5: Dependence index

A. Males

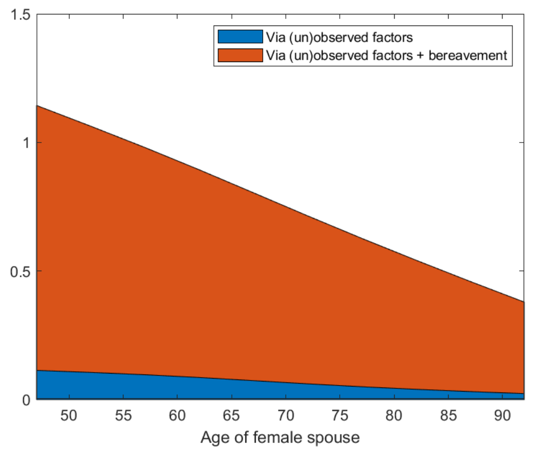


Objective lifetimes

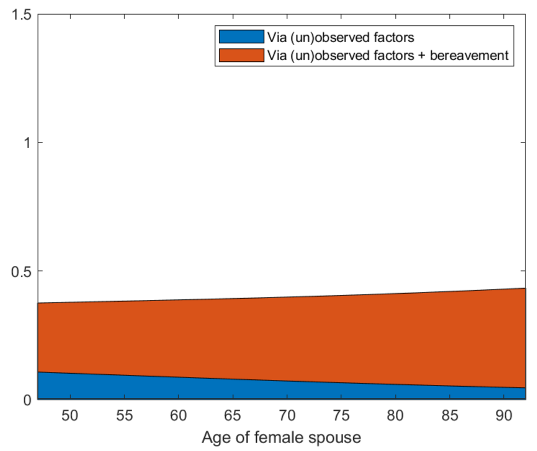


Subjective lifetimes

B. Females



Objective lifetimes



Subjective lifetimes

Note: The dependence index compares the immediate changes in log hazard rates of surviving spouses at the time of spousal bereavement. The index is calculated as $\ln \lambda(T_i|T_i \geq t_i, T_{3-i} \geq t_{3-i}) - \ln \lambda(T_i|T_i \geq t_i, T_{3-i} = t_{3-i})$, where $i = 1$ for males and $i = 2$ for females. Note that these log hazard rates are unconditional on (un)observed factors. Appendix C shows that the dependence index can be decomposed into additive terms. As shown in the figure, the orange areas show the part of the dependence index explained by the bereavement effect. The blue areas show the part of the dependence index explained by the correlation between (un)observed factors of spouses.

Next, we incorporate the bereavement effects. The corresponding predictions are illustrated in the middle of Figure 6. Once we introduce bereavement effects, there is a clear difference between the expected remaining lifetimes of people whose partners are alive and people whose partners are dead. Comparing the figures on the left with the figures in the middle, we can see that after incorporating a bereavement effect, the predictions of married and widowed people's expected remaining lifetimes decrease. Since we set the baseline hazards of widowed people equal to those of married people, it is evident that the figures on the left predict larger expected remaining life-years than those in the middle for widow(er)s. Married people's predicted life expectancy also falls once we incorporate the bereavement effect channel because the presence of the bereavement effect enables the possibility that married people are at risk of losing their partners in any future period.

On the right-hand side of Figure 6, we show the differences between the expected remaining lifetimes of those whose partners are alive and those whose partners are dead. The positive differences imply that spouses' remaining lifetimes are positively dependent, because people whose spouses are alive tend to live longer than those whose partner is dead.[15] For example, average males (females) who survived until age 50, would live around 5 (4.4) more years if their partners are alive. Out of this 5 (4.4) years of gain, 0.9 years, or 18% (20.5%), is explained by the correlated (un)observed factors, and 4.1 years (3.5 years), or 82% (79.5%), is explained by the fact that married people have not encountered any spousal bereavement.

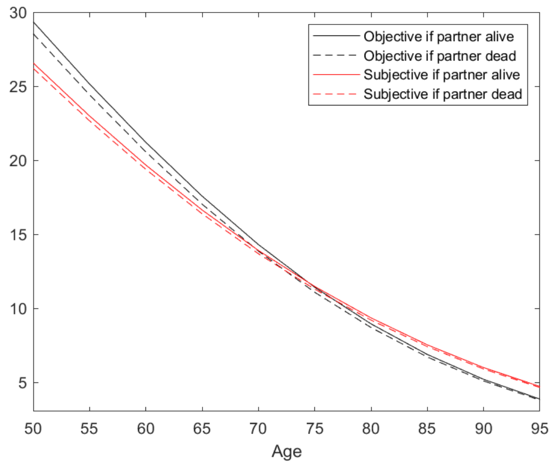
Spouses' perceived remaining lifetimes are also positively dependent. For example, males (resp. females) who reached age 50 perceive that they would live around 4.7 (resp. 2.2) more years if their partners are alive. Around 4.4 (resp. 1.9) years, or 94% (resp. 86%), of this gain from having their partners alive, is explained by the fact that they have not encountered any perceived bereavement effect, and the rest is explained by the correlated (un)observed factors of both spouses.

The dependence between the remaining lifetime dependence falls as people age. As shown in the figures, if males (resp. females) survive to age 95 with their partners, they would live 0.9 (resp. 0.6) more years compared to males (resp. females) who lost their partners. Moreover, at older ages, people overvalue the impact of having their partners alive on their actual life expectancy. For example, if males (females) survive to age 95 with their partners, they perceive they would live 1.4 more years (1 more year) than males (females) whose partners are dead. For both males and females with age 95 and for objective and subjective life expectancies, most of the gain from having the partner alive is explained by the fact that married people have not encountered any bereavement effect.

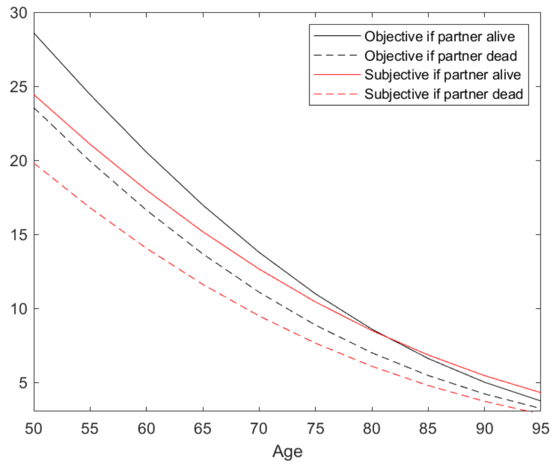
[15] The opposite is true if the difference is predicted to be negative. If the remaining lifetimes of spouses in a couple were independent, the difference between married and widowed people's life expectancies would be zero.

Figure 6: Expected remaining life years conditional on surviving to a certain age

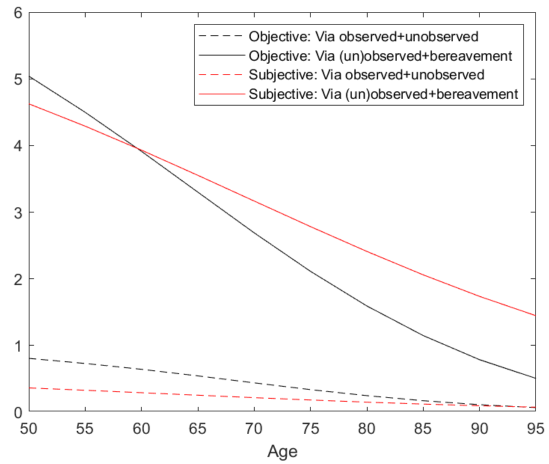
A. Males



dependent via observed + unobserved factors

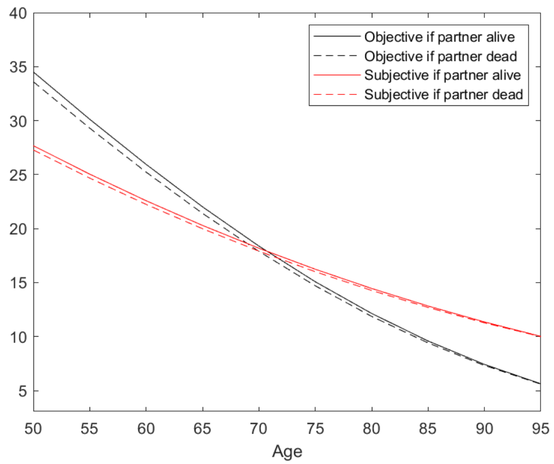


dependent via (un)observed+bereavement

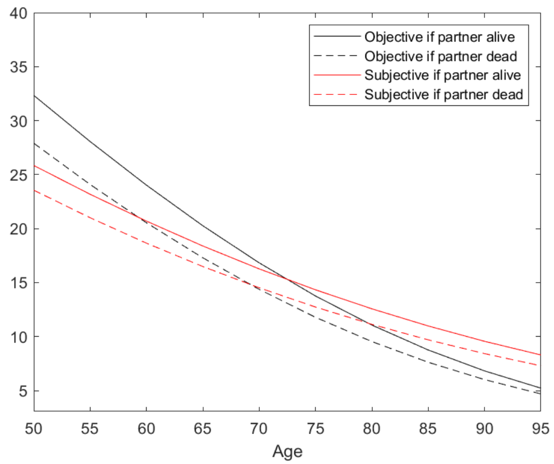


Difference between if partner is alive & if dead

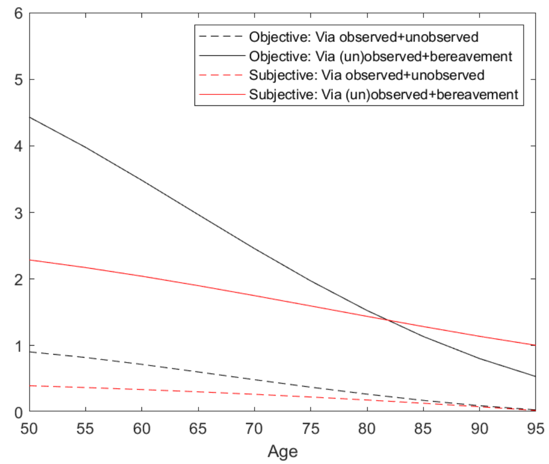
B. Females



dependent via observed + unobserved factors



dependent via (un)observed+bereavement



Difference between if partner is alive & if dead

5 Conclusion

This paper estimates the models for married couples' joint actual and perceived survival probabilities using longitudinal data. We estimate the objective model using couples' actual mortality data and estimate the subjective model using couples' reported subjective expectations to survive to certain target ages. Our model captures the dependence between remaining life times of spouses in a couple by allowing for a correlation between the (observed and unobserved) factors that explain mortality hazard rates, as well as a structural change in the surviving spouse's baseline hazard rates when the first spouse dies.

We find that the remaining actual and perceived lifetimes of spouses are positively correlated. Depending on age, males' (resp.females) are expected to live for around 0.6-5 (0.5-4.4) more years if their partners are alive compared to males (females) whose partners are dead. People also perceive that they would live longer if their partners are alive. For example, depending on age, males (females) perceive to live for 1.4-4.8 (1.6-2.3) more years if their partners are alive compared to males (females) whose partners are dead. The bereavement effect explains 85-98% of the actual life-expectancy difference between widowed and non-widowed people. In contrast, the bereavement effect explains 60-90% of the perceived life-expectancy difference between widowed and non-widowed people. Correlated (un)observables explain the rest.

We find that actual and perceived survival probabilities of married couples are significantly different. On average, couples whose spouses are younger than 75 and widowed individuals underestimate their expected remaining life years. Misinterpretation of their remaining expected life-years could let people allocate their assets not optimally over their life courses. For example, pessimistic young couples consume more than the optimal level, thus spending down their wealth too fast. Hence, they are at risk of being trapped in poverty and leaving a too small bequest to a surviving spouse.

People update their survival expectations after a spousal bereavement, but they underestimate the magnitude of the negative impact of the bereavement effect. This suggests that informing spouses about the bereavement effect when both spouses are alive would help them understand the bereavement effect's severe consequences. Better informed spouses are more likely to prepare against risks that a potential widow(er) would encounter after spousal bereavement. Moreover, the bereavement effects' substantial negative impact asks for a policy to support and monitor widowed people right after a spousal bereavement to improve their remaining lifetimes.

Our predictions of married couples' survival probabilities can be used to predict couples' life-cycle choices. Since we propose predictions based on couples' elicited survival expectations, it is more credible to use our model to compute the aspects of the joint posterior distribution of survival probability that capture perceived dependence. Researchers can then relate the predicted perceived survival probabilities to some aspects of married couples' life-cycle decisions, especially their decision before and after one of the spouses enter widowhood.

Throughout the paper, we assumed that married people are aware of their mortality distribution if they get widowed. Furthermore, we assumed that both spouses in a couple have the same expectations regarding each spouse's survival probability. In future research, these assumptions could be relaxed by collecting more information on respondents' survival expectations conditional on remaining married or becoming a widow(er) in all future periods. Also, further information on how spouses perceives their partners' survival expectations will help to identify the subjective joint survival distribution from each spouse's perspective.

References

Aquilino, W. S. (1993). Effects of spouse presence during the interview on survey responses concerning marriage. *Public Opinion Quarterly*, *57*(3), 358–376. pages 3

Aquilino, W. S., Wright, D. L., & Supple, A. J. (2000). Response effects due to bystander presence in CASI and paper-and-pencil surveys of drug use and alcohol use. *Substance Use & Misuse*, *35*(6-8), 845–867. pages 3

Bissonnette, L., Hurd, M. D., & Michaud, P.-C. (2017). Individual survival curves comparing subjective and observed mortality risks. *Health Economics*, *26*(12), e285–e303. pages 2, 3, 7, 8, 14, 21

Browning, M. (2000). The saving behaviour of a two-person household. *Scandinavian Journal of Economics*, *102*(2), 235–251. pages 2

Bruine de Bruin, W., Fischbeck, P. S., Stiber, N. A., & Fischhoff, B. (2002). What number is "fifty-fifty"?: Redistributing excessive 50% responses in elicited probabilities. *Risk Analysis: An International Journal*, *22*(4), 713–723. pages 7

Burgess, E. W., & Wallin, P. (1943). Homogamy in social characteristics. *American Journal of Sociology*, *49*(2), 109–124. pages 2

Carriere, J. F. (2000). Bivariate survival models for coupled lives. *Scandinavian Actuarial Journal*, *2000*(1), 17–32. pages 11

Casanova, M. (2010). Happy together: A structural model of couples' joint retirement choices. *Unpublished Manuscript, Department of Economics, UCLA*. pages 2, 25

De Bresser, J. (2019). Measuring subjective survival expectations–Do response scales matter? *Journal of Economic Behavior & Organization*, *165*, 136–156. pages 3, 6, 7, 14, 16

De Bresser, J., & Van Soest, A. (2013). Survey response in probabilistic questions and its impact on inference. *Journal of Economic Behavior & Organization*, *96*, 65–84. pages 3, 7, 14

De Nardi, M., French, E., & Jones, J. B. (2009). Life expectancy and old age savings. *American Economic Review*, *99*(2), 110–15. pages 25

Frees, E. W., Carriere, J., & Valdez, E. (1996). Annuity valuation with dependent mortality. *Journal of Risk and Insurance*, *63(2)*, 229–261. pages 11

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies*, *72*(2), 395–427. pages 25

Freund, J. E. (1961). A bivariate extension of the exponential distribution. *Journal of the American Statistical Association*, *56*(296), 971–977. pages 3

Gan, L., Gong, G., Hurd, M., & McFadden, D. (2015). Subjective mortality risk and bequests. *Journal of Econometrics*, *188*(2), 514–525. pages 2

Gourieroux, C., & Lu, Y. (2015). Love and death: A Freund model with frailty. *Insurance: Mathematics and Economics*, *63*, 191–203. pages 3, 26, 28, 39

Heimer, R. Z., Myrseth, K. O. R., & Schoenle, R. S. (2019). YOLO: Mortality beliefs and household finance puzzles. *The Journal of Finance*, *74*(6), 2957–2996. pages 2, 3, 21

Hollingshead, A. B. (1950). Cultural factors in the selection of marriage mates. *American Sociological Review*, *15*(5), 619–627. pages 2

Hurd, M. D. (1989). Mortality risk and bequests. *Econometrica*, *57(4)*, 779–813. pages 2

Hurd, M. D., McFadden, D. L., & Gan, L. (1998). Subjective survival curves and life cycle behavior. In *Inquiries in the economics of aging* (pp. 259–309). University of Chicago Press. pages 8

Iannario, M., & Piccolo, D. (2010). Statistical modelling of subjective survival probabilities. *Genus*, *66*(2). pages 2

Kapteyn, A., & Kooreman, P. (1992). Household labor supply: What kind of data can tell us how many decision makers there are. *European Economic Review*, *36*(2-3), 365–371. pages 13

Khwaja, A., Sloan, F., & Chung, S. (2007). The relationship between individual expectations and behaviors: Mortality expectations and smoking decisions. *Journal of Risk and Uncertainty*, *35*(2), 179–201. pages 2

Kleinjans, K. J., & Van Soest, A. (2014). Rounding, focal point answers and nonresponse to subjective probability questions. *Journal of Applied Econometrics*, *29*(4), 567–585. pages 3, 7, 14

Lancaster, T. (1990). *The Econometric Analysis of Transition Data* (Vol. 17). Cambridge University Press. pages 18

Lillard, L. A., & Willis, R. J. (2001). Cognition and wealth: The importance of probabilistic thinking. *Michigan Retirement Research Center Research Paper No. WP*, *7*. pages 8

Luciano, E., Spreeuw, J., & Vigna, E. (2008). Modelling stochastic mortality for dependent lives. *Insurance: Mathematics and Economics*, *43*(2), 234–244. pages 11

Ludwig, A., & Zimper, A. (2013). A parsimonious model of subjective life expectancy. *Theory and Decision*, *75*(4), 519–541. pages 2

Manski, C. F. (2004). Measuring expectations. *Econometrica*, *72*(5), 1329–1376. pages 3

Manski, C. F., & Molinari, F. (2010). Rounding probabilistic expectations in surveys. *Journal of Business & Economic Statistics*, *28*(2), 219–231. pages 3, 8, 22

Michaud, P.-C., Van Soest, A., & Bissonnette, L. (2019). Understanding joint retirement. *Journal of Economic Behavior & Organization*. pages 2, 13, 25

Salm, M. (2010). Subjective mortality expectations and consumption and saving behaviours among the elderly. *Canadian Journal of Economics/Revue Canadienne d'Économique*, *43*(3), 1040–1057. pages 2

Sanders, L., & Melenberg, B. (2016). Estimating the joint survival probabilities of married individuals. *Insurance: Mathematics and Economics*, *67*, 88–106. pages 3

Smith, V. K., Taylor, D. H., & Sloan, F. A. (2001). Longevity expectations and death: Can people predict their own demise? *American Economic Review*, *91*(4), 1126–1134. pages 2, 3, 18, 24

Spreeuw, J., & Owadally, I. (2013). Investigating the broken-heart effect: a model for short-term dependence between the remaining lifetimes of joint lives. *Annals of Actuarial Science*, *7*(2), 236–257. pages 3, 20

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press. pages 19, 38

Van den Berg, G. J. (2001). Duration models: specification, identification and multiple durations. In *Handbook of econometrics* (Vol. 5, pp. 3381–3460). Elsevier. pages 11

Van den Berg, G. J., Lindeboom, M., & Portrait, F. (2011). Conjugal bereavement effects on health and mortality at advanced ages. *Journal of Health Economics*, *30*(4), 774–794. pages 3, 20

Van der Klaauw, W., & Wolpin, K. I. (2008). Social security and the retirement and savings behavior of low-income households. *Journal of Econometrics*, *145*(1-2), 21–42. pages 2

Wang, Y. (2014). Dynamic implications of subjective expectations: Evidence from adult smokers. *American Economic Journal: Applied Economics*, *6*(1), 1–37. pages 2

Williams Jr, J. R. (2005). Depression as a mediator between spousal bereavement and mortality from cardiovascular disease: Appreciating and managing the adverse health consequences of depression in an elderly surviving spouse. *Southern Medical Journal*, *98*(1), 90–96. pages 3

Wittstein, I. S., Thiemann, D. R., Lima, J. A., Baughman, K. L., Schulman, S. P., Gerstenblith, G., … Champion, H. C. (2005). Neurohumoral features of myocardial stunning due to sudden emotional stress. *New England Journal of Medicine*, *352*(6), 539–548. pages 3

A Mid-point approximation rule

Let $\psi(\cdot)$ be the integral in the numerator of Eq. 12. Suppose that a married spouse has age a and he/she answers the question with the target age ta. We need to integrate $\psi(\tau)$ over τ where τ takes values from a to ta; however, the integral does not have a closed-form solution. Thus, we implement the following strategy to numerically approximate the integral $\psi(\cdot)$:

1. Create a vector $[a, a+D, a+2D, ..., ta-D, ta]$ where $D = \frac{(ta-a)}{\Delta}$. We set $\Delta = 10$.

2. Find the mid-points that are: $[a + \frac{1}{2}D, a + \frac{3}{2}D, ..., ta - \frac{3}{2}D, ta - \frac{1}{2}D]$. There are Δ number of mid-points.

3. Then evaluate function $\psi(\cdot)$ at each mid-point: $[\psi(a + \frac{1}{2}D), \psi(a + \frac{3}{2}D) ... \psi(ta - \frac{1}{2}D)]$

4. Approximate the integral as $\int_a^{ta} \psi(\tau) d\tau \approx D \cdot \Big(\psi(a + \frac{1}{2}D) + \psi(a + \frac{3}{2}D) + ... + \psi(ta - \frac{1}{2}D)\Big)$.

If the value of Δ goes to infinity, then the approximated value equals the true value.

B Predicting unobserved frailties

Once we estimate the parameters, we can impute the expected value (aka. the posterior mean) of unobserved random components. The posterior means of unobserved frailties are calculated by the algorithm discussed in Chapter 12 of Train (2009). The algorithm is implemented as follows:

1. For each couple, generate 500x6 i.i.d. draw from a standard normal distribution. Let Q be the matrix of generated random draws with dimension 500x6. Calculate $\tilde{\eta} = Q\hat{\Lambda}$ to induce correlation among random draws. Here $\hat{\Lambda}$ is the estimated lower-triangular Cholesky matrix.

2. Let $\tilde{\eta}_q$ be the qth row of $\tilde{\eta}$, and let us define $\tilde{\eta}_q = [\tilde{\eta}^o_q, \tilde{\eta}^s_q, \tilde{\eta}^R_q]$. Then for each qth row of $\tilde{\eta}$, predict the conditional likelihood contribution of couple n as the following way:

$$L_{n,q}(\hat{\gamma}|d_n, x_n, z_n, \tilde{\eta}^o_{n,q}, \tilde{\eta}^s_{n,q}, \tilde{\eta}^R_{n,q}) = L^o_{n,q}(\hat{\gamma}^o|d_n, x_n, \tilde{\eta}^o_{n,q}) \cdot L^s_{n,q}(\hat{\gamma}^s, \hat{\gamma}^R|d_n, x_n, z_n, \tilde{\eta}^s_{n,q}, \tilde{\eta}^R_{n,q}) \tag{31}$$

   Here $\hat{}$ indicates the estimates of the corresponding parameters. The conditional likelihood of couple n is shown in Eq. 24.

3. Generate the weighting vector for each couple as

$$w_{n,q} = \frac{L_{n,q}(\hat{\gamma}, d_n, x_n, z_n, \tilde{\eta}^o_{n,q}, \tilde{\eta}^s_{n,q}, \tilde{\eta}^R_{n,q})}{\sum_{\iota=1}^{500} L_{n,\iota}(\hat{\gamma}, d_n, x_n, z_n, \tilde{\eta}^o_{n,q}, \tilde{\eta}^s_{n,q}, \tilde{\eta}^R_{n,q})} \tag{32}$$

4. Calculate the posterior mean of unobserved frailties of couple n as

$$\hat{\eta}_n = \sum_{q=1}^{500} w_{n,q}\tilde{\eta}_{n,q} \tag{33}$$

Here $\hat{\eta}_n$ is a column vector with 6 elements. We use these posterior means, $\hat{\eta}_n$, to predict each couples' survival probabilities, and their expected remaining life-years.

C Dependence index

To simplify the notation, let $\theta_i = (x_i, \eta_i)$ for $i = 1, 2$ and $\theta = [\theta_1, \theta_2]$. Here, I do not make a distinction between the objective and subjective hazards. The dependence index is calculated for both objective and subjective models and the results are reported in Section 4.

Let $g(\theta)$ be the p.d.f of θ. Let $\gamma_1^{GL}(t_1, t_2)$ and $\gamma_2^{GL}(t_1, t_2)$ be the dependence index of Gourieroux and Lu (2015) from the perspective of male and female spouses when spouses have ages (t_1, t_2). Without loss of generality, we show that the derivation of the index from the perspective of a male spouse (aka. spouse 1). The mortality hazards of spouse 1 unconditional on θ

- if both spouses are alive at (t_1, t_2), then

$$\lambda_1(t_1|T_1 \geq t_1, T_2 \geq t_2) = \left(\frac{\partial S(t_1, t_2)}{\partial t_1} \Big/ S(t_1, t_2) \right) \equiv \frac{S_1(t_1, t_2)}{S(t_1, t_2)} \tag{34}$$

- if spouse 2 dies at age t_2 and spouse 1 survives at age t_1, then

$$\lambda_{1|2}(t_1|T_1 \geq t_1, T_2 = t_2) = \left(\frac{\partial S(t_1, t_2)}{\partial t_1 \partial t_2} \Big/ \frac{\partial S(t_1, t_2)}{\partial t_2} \right) \equiv \frac{f_{1|2}(t_1, t_2)}{S_2(t_1, t_2)} \tag{35}$$

$\gamma_1^{GL}(t_1, t_2)$ is defined as the ratio of mortality hazard in 35 to 34. This implies

$$\begin{aligned} \gamma^{GL}(t_1, t_2) &= \frac{\lambda_{1|2}(t_1|T_1 \geq t_1; T_2 = t_2)}{\lambda_1(t_1|T_1 \geq t_1; T_2 \geq t_2)} = \frac{f_{1|2}(t_1, t_2)}{S_1(t_1, t_2)} \cdot \frac{S(t_1, t_2)}{S_2(t_1, t_2)} && (36) \\ &= \frac{\mathbb{E}_\theta(\lambda_1^q(t_1|\theta_1)\lambda_2^p(t_2|\theta_2)S_0(t_1, t_2|\theta))}{\mathbb{E}_\theta(\lambda_1^p(t_1|\theta_1)S_0(t_1, t_2|\theta))} \cdot \frac{\mathbb{E}_\theta(S_0(t_1, t_2|\theta))}{\mathbb{E}_\theta(\lambda_2^p(t_2|\theta_2)S_0(t_1, t_2|\theta))} && (37) \end{aligned}$$

Let $\mathbf{S}_0(t_1, t_2|\theta) = \frac{S(t_1, t_2|\theta))}{\mathbb{E}_\theta(\mathbf{S}(t_1, t_2|\theta))}$, then we can re-define the index in Eq. 37 as:

$$\begin{aligned} \gamma^{GL}(t_1, t_2) &= \frac{\mathbb{E}_\theta(\lambda_1^q(t_1|\theta_1)\lambda_2^p(t_2|\theta_2)\mathbf{S}_0(t_1, t_2|\theta))}{\mathbb{E}_\theta(\lambda_1^p(t_1|\theta_1)\mathbf{S}_0(t_1, t_2|\theta))} \cdot \frac{1}{\mathbb{E}_\theta(\lambda_2^p(t_2|\theta_2)\mathbf{S}_0(t_1, t_2|\theta))} && (38) \\ &= \frac{\mathbb{E}_{\theta|T \geq t}(\lambda_1^q(t_1|\theta_1)\lambda_2^p(t_2|\theta_2))}{\mathbb{E}_{\theta|T \geq t}(\lambda_1^p(t_1|\theta_1))\, \mathbb{E}_{\theta|T \geq t}(\lambda_2^p(t_2|\theta_2))} && (39) \end{aligned}$$

In the last line, $(T \geq t)$ stands for $(T_1 \geq t_1, T_2 \geq t_2)$ and the p.d.f θ conditional on $T \geq t$ is

$$g(\theta|T \geq t) = \frac{S_0(t_1, t_2|\theta)}{\mathbb{E}_\theta(S_0(t_1, t_2|\theta))} g(\theta). \tag{40}$$

After re-arranging the term in the second line of Eq. 39, one obtains the following:

$$\gamma_1^{GL}(t_1, t_2) = \frac{\mathbb{E}_{\theta|T \geq t}\left(\frac{\lambda_1^q(t_1|\theta_1)}{\lambda_1^p(t_1|\theta_1)} \frac{\lambda_1^p(t_1|\theta_1)\lambda_2^p(t_2|\theta_2)}{\mathbb{E}_{\theta|T \geq t}(\lambda_1^p(t_1|\theta_1)\lambda_2^p(t_2|\theta_2))} \right) \mathbb{E}_{\theta|T \geq t}(\lambda_1^p(t_1|\theta_1)\lambda_2^p(t_2|\theta_2))}{\mathbb{E}_{\theta|T \geq t}(\lambda_1^p(t_1|\theta_1))\, \mathbb{E}_{\theta|T \geq t}(\lambda_2^p(t_2|\theta_2))} \tag{41}$$

Using the assumption that conditional on θ the hazard rates follow the MPH form, Eq. 41 can be simplified further as the following way:

$$\gamma^{GL}(t_1,t_2) = \frac{\exp(\beta_1^q + \alpha_1^q t_1)}{\exp(\beta_1^p + \alpha_1^p t_1)} \cdot \frac{\mathbb{E}_{\theta|T\geq t}\Big(\exp(x_1\beta_1 + \eta_1)\cdot \exp(x_2\beta_2+\eta_2)\Big)}{\mathbb{E}_{\theta|T\geq t}(\exp(x_1\beta_1+\eta_1))\,\mathbb{E}_{\theta|T\geq t}(\exp(x_2\beta_2+\eta_2))} \tag{42}$$

$$= \underbrace{\frac{\exp(\beta_1^q + \alpha_1^q t_1)}{\exp(\beta_1^p + \alpha_1^p t_1)}}_{\equiv A} \cdot \underbrace{\Big[\frac{COV_{\theta|T\geq t}\Big(\exp(x_1\beta_1 + \eta_1)\cdot \exp(x_2\beta_2+\eta_2)\Big)}{\mathbb{E}_{\theta|T\geq t}(\exp(x_1\beta_1+\eta_1))\,\mathbb{E}_{\theta|T\geq t}(\exp(x_2\beta_2+\eta_2))} + 1\Big]}_{\equiv B} \tag{43}$$

In Section 4, we report the log of $\gamma^{GL}(t_1,t_2)$ because the log version of the index can be decomposed as a summation of the following two terms:

$$\ln \gamma^{GL}(t_1,t_2) = \ln A + \ln B \tag{44}$$

The first multiplier $\ln A$ captures the magnitude of the dependence due to a bereavement effect. Because for any relation between θ_1 and θ_2, the first multiplier equals zero if there is no bereavement effects. The value of the first multiplier positive (resp. negative) implies that bereavement effect increases (resp. decreases) the mortality hazard rates of surviving spouses. The second multiplier, $\ln B$ captures the dependence via the correlated (un)observed factors that affect spouses' mortality hazard rates. The second multiplier equals zero if there is no dependence via (un)observed factors. The value of the second multiplier that is positive (resp. negative) implies a positive (resp. negative) correlated (un)observed factors.