Subjective Survival Probabilities and Life Tables: Evidence from Europe[*]

Franco Peracchi[†] and Valeria Perotti

Presented at the Living to 100 Symposium

Orlando, Fla.

January 8–10, 2014

[*] Corresponding author: Valeria Perotti, The World Bank, 1818 H Street NW, Washington DC 20433 USA, tel: +1 (202) 458 7488, e-mail: valeria.perotti@gmail.com. We thank Anne Case, Graziella Caselli, Luigi Guiso, Annamaria Lusardi, and participants in Society of Actuaries Living to 100 Conference, the Chicago Booth Center for Population Economics Workshop, and the International SHARE User Conference for helpful comments. This paper uses data from SHARE release 2.2.0. SHARE data collection was primarily funded by the European Commission, with additional funding by the U.S. National Institute on Aging and various national sources. The findings, interpretations and conclusions expressed in this paper are those of the authors and do not necessarily represent the views of the World Bank, its Executive Directors, or the governments they represent.

† Tor Vergata University and EIEF

Subjective survival probabilities and life tables: Evidence from Europe*

Franco Peracchi
Tor Vergata University and EIEF

Valeria Perotti
The World Bank

This version: February 2014

Abstract

Understanding the variability of survival probabilities, both between and within cohorts, is important in order to understand life-cycle decisions under uncertainty. In this paper we analyze the subjective probabilities of survival to specific target ages provided by respondents to the Survey of Health, Ageing and Retirement in Europe (SHARE). To evaluate how these probabilities compare with objective data from life tables, and avoid the problems associated with a naive use of period life tables, we construct cohort life tables from the sequence of period life tables available in the Human Mortality Database and use them to compute actuarial probabilities of survival to the same target ages. We find that male subjective survival probabilities are close to the probabilities computed from our cohort life tables, whereas female subjective probabilities are always lower. We also find that subjective survival probabilities are on average higher for more educated people, those with higher household income, and those with better health. This evidence suggests that both income and health matter for own assessments of subjective survival.

Keywords: Survival; Subjective probabilities; Cohort life tables; SHARE data

JEL Classifications J11; J14

* Corresponding author: Valeria Perotti, The World Bank, 1818 H Street NW, Washington DC 20433 USA, tel: +1 (202) 458 7488, e-mail: `valeria.perotti@gmail.com`. We thank Anne Case, Graziella Caselli, Luigi Guiso, Annamaria Lusardi, and participants in Society of Actuaries Living to 100 Conference, the Chicago Booth Center for Population Economics Workshop, and the International SHARE User Conference for helpful comments. This paper uses data from SHARE release 2.2.0. SHARE data collection was primarily funded by the European Commission, with additional funding by the U.S. National Institute on Aging and various national sources. The findings, interpretations and conclusions expressed in this paper are those of the authors and do not necessarily represent the views of the World Bank, its Executive Directors, or the governments they represent.

1 Introduction

Understanding the variability of survival probabilities, both between and within cohorts, is important in order to understand life-cycle decisions under uncertainty, such as labor supply, consumption, savings and portfolio choices.

Not only do survival probabilities display clear trends over time, but they are also highly heterogeneous within a population, depending on both observed and unobserved characteristics. Several studies have been carried out to assess how survival changes with observed individual characteristics such as education, income, employment or health status. Because this information is typically not available in life tables, which at best contain average survival probabilities for a few coarsely defined population subgroups, researchers have tried to estimate survival probabilities from longitudinal sample surveys. As noted by Delavande and Rohwedder (2011), this approach is problematic because it would require large samples of individuals interviewed over a long period of time with low attrition rates.

An attractive alternative is to directly ask survey respondents about their subjective probabilities of survival to certain target ages. According to Bernheim (1990), eliciting subjective probabilities is better than asking about expected values because survey respondents may lack a clear and common definition of "expectation", thus limiting interpretation and comparability of responses. Further, as pointed out by Dominitz and Manski (1997), subjective probabilities can be checked for internal consistency using the laws of probability, and can be directly compared across individuals and with known event frequencies. The development of the literature on subjective probabilities is reviewed by Manski (2004), who shows how the initial skepticism has been dispelled by the empirical evidence on the ability and willingness of survey respondents to provide probabilistic assessments of future events.

Subjective survival probabilities are numbers between 0 and 100 that survey respondents attach to the event of surviving to a specified target age, with 0 corresponding to no chance of survival and 100 to sure survival. Considering individual-specific measures of expected longevity instead of actuarial data is justified by three observations. First, actuarial mortality forecasts for a given country are usually computed by taking into account only a limited set of variables, typically age, gender, cohort and, sometimes, geographical region, while subjective survival probabilities are collected by large nationally-representative household surveys that provide information on a wide array of socio-demographic characteristics which are relevant for predicting mortality and may be associated with considerable heterogeneity in beliefs. Second, subjective longevity expectations have been found to be reasonable good predictors of future mortality at the individual level (Smith, Taylor and Sloan 2001). Third, individuals are likely to take important decisions, such as when to retire or whether to purchase life insurance, based on their own longevity expectations rather than what is predicted by demographers, regardless of whether they have more or less accurate information.

Since life tables remain the most easily available source of information on mortality, understanding their relationship to subjective survival probabilities elicited from surveys is crucial. Starting from the seminal work by Hamermesh (1985), several papers have compared subjective and life-table survival probabilities using either a single period life table (Hurd and McGarry 1995, 2002) or time averages of period life tables (Guiso et al. 2005; Hurd et al. 2005; Balia 2013). An alternative, but less direct approach, is to compare self-reported life expectancy (the answer to the question "About how long do you think you will live?") to actuarial life expectancy (Puri and Johnson 2007). Unfortunately, all these papers ignore the fact that period life tables cannot be directly compared to subjective survival probabilities or self-reported life expectancy because the information contained in a period life table is cross-sectional and refers to a different cohort for each age. So far, only a few authors have based the comparison on cohort life tables (Ludwig and Zimper 2013; Perozek 2008; Peracchi and Perotti 2009).

The aim of this paper is to compare subjective survival probabilities obtained from a population survey to survival probabilities obtained from properly constructed cohort life tables. For the former we use data from the first wave of the Survey of Health, Ageing and Retirement in Europe (SHARE), a nationally representative survey of people aged 50+ living in several European countries. The SHARE data have already been employed by other authors. For example, Guiso et al. (2005) and Balia (2013) employ them to compare subjective survival probabilities to survival probabilities from period life tables, while Delavande and Rohwedder (2011) employ them to study differentials in subjective survival probabilities by wealth, income and education. For constructing the cohort life tables we use a sequence of country- and gender-specific period life tables obtained from the Human Mortality Database (HMD).

The remainder of this paper is organized as follows. Section 2 describes our data and our procedure for constructing cohort life tables. Section 3 describes our approach to modeling survival probabilities. Section 4 presents our empirical results. Finally, Section 5 discusses our results and concludes.

2 Data

We begin by describing the micro-level information collected by SHARE on subjective survival probabilities. We then discuss the available life-table data, the way they should be used for comparison with subjective survival probabilities, and our procedure for constructing cohort life tables.

2.1 The SHARE data

Our data on subjective survival probabilities are obtained from the first wave of SHARE, a cross-national longitudinal survey of non-institutionalized individuals aged 50+, carried out in 2004 in 11 European countries, namely Austria, Belgium, Denmark, France, Germany, Greece, Italy, the Netherlands, Spain, Sweden, and Switzerland.[1]

The design of the survival probability questions is described in Guiso et al. (2005). Before being asked about own survival, interviewees are asked the following warm-up question in order to become familiar with the response format:

> *"I have some questions about how likely you think various events might be. When I ask a question I'd like for you to give me a number from 0 to 100. Let's try an example together and start with the weather. What do you think the chances are that it will be sunny tomorrow? For example, '90' would mean a 90 percent chance of sunny weather. You can say any number from 0 to 100."*

The information on subjective survival probabilities is then collected by asking the following question:

> *"What are the chances that you will live to be age T or more?"*

The target age T depends on the age of the respondent: it is equal to 75 for those aged 50–65, to 80 for those aged 66–70, to 85 for those aged 71–75, to 90 for those aged 76–80, to 95 for those aged 81–85, to 100 for those aged 86–95, to 105 for those aged 96–100, and to 110 for those aged 101–105.

As pointed out by Guiso et al. (2005), in order to estimate the whole distribution of the random variable "age at death" one would need to ask several questions using different target ages for the same person, as done for example in the U.S. Health and Retirement Study (HRS). To keep the questionnaire short, SHARE did not adopt this strategy and asked only one question.

[1] We use data from SHARE Release 2.2.0, as of August 19, 2009. A detailed description of the survey can be found in Börsch-Supan et al. (2005).

To reduce the impact of cognitive problems, we exclude the oldest old from our analysis and focus on people between 50 and 80 years of age. Thus, we consider the subjective probabilities of survival to ages 75, 80, 85, and 90. Another reason for excluding the oldest old is that the fraction of institutionalized individuals, who are not represented in the SHARE survey, is much higher among the oldest old, especially in some countries.

As for nonsampling error, the fraction of nonrespondents to the survival probability question is usually below 10%, except in Italy (10.1%), Spain (12.3%) and France (18.1%). The low rate of nonresponse is crucial for our analysis. As remarked by Manski (2004), "showing that respondents are willing and able to respond to probabilistic questions is an obvious prerequisite for substantive interpretation of the data" (p. 1342).

In principle, the longitudinal dimension of SHARE may be exploited to control for unobserved time-invariant determinants of subjective survival probabilities. However, the high attrition rate of the survey (ranging from 25% in Belgium to 41% in Spain) suggests that it may be better to limit the analysis to its first wave in order to avoid potential biases. In fact, if attrition is partly due to mortality among wave-1 respondents, and if actual and subjective survival probabilities are correlated, then attrition and subjective survival probabilities are also likely be correlated. This correlation would lead to sample selection bias that may offset the advantage of controlling for unobserved heterogeneity.

Another reason for not exploiting the longitudinal information in SHARE is panel conditioning, namely the effect that repeated interviews may have on respondents' behavior (Lazarsfeld 1940; Sturgis et al. 2009). For example, being asked questions about own survival may raise interest in the issue, hence wave-1 respondents may acquire information and learn more about their survival chances before they are interviewed again in wave 2. Thus, their answer in wave 2 may differ from that in wave 1 even in the absence of changes in health status or in other individual characteristics.

To assess the predictive validity of subjective survival probabilities in SHARE, we follow Hurd and McGarry (2002) and Winter (2008) and estimate several logit regression models where the outcome variable is a binary indicator equal to one if the respondent died between the two waves and to zero otherwise, and the covariates include the subjective survival probability and standard socio-demographic characteristics such as age, gender, education, and self-reported health status. Consistently with the results in Winter (2008), we find that the subjective survival probability is indeed a good predictor of mortality, although its statistical significance is somewhat reduced once we introduce additional controls for self-reported conditions and diseases.

Table 1 reports descriptive statistics for the sample of respondents to the survival probability question in wave 1. The overall sample consists of 18,190 individuals (8,429 males and 9,761 females), while the national samples range from a minimum of 796 individuals in Switzerland and 1,324 individuals in Denmark to a maximum of 2,594 individuals in Sweden and 3,216 individuals in Belgium. The table reports pooled means for all the variables that we use in the analysis and standard deviations (SD) for the non-binary variables. Max grip is the maximum of the measurements (two on each hand) taken using a hand-grip dynamometer, while obese is the fraction of people with body-mass index (the ratio of self-reported weight in kg. to the square of self-reported height in meters) greater or equal to 30. Notice that, compared to men, the women in our sample are less educated, have a lower household income and a weaker grip strength, are more obese, and have a lower prevalence of heart attack and chronic lung disease but a higher prevalence of arthritis and osteoporosis.

A common problem with surveys collecting subjective probabilities is heaping of responses at "focal" values, such as 0, 50, and 100. This problem is discussed in detail in Gan et al. (2005) for the HRS, Hurd et al. (2005) and Balia (2013) for SHARE, and Peracchi and Perotti (2009) for the Italian PLUS survey. Bruine de Bruin et al. (2002) argue that "50" answers may reflect epistemic uncertainty rather than probabilistic thinking. An alternative interpretation is the modal response hypothesis (MRH) of Lillard and Willis (2001), where respondents are uncertain about the objective probability, so they provide the mode of the distribution of all possible values. Hill et al. (2004) estimate a model based on the MRH for subjective survival probabilities in the HRS and find that uncertainty is lower among respondents with higher education and higher cognitive scores. Manski and Molinari (2010) argue that focal responses to probability questions should be interpreted as interval data arising from rounding. They use the information in the 2006 wave of the HRS to identify various patterns of response that differ in the extent of rounding. They then replace each individual's response with an interval that contains the original response and whose width is based on the individual's rounding practice identified in the response patterns. This approach has the advantage of providing more reliable evidence compared to studies that simply ignore the rounding problem. However, it becomes much more difficult to employ regression methods to study the relationship between these interval probabilities and standard demographic variables.

In our sample, about 25% of male and female respondents answer "50 percent" and this percentage does not change much with target age. Further, for both men and women, the fraction of people answering "100 percent" declines with target age, while the fraction answering "0 percent" increases with target age. After estimating a logit model for the probability of answering "50 percent", we find that the probability of this particular focal response increases with age and is lower for individuals in good self-reported health status, but is not systematically associated with other personal characteristics such as gender, education, household income, or self-reported chronic conditions.

2.2 The Human Mortality Database

To construct the life-table counterparts of subjective survival probabilities, we exploit the information contained in the Human Mortality Database (HMD), a joint project by the University of California at Berkeley and the Max Planck Institute for Demographic Research, which provides data on objective mortality for a large number of countries. Because the HMD does not provide data for Greece whereas data for unified Germany are available only from 1991 onwards, we exclude Greece and Germany from our analysis.

The core data contained in the HMD is a set of period life tables, reporting age- and gender-specific death rates in a given year for a given country. A period life table for year t contains the probability q_{xt} of dying at age x in year t conditional on survival to age $x-1$, for all ages x in the range $[0, \overline{x}]$, where $\overline{x}$ is usually set equal to 110. To a first approximation, an individual of age x in year t was born in year $b = t - x$, so the information contained in a period life table refers to a different birth cohort for each age. For example, the 1970 period life table reports the probability of dying at age $0, 1, \ldots, 110$ in year 1970. As a consequence, the probability of dying at age 0 reported in that table refers to individuals born in 1970, while the probability of dying at age 1 refers to individuals born in 1969, and so on.

A different type of life tables contained in the HMD are cohort life tables, which provide age-specific death rates for a given birth cohort. A sequence of cohort life tables would be the appropriate source of data in order to construct the actuarial survival probabilities for SHARE respondents. In fact, the proper life-table counterpart of the subjective probability of survival to age T provided by a SHARE respondent of age a born in year b would be the product

$$P_T = \prod_{x=a}^{T-1} (1 - q_{x,b+x}), \tag{1}$$

where $q_{x,b+x}$ is taken from the life table of the cohort born in year b. Being based on historical data, cohort life tables are usually available only for cohorts that are nearly extinct. For example, the HMD database contains cohort life tables only for the cohorts born before 1917. Because we study the expectations of SHARE respondents aged 50–80 and interviewed in 2004, we need to forecast future life-table death rates for the cohorts born between 1924 and 1954.

2.3 Construction of cohort life tables

Although the period life tables in the HMD are not suited for direct comparisons with SHARE subjective survival probabilities, they can be used to construct cohort life tables for all the cohorts covered by SHARE.

Suppose that one has available a set of period life tables spanning the period from year $\underline{t}$ to year $\bar{t}$. Then, by re-arranging the death rates for a given cohort b, we obtain a life table showing the death rates for that cohort over the age range from $\max(\underline{t} - b, 0)$ to $\bar{t} - b$. Notice that we obtain a different segment of the life cycle for each cohort. For example, since we confine attention to life tables spanning the period from $\underline{t} = 1970$ to $\bar{t} = 2004$,[2] for the 1924 cohort we have death rates from age 46 to age 80, for the 1925 cohort we have death rates from age 45 to age 79, and so on until the 1954 cohort, for which we have death rates from age 16 to age 50. Since the target age T in the SHARE questionnaire is higher than the oldest age $\bar{t} - b$ for which death rates are available for a given cohort b, we need to forecast mortality for all ages from $\bar{t} - b + 1$ to T. We do this by using a forecasting model estimated on past death rates.

To forecast mortality beyond the ages available for each cohort, we adopt the method proposed by Lee and Carter (1992), which specifies the death rate at age x in year t, q_{xt}, as a log-linear function of a mortality index χ_t,

$$\log q_{xt} = \beta_{0x} + \beta_{1x}\chi_t + \epsilon_{xt}. \tag{2}$$

The mortality index χ_t is identified by singular value decomposition methods and is modeled as a low-order ARIMA process. This model is fairly simple, can easily be estimated,[3] and generates probabilistically consistent forecasts and forecast intervals. If the mortality index follows a random-walk with drift, then the Lee-Carter model produces point forecasts that are very similar to those obtained from a linear cohort trend model for the age-specific log-odds of death. Dowd et al. (2010)

[2] It turns out that extending the period further into the past, e.g. 1950–2004, would create problems in estimating models with age-specific trends for men, as their age-specific death rates are stable or even increasing until about 1970 in many European countries and, unlike women, show a clear downward trend only after 1970. We use data until 2004 because this is the year of the first wave of SHARE, and because we want to keep the last few years of available data (up to 2009 for most countries) for assessing the quality of our out-of-sample forecasts.

[3] We use the Stata routine provided by the Princeton University Office of Population Research at `http://data.princeton.edu/eco572/LeeCarter.html`.

show that more sophisticated forecasting models do not provide substantial improvements over the Lee-Carter method.

Thus, to construct the life-table counterparts of subjective survival probabilities, we proceed in three steps. First, we estimate the Lee-Carter model, separately by country and gender, using the HMD data for the period 1970–2004. The data available after 2004 are kept aside to assess the forecasting accuracy of the estimated model. Next, we use the estimates to derive forecasts of q_{xt} for all years beyond 2004. Finally, we compute the life-table counterpart of subjective survival probabilities as in (1), that is, as the product of the forecasts for each cohort. For example, to construct the life-table counterpart of the subjective probability of survival to age 75 for individuals aged 50 in 2004, we forecast the death rates of the 1954 cohort from age 51 to age 75. For individuals aged 60 in 2004, we only need to forecast the death rates of the 1944 cohort from age 61 to age 75. Notice that the model for death rates at age 50 is estimated using data for the cohorts born between 1920 and 1954, the model for death rates at age 51 is estimated using data for the cohorts born between 1919 and 1953, and so on until age 89, for which the model is estimated using data for the cohorts born between 1880 and 1914. Thus, all age-specific models are estimated using the same number of years, but the models for mortality rates at older ages are based on data from older cohorts.

The number of death rates to be forecasted for each cohort depends on the distance between the age of the cohort at the time of the SHARE interview and the target age of the survival probability question for the same cohort. For SHARE participants aged 50–65, the target age is 75 and therefore the minimum distance is 10 years and the maximum distance is 25 years. For participants in older age groups (66–70, 71–75 and 76–80), the minimum distance is 10 years and the maximum distance is 14 years (the target ages are 80, 85 and 90 respectively). Accordingly, death rates for older ages, which are more difficult to forecast, have shorter forecasting periods and are never forecasted beyond year 2019. On the other hand, the furthest year in which we need predictions for death rates for age 75 is 2030.

In order to assess the forecasting accuracy of the Lee-Carter model, we compare the projected survival rates obtained from the model to the actual survival rates from 2005 till the last year of the available HMD data (2008 for Austria and Italy, 2010 for Sweden, and 2009 for all the other countries). We find that the projections are very close to the actual values, although the model tends to slightly underestimate survival at very old ages, especially for men in Belgium, Denmark, the Netherlands and Sweden.

Figures 1 and 2 show the time trend of the observed log-odds of death by age and gender over the available period (solid curves), along with the forecasts from the Lee-Carter model (dashed curves). The vertical line corresponds to year 2004. The gaps between the actual rates and the forecasted rates denote combinations of age and year for which a forecast is not needed, as we are only interested in forecasting the death rates of the cohorts in the SHARE sample between their current age and the target age of the survival probability question. In some cases, for example in Denmark and the Netherlands, the forecasts are rather flat, reflecting the relative stability of survival probabilities over the last two decades.

3 Modeling survival probabilities

The procedure described in the previous section allows us to construct the proper life-table counterparts of the subjective probabilities of survival to a given age provided by SHARE respondents. This section describes how we interpret and model these two sets of probabilities.

Let π_i denote the *objective* probability that the ith individual in a given cohort will survive to a certain target age, let P_i denote the *life-table* survival probability for the demographic group to which the ith individual belongs, and let p_i denote her *subjective* survival probability. Also let X_i denote the set of individual characteristics collected by the SHARE survey. We partition this set as $X_i = (X_{i1}, X_{i2})$, where X_{i1} is the subset of characteristics which are also available in the life tables and X_{i2} is the subset of of additional characteristics which are only available in the SHARE survey. The first subset includes age, target age, gender, birth year, and country, while the second subset includes variables that are likely to affect individual beliefs about own survival but cannot be controlled for when using life tables, such as educational attainments, household income, health status, and whether parents are still alive.

Projecting π_i linearly on the full set of covariates X_{i1} and X_{i2}, objective individual survival probabilities can be represented as

$$\pi_i = \alpha_0 + \alpha_1^\top X_{i1} + \alpha_2^\top X_{i2} + u_i, \tag{3}$$

where u_i is the projection error which, by construction, has mean zero and is uncorrelated with all the variables in X_{1i} and X_{i2}. Since the π_i are unobservable, this model cannot be estimated.

Life-table survival probabilities may be interpreted as the conditional mean of objective individual survival probabilities given the subset of covariates X_{i1} available in the life tables, that is, $P_i = \mathrm{E}(\pi_i|X_{i1})$. Projecting P_i linearly on the subset of covariates in X_{i1}, life-table survival probabilities can be represented as

$$P_i = \gamma_0 + \gamma_1^\top X_{i1} + U_i, \tag{4}$$

where U_i is a random error that has mean zero and is uncorrelated with X_{i1}. Notice that if the conditional mean of X_{i2} given X_{i1} is linear, namely $\mathrm{E}(X_{i2}|X_{i1}) = \mu + B^\top X_{i1}$, and U_i is mean independent of X_{i1}, then $\gamma_0 = \alpha_0 + \mu^\top \alpha_2$ and $\gamma_1 = \alpha_1 + B\alpha_2$.

When forming their beliefs about own survival, individuals have access to a much larger information set than X_{i1}. For example, they have information on important predictors of survival, such as their long-run income prospects, their current health and the actual survival of their parents. Many of these variables are contained in X_{i2}, which is not observed by the demographers who compile the life tables. Projecting p_i linearly on the full set of covariates X_{i1} and X_{i2}, subjective survival probabilities can be represented as

$$p_i = \beta_0 + \beta_1^\top X_{i1} + \beta_2^\top X_{i2} + v_i, \tag{5}$$

where v_i is the error in the linear projection of p_i on X_{i1} and X_{i2} which, by construction, has mean zero and is uncorrelated with all the variables in X_{i1} and X_{i2}. Unlike model (3), model (5) can be estimated using the full set of available covariates.

If we now assume that $\mathrm{E}(v_i|X_{i1}, X_{i2}) = 0$ and again assume that $\mathrm{E}(X_{i2}|X_{i1}) = \mu + B^\top X_{i1}$, then

$$p_i = \delta_0 + \delta_1^\top X_{i1} + V_i, \tag{6}$$

where $\delta_0 = \beta_0 + \mu^\top \beta_2$ and $\delta_1 = \beta_1 + B\beta_2$ are the analogs of the expressions for γ_0 and γ_1 in (4), and V_i is a random error that has mean zero and is uncorrelated with X_{i1}. By comparing the estimates of β_1 from model (5) with those of δ_1 from model (6), we can evaluate the impact of omitting the variables in X_{i2}. In turn, this may help us understand the impact of omitting them in (4). We can also analyze which characteristics in X_{i1} and X_{i2} help explain differences between subjective and life-table survival probabilities by estimating a probit model for the probability that $p_i - P_i > 0$, that is, subjective survival probabilities exceed life-table survival probabilities.

4 Empirical results

We now present the results obtained by estimating models (4)–(6). Because of the limitations of our data, what we cannot do is to produce micro-level estimates of model (3) for the objective probability of survival, and compare them to the estimates of our models for the life-table and the subjective probabilities of survival.

4.1 Age patterns of survival probabilities

Figures 3 and 4 compare group-specific averages of subjective and life-table survival probabilities for groups defined by country, gender, age, and target age. The solid line corresponds to the age average (a 3-year centered moving average) of subjective survival probabilities, the shaded area corresponds to its two-standard error band, the dotted line corresponds to the survival probabilities based on cohort life tables obtained from the Lee-Carter forecasting model (which we call *longitudinal* for short), and the broken line corresponds to the survival probabilities based on the 2004 life table only (which we call *cross-sectional* for short). To ensure comparability, subjective survival probabilities have been weighted using country-specific population proportions by age and gender in 2004.

Longitudinal probabilities are always higher than cross-sectional probabilities, and are always higher for women than for men. This means that using a single life table, even the most recent, to predict survival to a given age for non-extinct cohorts leads to underestimate survival probabilities, especially for countries where the downward trend in mortality is steeper.

For a fixed target age, both subjective and life-table survival probabilities are generally increasing with age. In other words, a shorter distance between the current age of the respondent and the given target age is associated with higher subjective and life-table probabilities of surviving to that target age. Further, subjective and life-table survival probabilities both decrease as we increase the target age. While for life-table probabilities all this is not surprising, because of the way they have been constructed, it is reassuring to find consistency between subjective probabilities and the basic formula (1).

A comparison of subjective and longitudinal survival probabilities reveals important differences by gender. For men, longitudinal probabilities tend to lie within the two-standard error band around subjective probabilities. The exceptions are Belgium and France, where subjective probabilities seem to understate survival at younger ages (that is, longitudinal probabilities are above the two-standard error band around subjective probabilities), and Denmark and the Netherlands, where subjective probabilities seem to overstate survival at older ages (that is, longitudinal probabilities are below the two-standard error band around subjective probabilities).

For women, on the contrary, overstatement is never observed, while understatement is observed at younger ages in all countries except Denmark. Further, except for Denmark, the differences between longitudinal and subjective probabilities tend to get smaller at older ages. These results are broadly in line with the cross-sectional evidence in Guiso et al. (2005) and Hurd et al. (2005) for Europe, and the longitudinal evidence in Perozek (2008) for the United States. Perozek (2008) offers an interpretation of this result based on the idea that individuals are the most qualified to predict their own survival, and suggesting that the longevity gender gap is in fact narrowing compared to the predictions from the life tables.

One may wonder whether gender differences in subjective survival probabilities reflect systematic differences in the way men and women form their probability assessments. To address this issue, we compare the answers that men and women give to the warm-up question about the weather ("What do you think the chances are that it will be sunny tomorrow?"). After estimating a regression model for the answer to this warm-up question, we find no systematic differences by gender, age or education.

4.2 Regression results for life-table survival probabilities

Table 2 shows the least squares estimates of the regression model (4) for the longitudinal probability of survival to a given age, computed using cohort life tables. The basic covariates in X_{i1} include a linear term in age, indicators for target age, and their interactions with the linear age term. The model is estimated separately for men and women, first by country and then, in the last column of the table, by pooling the data and adding to the basic covariates in X_{i1} a set of country fixed effects, with Italy as the reference country. The number of observations in each regression is just the number of available demographic groups. Country-specific weights are used to reflect the importance of each group in the population.

Because individuals of the same age interviewed in the same year belong approximately to the same cohort and we confine attention to the cross-sectional sample of wave-1 respondents, we cannot distinguish between age and cohort effects. In principle, we could take advantage of the fact that, in some countries, wave-1 interviews were conducted in both 2004 and 2005. We do not exploit this source of variation because it is only available for France and Belgium and, in any case, the variation would be small.

The regression results confirm our previous graphical evidence that life-table survival probabilities increase with age, decrease with target age and, everything else equal, are higher for women than for men. The coefficients on the constant term, which provide estimates of the objective probability of surviving to age 75 for a person aged 63 (the average age in the SHARE sample), show an interesting pattern: for both men and women, France is at the top end, while Denmark and the Netherlands are at the bottom end. Notice that the estimated slope coefficients are smaller in absolute value for women than for men, and are also numerically very similar across countries.

The latter finding, which is in line with the results in White (2002) of increasing cross-country convergence of mortality profiles (but not necessarily mortality levels), justifies the practice that we adopt in the remainder of this paper of pooling the data and capturing heterogeneity across countries through a set of country fixed effects.

4.3 Regression results for subjective survival probabilities

Table 3 shows the results from several regression models estimated separately by least squares for men and women using the pooled data. To ensure comparability with the results in the previous section, the individual observations are weighted using country-specific population proportions by age and gender in 2004.

For each gender, column (1) of the table presents the estimates of model (6) which only includes a constant and the variables in X_{i1}, namely a linear term in the deviation of age from age 63, indicators for target age and their interactions with the linear age term, plus a set of country dummies. Column (2) presents the estimates obtained by adding to X_{i1} the log of de-meaned household income (the deviation of household income from Euro 50,000, which is about equal to the average household income in our sample) and an indicator for upper-secondary education completed. Column (3) presents the estimates obtained when X_{i2} also includes a number of health-related variables, namely an indicator for poor or bad self-reported health status, an objective measure of grip strength (the deviation of maxgrip from its gender specific average), an indicator for being obese, and indicators for the parents being alive. Finally, column (4) presents the estimates obtained by adding to X_{i2} a set of binary indicators for self-reported chronic conditions. Our list of self-reported conditions roughly corresponds to that used by Case and Paxson (2005) and includes indicators for hearth attack, high blood pressure, high blood cholesterol, stroke, diabetes, chronic lung disease, asthma, arthritis, osteoporosis, hip or femoral fracture, cancer, and other conditions. The coefficient on the constant term in column (4) represents the expected subjective probability of survival to age 75 for an Italian aged 63, who did not complete secondary education, has a household income of Euro 50,000, a grip strength of 44 if male and 27 if female, is not obese, has no parent still alive, and suffers of no chronic conditions.

First compare the coefficients in the “short” regression in column (1) with those from the pooled regression in the last column of Table 2. For men, the coefficients on the constant terms are quite similar, while the age profile of life-table probabilities seems to be steeper than the age profile of subjective probabilities, especially for older target ages. On the other hand, life-table probabilities decrease with target age considerably faster than subjective probabilities. For women, instead, the coefficient on the constant term is much smaller and the age pattern is flatter for subjective probabilities than for life-table probabilities, although the age coefficient is higher for subjective probabilities at older ages.

The coefficients on the variables in X_{i1} change little when, in column (2), we add as regressors the indicator for educational attainments and the log of household income. The coefficients on these two variables are positive and strongly statistically significant for both men and women. Hurd and McGarry (1995, 2002) also find a positive effect of education on subjective survival probabilities, while Lleras-Muney (2005) finds a similar effect on actual survival. Notice that the coefficient on education is much larger for women than for men, while the coefficient on log household income is somewhat smaller for women than for men. Although adding health-related indicators in columns (3) and (4) lowers the coefficients on education and income relative to column (2), they remain positive and statistically significant, particularly for women. This means that education and income are important predictors of subjective survival even after controlling for self-reported health and family history.

We see little evidence of gender differences in the coefficients on the health-related variables in column (3). The coefficient on the indicator for poor or bad self-reported health is negative, that on grip-strength is positive, and both are strongly statistically significant. The coefficient on being obese is instead negative but not statistically significant. Information on parents' survival also appear to matter, especially for men, as subjective probabilities are significantly higher for respondents whose parents are still alive. Similar results have been found by Hamermesh (1985) and Hurd and McGarry (1995, 2002) using data for the United States, while Balia (2013) finds no such evidence for Italy in the first wave of SHARE.

Now consider the "long" regression (4), which also includes the indicators of chronic conditions. As expected, the coefficients on these indicators are generally negative. Further, the coefficients which are strongly statistically significant are all negative and are those associated with severe chronic conditions, such as a heart attack or cancer. In line with the results in Case and Paxson (2005), the association between diseases and survival probabilities is different for men and women. In particular, the coefficients on self-reported heart attack, diabetes, asthma and cancer are much larger for men than for women, whereas the opposite is true for the coefficients on high blood pressure and arthritis.

Comparing the "short" regressions (1) and (2) to the "long" regressions (3) and (4), the coefficients on the variables in X_{i1} change little. A notable exception is the fact that, for men, the linear age term in the "long" regressions is larger and strongly statistically significant. A possible explanation is that the age coefficient in the "short" regressions reflects two opposite effects: one is the positive effect of a shorter distance from the target age (for example, the probability of reaching age 75 conditional on survival to age 61 is higher than the probability of reaching the same age conditional on survival to age 60, all else equal), the other is age-related deterioration in health status, which has a potentially negative effect on survival probabilities. In the "long" regressions, the coefficient on age is much higher because the second effect is weakened by the presence of health-related variables and indicators of chronic conditions.

4.4 Differences between subjective and life-table survival probabilities

Table 4 presents the estimates of a probit model for the probability that the difference between subjective and life-table survival probabilities is positive. As we did for the regression models discussed earlier, we first estimate the model with the basic set of covariates X_{i1} and then progressively include education, household income and the health-related variables.

Interestingly, the coefficient on the constant term is always negative and statistically significant for women, while for men it is positive but statistically significant only after controlling for the health-related variables. Relative to a target age of 75, the probability of a positive difference between subjective and life-table probabilities tends to be higher at higher target ages for men, while for women results are not so clear. This probability is also higher for women with secondary or higher educational attainments, and for men whose parents are still alive. For both men and women, household income and chronic conditions do not seem to matter much, while reported poor health status and lower levels of grip strength are associated with a lower probability of a positive difference between subjective and life-table survival probabilities.

5 Discussion

In this paper we compared subjective probabilities of survival to a target age reported by SHARE respondents to objective survival probabilities from cohort life tables constructed using the set of period life tables available in the HMD. By comparing age-specific averages, we find that male subjective survival probabilities are relatively close to the probabilities computed from cohort life tables, whereas female subjective probabilities are always lower than life-table probabilities. The distance between subjective and life-table probabilities increases with target age for men and decreases with target age for women. Further, gender differences persist even after controlling for a broad set of individual characteristics.

Two alternative interpretations are possible for these results. If individuals provide more accurate forecasts of their own survival than life tables, as argued by Perozek (2008), then our findings imply that the longevity gender gap is narrowing, with men's survival probabilities approaching that of women.

Alternatively, if life tables provide more accurate forecasts than individuals, then women appear to considerably underestimate their own survival. This second interpretation implies that women tend to behave myopically, with important consequences for financial decisions, such as savings and planning for retirement, and the life insurance world, such as product development, pricing and risk management for retirement products.

The second interpretation is consistent with the results in Puri and Robinson (2007) who find that men are more likely to be optimistic about survival than women, and that optimism is positively correlated with the probability of saving and the amount saved. In fact, empirical evidence from the U.S. shows that older women have very low levels of financial literacy, and that the majority of them have not planned for retirement (Lusardi and Mitchell 2008). In addition, poverty rates for older non-married women are much higher than for married women or nonmarried men (Karamcheva and Munnell 2007).

The interpretation that women may underestimate their survival probabilities is also in line with the evidence that women tend to report worse health than men but are less likely to die at each age. Case and Paxson (2005) find that gender differences in self-rated health in the U.S. can be entirely explained by differences in the distribution of chronic conditions, in particular the higher prevalence among women of painful but not life-threatening conditions such as arthritis, headache and some respiratory conditions. Peracchi and Rossetti (2012) find that this is partly true also in Europe. Thus, it seems that the paradox of worse self-rated health for women and higher mortality for men cannot be solely attributed to systematic gender differences in reporting style and has a fairly straightforward explanation. Are there similar "objective" reasons why women may underestimate their own survival probabilities? One possibility is that women tend to live longer than men, but to live more years with serious chronic conditions. Overall, there are not large gender differentials in survival in healthy conditions, and this may be relevant for our results if SHARE respondents think of survival as survival in healthy conditions.

This paper also studies the relationship between subjective survival probabilities and a detailed set of individual characteristics. In particular, we find that subjective survival probabilities are on average higher for more educated people, people whose household income is greater, people with better subjective and objective health status (as measured by the grip strength), and people without chronic conditions. This evidence suggests that both income and health matter for own assessments of subjective survival,

More generally, our results suggest the need of going beyond life-tables, which only provide an aggregate description of survival for coarsely defined demographic groups. This is especially important if one is interested in understanding the distributional effects of policies in crucial domains such as social security, public health or long-term care.

References

Balia, S. (2013). "Survival Expectations, Subjective Health and Smoking: Evidence from SHARE." *Empirical Economics*, September: 1-28.

Bernheim, B.D. (1990). "How Do the Elderly Form Expectations: An Analysis of Responses to New Information." In *Issues in the Economics of Aging*, edited by D.A. Wise. Chicago: University of Chicago Press.

Börsch-Supan, A., A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegrist and G. Weber, eds. (2005). *Health, Ageing and Retirement in Europe: First Results From the Survey of Health, Ageing and Retirement in Europe*. Mannheim: MEA.

Bruine de Bruin, W., P.S. Fischbeck, N.A. Stiber and B. Fischhoff (2002). "What Number is 'Fifty-Fifty'? Redistributing Excessive 50% Responses in Elicited Probabilities." *Risk Analysis*, 22: 713–723.

Case, A. and C. Paxson (2005). "Sex Differences in Morbidity and Mortality." *Demography*, 42: 189–214.

Delavande, A. and S. Rohwedder (2011). "Differential Mortality in Europe and the United States: Estimates Based on Subjective Probabilities of Survival." *Demography*, 48: 1377–1400.

Dominitz, J. and C.F. Manski (1997). "Using Expectations Data to Study Subjective Income Expectations." *Journal of the American Statistical Association*, 92: 855–867.

Dowd, K., A.J. G. Cairns, D. Blake, G.D. Coughlan, D. Epstein and M. Khalaf-Allah (2010). "Backtesting Stochastic Mortality Models: An Ex Post Evaluation of Multiperiod-Ahead Density Forecasts." *North American Actuarial Journal*, 14: 281–298.

Gan, L., M.D. Hurd and D. McFadden (2005). "Individual Subjective Survival Curves." In *Analyses in the Economics of Ageing*, edited by D.A. Wise. Chicago: University of Chicago Press.

Guiso, L., A. Tiseno and J. Winter (2005). "Expectations." In *Health, Ageing and Retirement in Europe: First Results From the Survey of Health, Ageing and Retirement in Europe*, edited by A. Börsch-Supan, A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegrist and G. Weber. Mannheim: MEA.

Hamermesh, D.S. (1985). "Expectations, Life expectancy, and Economic Behavior." *Quarterly Journal of Economics*, 100: 389–408.

Hill, D., M. Perry and R.J. Willis (2004). "Estimating Knightian Uncertainty From Survival Probability Questions on the HRS." Paper prepared for the Conference "Labor and Social Interaction in Economics," December 16–17, 2004. Eitan Berglas School of Economics, Tel Aviv University.

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

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

Hurd, M. S. Rohwedder and J. Winter (2005). "Subjective Probabilities of Survival: An International Comparison." Unpublished manuscript. Santa Monica, CA: RAND Corporation.

Karamcheva, N. and A.H. Munnell (2007). "Why Are Widows so Poor?" Brief No. 7–9. Center for Retirement Research, Boston College.

Lazarsfeld, P.F. (1940). "'Panel' Studies." *Public Opinion Quarterly*, 4: 122–128.

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

Lillard, L.A. and R.J. Willis (2001). "Cognition and Wealth: The Importance of Probabilistic Thinking." Working Paper No. 007. Retirement Research Center, University of Michigan.

Lleras-Muney, A. (2005). "The Relationship Between Education and Adult Mortality in the United States." *Review of Economic Studies*, 72: 189–221.

Ludwig, A. and A. Zimper (2013). "A Parsimonious Model of Subjective Life Expectancy." *Theory and Decision*, 75: 519–541.

Lusardi, A. and O.S. Mitchell (2008). "Planning and Financial Literacy: How Do Women Fare?" *American Economic Review: Papers & Proceedings*, 98: 413–417.

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

Manski, C.F. and F. Molinari (2010). "Rounding Probabilistic Expectations in Surveys." *Journal of Business & Economic Statistics*, 28: 219–231.

Peracchi, F. and V. Perotti (2009). "Subjective Survival Probabilities and Life Tables: An Empirical Analysis of Cohort Effects." *Genus*, 65: 23–57.

Peracchi, F. and C. Rossetti (2012). "Heterogeneity in Health Responses and Anchoring Vignettes" *Empirical Economics*, 42: 513–538.

Perozek, M. (2008). "Using Subjective Expectations to Forecast Longevity: Do Survey Respondents Know Something We Don't Know?" *Demography*, 45: 95–113.

Puri, M. and D.T. Robinson (2007). "Optimism and Economic Choice." *Journal of Financial Economics*, 86: 71–99.

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

Sturgis, P., N. Allum and I. Brunton-Smith (2009). "Attitudes Over Time: The Psychology of Panel Conditioning." In *Methodology of Longitudinal Surveys*, edited by P. Lynn. Chichester: Wiley.

White, K.M. (2004). "Longevity Advances in High-Income Countries, 1955–96." *Population and Development Review*, 28: 59–76.

Winter, J. (2008). "Expectations and Attitudes." In *Health, Ageing and Retirement in Europe (2004–2007): Starting the Longitudinal Dimension*, edited by A. Börsch-Supan, A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegrist and G. Weber. Mannheim: MEA.

Table 1: Descriptive statistics for the sample of respondents to the survival probability question in wave 1 of SHARE.

	Men			Women		
	Obs.	Mean	SD	Obs.	Mean	SD
Age (yrs)	8429	62.8	8.3	9761	62.6	8.4
Upper secondary educ. (%)	8360	51.3		9683	42.6	
Household gross income (1,000 EUR)	8429	49.6	57.0	9761	44.1	54.3
Target age:						
Age 75	8429	62.8		9761	64.1	
Age 80	8429	15.8		9761	14.5	
Age 85	8429	12.4		9761	12.3	
Age 90	8429	9.0		9761	9.1	
Distance to target age (yrs)	8429	15.6	4.5	9761	15.7	4.6
Subjective probability (%)	8429	63.3		9761	64.4	27.9
Poor SRH (%)	8427	23.6		9761	27.6	
Max grip	8127	44.3	10.1	9225	27.1	7.2
Obese (%)	8376	16.2		9560	18.2	
Father alive (%)	8405	9.8		9726	10.1	
Mother alive (%)	8357	24.6		9660	25.4	
Chronic conditions (%):						
Heart attack	8420	14.3		9757	8.2	
High blood pressure	8420	28.3		9757	31.6	
High blood cholesterol	8420	21.5		9757	20.6	
Stroke	8420	3.8		9757	2.7	
Diabetes	8420	9.8		9757	8.4	
Chronic lung disease	8420	6.0		9757	4.4	
Asthma	8420	4.5		9757	5.2	
Arthritis	8420	13.8		9757	24.9	
Osteoporosis	8420	1.5		9757	11.3	
Hip/femoral fracture	8420	1.6		9757	1.5	
Cancer	8420	4.7		9757	6.1	
Other conditions	8429	26.0		9761	28.2	
Country: (%)						
Austria	8429	8.5		9761	9.4	
Belgium	8429	17.9		9761	17.5	
Denmark	8429	7.6		9761	7.0	
France	8429	12.2		9761	12.4	
Italy	8429	11.6		9761	11.6	
Netherlands	8429	13.7		9761	13.3	
Spain	8429	9.4		9761	10.6	
Sweden	8429	14.6		9761	14.0	
Switzerland	8429	4.5		9761	4.2	

Table 2: Coefficients of the best linear predictor of life-table survival probabilities, by gender and country (AT: Austria, BE: Belgium, DK: Denmark, FR: France, IT: Italy, NL: Netherlands, ES: Spain, SE: Sweden, CH: Switzerland).

	Men									
	AT	BE	DK	FR	IT	NL	ES	SE	CH	Pooled
Constant	76.573	78.332	69.852	79.611	78.866	73.400	78.191	79.402	80.869	79.733
Age (63=0)	0.257	0.135	0.391	0.288	0.137	0.187	0.280	0.162	0.109	0.216
Target age 80	-15.862	-16.667	-20.522	-13.331	-15.224	-20.329	-15.338	-15.906	-14.216	-16.485
Target age 85	-39.120	-42.294	-45.979	-34.910	-37.536	-48.898	-37.827	-40.708	-37.211	-39.979
Target age 90	-65.252	-70.195	-69.528	-61.655	-64.530	-72.986	-64.285	-69.106	-65.449	-65.937
Age × Target age 80	0.620	0.591	0.888	0.546	0.668	0.841	0.602	0.636	0.586	0.684
Age × Target age 85	0.762	0.867	1.041	0.695	0.881	1.209	0.793	0.883	0.850	0.842
Age × Target age 90	0.771	1.021	1.008	0.717	1.013	1.123	0.888	0.971	0.978	0.873
# observations	31	31	31	31	31	31	31	31	31	279
# covariates	7	7	7	7	7	7	7	7	7	15
Adj. R^2	0.999	0.999	0.998	0.999	0.999	0.999	0.999	0.999	0.999	0.992
	Women									
	AT	BE	DK	FR	IT	NL	ES	SE	CH	Pooled
Constant	87.988	89.141	78.138	91.202	90.210	86.067	91.154	87.044	89.518	91.348
Age (63=0)	0.101	0.109	0.449	0.109	0.048	0.231	0.018	0.188	0.118	0.156
Target age 80	-10.439	-9.259	-14.316	-7.193	-8.293	-11.246	-8.337	-10.543	-8.788	-9.721
Target age 85	-29.755	-27.339	-35.561	-21.236	-23.774	-30.958	-25.247	-29.833	-25.638	-28.064
Target age 90	-59.676	-55.692	-60.805	-47.473	-50.135	-60.472	-54.616	-58.927	-53.881	-56.443
Age × Target age 80	0.289	0.247	0.662	0.203	0.261	0.387	0.216	0.353	0.255	0.289
Age × Target age 85	0.401	0.381	0.874	0.254	0.358	0.530	0.338	0.513	0.370	0.477
Age × Target age 90	0.615	0.593	0.858	0.462	0.576	0.734	0.632	0.698	0.551	0.674
# observations	31	31	31	31	31	31	31	31	31	279
# covariates	7	7	7	7	7	7	7	7	7	15
Adj. R^2	1.000	1.000	0.999	1.000	1.000	0.999	1.000	1.000	1.000	0.993

Table 3: Estimated regression models for subjective survival probabilities by gender.

	Men				Women			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Constant	73.393 **	73.646 **	78.496 **	79.628 **	70.428 **	70.527 **	75.503 **	77.015 **
Age	0.122	0.158	0.515 **	0.581 **	-0.154	-0.075	0.099	0.149
Target age 80	-7.510 †	-8.503 *	-9.815 *	-9.524 *	-12.828 **	-12.557 **	-14.214 **	-13.547 **
Target age 85	-16.048 †	-14.526	-12.107	-11.162	-22.846 *	-21.925 *	-23.091 *	-23.491 **
Target age 90	7.224	9.925	22.886	25.473	-48.673 **	-49.491 **	-46.920 **	-46.278 **
Age × Target age 80	-0.315	-0.039	0.204	0.144	1.359 †	1.285 †	1.720 *	1.624 *
Age × Target age 85	-0.085	-0.216	-0.530	-0.610	1.172	1.071	1.189	1.240
Age × Target age 90	-2.212	-2.397 †	-3.164 *	-3.332 *	1.960 †	2.007 †	1.821	1.792
Upper secondary educ.		3.014 **	0.535	0.653		5.695 **	3.293 **	3.193 **
Log household income		1.582 **	1.045 *	1.074 *		1.165 **	0.759 *	0.732 *
Poor SRH			-13.338 **	-10.464 **			-11.390 **	-9.607 **
Max grip			0.162 **	0.142 **			0.218 **	0.181 *
Obese			-1.674	-1.118			-0.026	0.878
Father alive			5.600 **	5.651 **			0.436	0.492
Mother alive			3.032 **	2.885 **			1.681	1.645
Heart attack				-5.127 **				-3.431 †
High blood pressure				-0.583				-2.578 *
High blood cholesterol				0.470				-0.136
Stroke				-4.335				-1.751
Diabetes				-2.705				-0.517
Chronic lung disease				-4.116 *				-4.871 †
Asthma				-3.501				1.352
Arthritis				-0.533				-1.555
Osteoporosis				-2.526				-1.154
Hip/femoral fracture				4.953				-3.176
Cancer				-7.033 **				-1.841
Other conditions				-1.912 †				-0.320
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# observations	8400	8275	7851	7844	9711	9564	8753	8750
# covariates	15	17	22	34	15	17	22	34
Adj. R^2	0.080	0.087	0.137	0.145	0.083	0.097	0.136	0.141
BIC	79683.4	78468.9	73798.9	73753.0	92078.0	90539.2	82378.9	82400.9

Note: † $p < .10$; * $p < .05$; ** $p < .01$

Table 4: Estimated probit models for the probability of a positive difference between subjective and life-table survival probabilities.

	Men				Women			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Constant	0.035	0.037	0.202 **	0.224 **	-0.457 **	-0.451 **	-0.273 **	-0.251 **
Age	-0.002	-0.000	0.013 *	0.015 *	-0.010 †	-0.007	-0.000	0.000
Target age 80	0.173	0.134	0.100	0.114	-0.349 †	-0.369 †	-0.545 **	-0.521 *
Target age 85	1.259 **	1.270 **	1.373 **	1.397 **	-0.080	-0.031	0.034	0.039
Target age 90	1.110	1.093	1.713 *	1.781 *	0.513	0.491	0.882	0.911
Age × Target age 80	-0.025	-0.017	-0.010	-0.013	0.071 †	0.074 †	0.108 **	0.105 *
Age × Target age 85	-0.112 **	-0.114 **	-0.125 **	-0.127 **	0.042	0.037	0.032	0.033
Age × Target age 90	-0.049	-0.049	-0.085	-0.089	0.019	0.021	-0.005	-0.006
Upper secondary educ.		0.117 *	0.036	0.039		0.194 **	0.120 *	0.117 *
Log household income		0.045 *	0.023	0.022		0.045 *	0.031	0.030
Poor SRH			-0.490 **	-0.405 **			-0.386 **	-0.358 **
Max grip			0.007 **	0.006 *			0.010 **	0.010 **
Obese			-0.099 †	-0.083			0.009	0.033
Father alive			0.234 **	0.242 **			-0.020	-0.020
Mother alive			0.115 *	0.109 †			0.069	0.067
Heart attack				-0.164 *				-0.075
High blood pressure				0.020				-0.100 *
High blood cholesterol				0.018				0.048
Stroke				-0.148				-0.008
Diabetes				-0.113				0.067
Chronic lung disease				-0.153				-0.136
Asthma				-0.040				0.108
Arthritis				-0.035				-0.087 †
Osteoporosis				-0.042				0.059
Hip/femoral fracture				0.206				0.063
Cancer				-0.142				0.034
Other conditions				-0.059				0.018
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# observations	8400	8275	7851	7844	9711	9564	8753	8750
# covariates	15	17	22	34	15	17	22	34
Pseudo R^2	0.014	0.017	0.048	0.051	0.041	0.047	0.062	0.064
BIC	11620.5	11434.4	10569.4	10635.7	11844.3	11649.0	10591.3	10668.1

Note: † $p < .10$; * $p < .05$; ** $p < .01$

Figure 1: Male log-odds of death, by country and age (HMD data). The solid curve is for the observed log-odds, the dashed curve for the forecasts from the Lee-Carter model. The vertical line corresponds to year 2004.

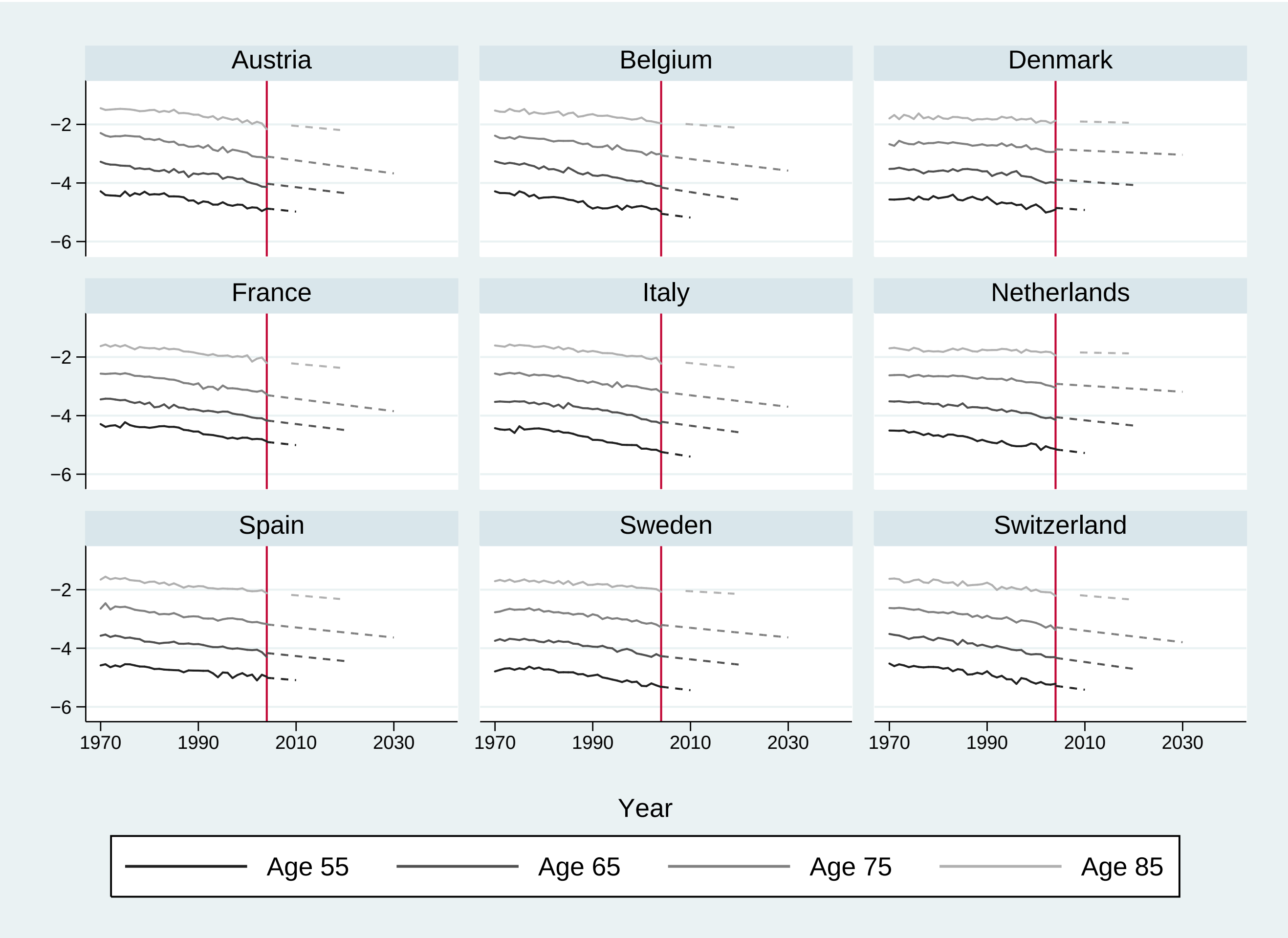

Figure 2: Female log-odds of death, by country and age (HMD data). The solid curve is for the observed log-odds, the dashed curve for the forecasts from the Lee-Carter model. The vertical line corresponds to year 2004.

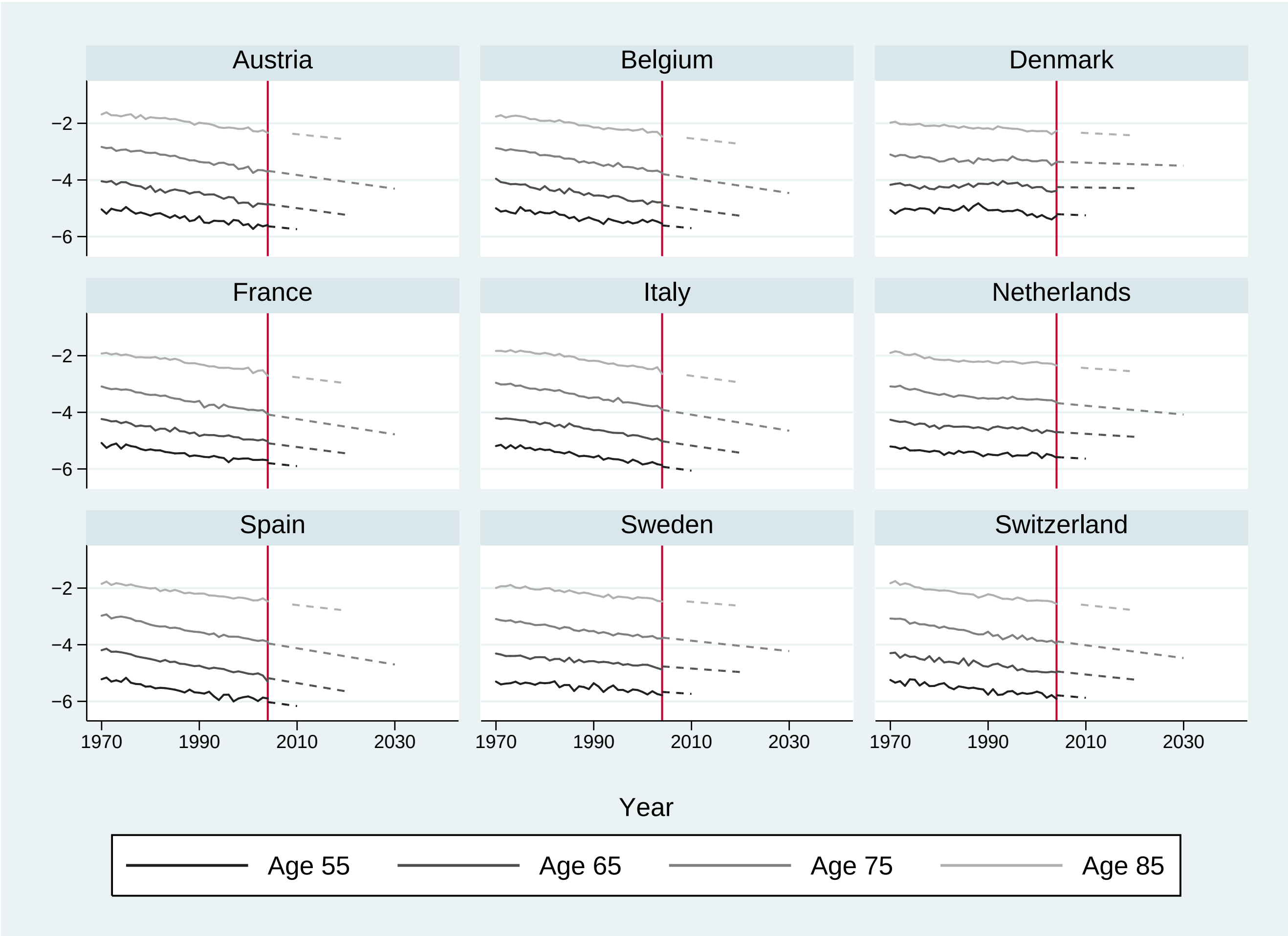

Figure 3: Male average life-table predictions and SHARE subjective survival probabilities (3-year centered moving average), by country and target age.

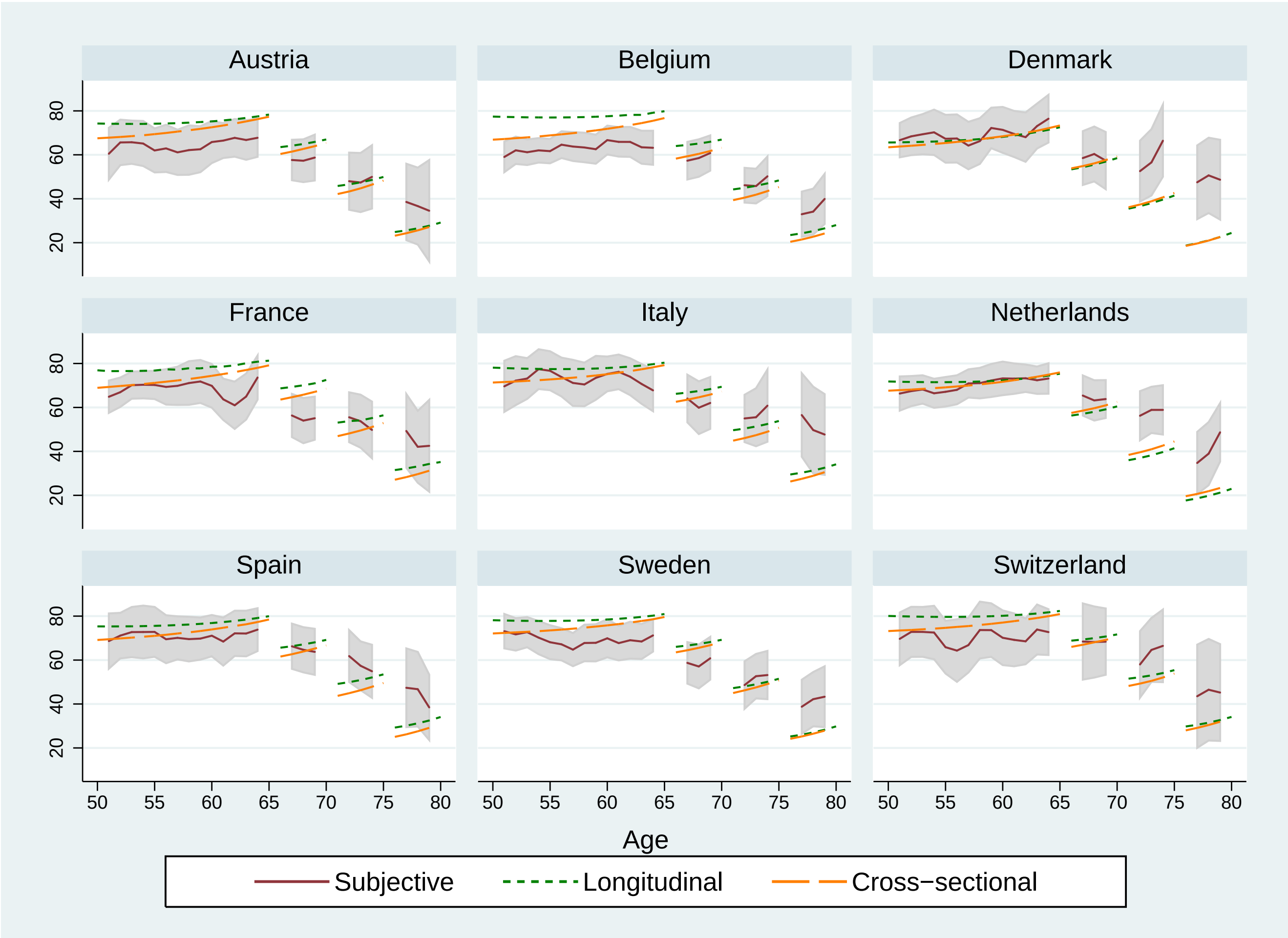

Figure 4: Female average life-table predictions and SHARE subjective survival probabilities (3-year centered moving average), by country and target age.