Subjective survival expectations and morbidity patterns of European citizens aged 50 years and older. A cross-sectional and longitudinal perspective based on SHARE data.

Apostolos Papachristos FIA CERA

Department of Statistics and Insurance Science

University of Piraeus

A Thesis submitted to the University of Piraeus in

accordance with the requirements for the degree of

Doctor of Philosophy.

Piraeus 2021

Αναμενόμενη Υποκειμενική Επιβίωση και επίπεδα νοσηρότητας Ευρωπαίων ηλικίας 50 ετών και άνω. Μια μελέτη βασισμένη στα δεδομένα του SHARE.

Απόστολος Παπαχρήστος FIA CERA

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

Πανεπιστήμιο Πειραιά

Πειραιάς 2021

Abstract

Individuals vary their perceptions towards own survival over their lifetime. Some are consistently more optimists, and they believe that the chances of own survival are high. The systematic investigation of the characteristics of both optimist and pessimist individuals will reveal the underlying reasons for differences in survival beliefs.

The results of this Thesis show that there are common factors affecting self-reported subjective survival probabilities. On the one hand, factors such as a large number of children, higher socio-economic status, frequent physical activity, frequent consumption of fruits or vegetables and eggs or legumes as well as better quality of life are associated with higher subjective survival, younger biological age and longevity. On the other hand, factors such as poor self-rated health, more limitations in Activities of Daily Living, a larger number of chronic diseases, poor memory, poor writing skills and depression are associated with lower subjective survival, older biological age and higher mortality.

One of the research objectives of this Thesis is the development of a method for the quantification of 'Self-perceived age'. 'Self-perceived age' incorporates survival information from the general population life tables as well as the self-reported subjective survival probabilities. The results of the last part of this Thesis indicate clearly that Subjective Survival Probabilities and 'Self-perceived age' are both strong and independent predictors of mortality. This implies that both Subjective Survival Probabilities and 'Self-perceived age' include 'survival information' important for predicting actual survival. Finally, the patterns of 'Self-perceived age' are also comparable to the patterns of Biological Age and Subjective Age.

The introduction of the new concept of 'Force of subjective mortality' allows the calculation of subjective survival probabilities which correspond to shorter time horizons. For instance, the average SHARE Wave 6 respondent reports a subjective survival probability which corresponds to the next 14 years. However, in order to assess the explanatory power of subjective survival probabilities on actual mortality, the subjective survival probabilities for the next SHARE Wave (i.e. 2 years) are required. These probabilities can be estimated using the 'Force of subjective mortality'. The results indicate clearly that both Subjective Survival Probabilities which correspond to the next two years (i.e. SHARE Wave 7) and 'Self-perceived age' are both strong and independent predictors of mortality and include information important for predicting actual survival.

Περίληψη

Οι προσδοκίες των ανθρώπων για την μελλοντική επιβίωση τους μεταβάλλονται κατά τη διάρκεια της ζωής τους. Κάποιοι είναι σταθερά πιο αισιόδοξοι και πιστεύουν ότι οι πιθανότητες της μελλοντικής επιβίωσης τους είναι υψηλές. Η συστηματική διερεύνηση των χαρακτηριστικών τόσο των αισιόδοξων όσο και των απαισιόδοξων ανθρώπων μπορεί να αποκαλύψει τους λόγους διαφορών της αναμενόμενης Υποκειμενικής Επιβίωσης.

Τα αποτελέσματα αυτής της διατριβής δείχνουν ότι υπάρχουν κοινοί παράγοντες που επηρεάζουν τις Υποκειμενικές Πιθανότητες Επιβίωσης. Αφενός, παράγοντες όπως ένας μεγάλος αριθμός παιδιών, υψηλότερη κοινωνικό-οικονομική κατάσταση, συχνή σωματική δραστηριότητα, συχνή κατανάλωση φρούτων ή λαχανικών και αυγών ή οσπρίων καθώς και καλύτερη ποιότητα ζωής σχετίζονται με υψηλότερες Υποκειμενικές Πιθανότητες Επιβίωσης, νεανική Βιολογική Ηλικία καθώς και μακροζωία. Από την άλλη πλευρά, παράγοντες όπως κακή υγείας, περισσότεροι περιορισμοί στις Δραστηριότητες της καθημερινής ζωής, μεγαλύτερος αριθμός χρόνιων ασθενειών, κακή μνήμη, κακές δεξιότητες γραφής και κατάθλιψη σχετίζονται με χαμηλότερες Υποκειμενικές Πιθανότητες Επιβίωσης, εξελιγμένη Βιολογικής Ηλικίας και υψηλότερη θνησιμότητα .

Ένας από τους ερευνητικούς στόχους της παρούσας διατριβής είναι η εξέλιξη μιας καινοτόμου μεθόδου για την εκτίμηση της «Υποκειμενική Ηλικία Επιβίωσης», με βάση τις Υποκειμενικές Πιθανότητες Επιβίωσης και τους πινάκες θνησιμότητας του γενικού πληθυσμού. Τα αποτελέσματα της διατριβής δείχνουν σαφώς ότι τόσο οι Υποκειμενικές Πιθανότητες Επιβίωσης όσο και η «Υποκειμενική Ηλικία Επιβίωσης» είναι ανεξάρτητοι παράγοντες που προβλέπουν θνησιμότητα. Αυτό συνεπάγεται ότι οι δύο ποσότητες αυτές περιλαμβάνουν «πληροφορίες επιβίωσης» σημαντικές για την πρόβλεψη της πραγματικής επιβίωσης και μπορούν να χρησιμοποιηθούν για την εκτίμηση Δεικτών Γήρανσης του πληθυσμού. Τέλος, τα πρότυπα της «Υποκειμενικής Ηλικίας Επιβίωσης» είναι επίσης συγκρίσιμα με τα πρότυπα της Βιολογικής Ηλικίας και της Υποκειμενικής Ηλικίας.

Η εισαγωγή της έννοιας «Υποκειμενική Ένταση Θνησιμότητας» επιτρέπει τον επαν-υπολογισμό των Υποκειμενικών Πιθανοτήτων Επιβίωσης για μικρότερο χρονικό ορίζοντα εκτίμησης. Για παράδειγμα οι συμμετέχοντες στο SHARE Wave 6 εκτιμούν την πιθανότητα επιβίωσης τους για τα επόμενα 14 έτη, κατά μέσο όρο. Ωστόσο, για την εκτίμηση της ακρίβειας των Υποκειμενικών Πιθανοτήτων Επιβίωσης μικρότεροι χρονικοί ορίζοντες εκτίμησης απατούνται (π.χ. 2 έτη μέχρι το επόμενο SHARE Wave). Οι Υποκειμενικές Πιθανότητες Επιβίωσης με μικρότερο χρονικό ορίζοντα εκτιμώνται με την χρήση της «Υποκειμενική Ένταση Θνησιμότητας».

Τα αποτελέσματα δείχνουν ότι αφενός οι Υποκειμενικές Πιθανότητες Επιβίωσης που αντιστοιχούν σε χρονικό ορίζονται 2 ετών (μέχρι το SHARE Wave 7) και αφετέρου η «Υποκειμενική Ηλικία Επιβίωσης», είναι ανεξάρτητοι παράγοντες που προβλέπουν θνησιμότητα και περιλαμβάνουν πληροφορίες σημαντικές για την πρόβλεψη της πραγματικής επιβίωσης.

Dedication and acknowledgements

This Thesis is dedicated to my parents and all members of my family. I would like to thank my direct supervisor Professor Georgia Verropoulou and the supervising committee for their advice, guidance and support on this research project.

Table of Contents

List of tables

List of figures

1. Chapter 1 – Introduction

1.1. Subjective survival expectations

Individuals vary their perceptions towards own survival over their lifetime. Some are consistently more optimists, and they believe that the chances of own survival are high (Lyubomirsky 2001). The systematic investigation of the characteristics of both optimist and pessimist individuals will reveal the underlying reasons for differences in survival beliefs. It has been noted that estimation of future survival is partially based on past experiences, socio-economic status as well as environmental influences (Griffin, Loh and Hesketh 2013).

While past experiences are likely to be unique for each individual, other factors - such as socio-economic status - may be common for groups of people. Individuals who belong in these 'homogeneous' groups may share similar understanding about factors affecting future survival. This common awareness is acquired through media and health campaigns (Griffin, Loh and Hesketh 2013) For example; the fact that smoking reduces life expectancy is widely understood (Doll et. al. 1994). The identification of common factors affecting subjective survival will allow us to understand the social profile as well as the reasons for which individuals are more or less optimistic regarding their own survival.

Psychological factors are certainly important in forming survival beliefs. For example, depression is associated with lower survival expectations (Rarrange et al. 2016). On the other hand, Freund and Baltes (2002) noted that individuals who use life-management behaviors, such as the Selection-Optimization-Compensation model, also report higher subjective well-being. In other words, individuals who manage to achieve a successful aging report higher life satisfaction (Baltes and Carstensen 1996) and thus higher survival expectations. It is worth noting that survival expectations are incorporated in financial decision making. Teppa (2011) noted that people who expect to live longer prefer to purchase annuities, instead of receiving lump sum payments. However, individual-specific judgment is the main underlying reason for the variability observed in survival beliefs (Hamermesh 1985).

1.2. Objectives of the study

This study has four objectives.

Objective 1: The identification of common factors driving subjective survival probabilities

The first objective is the identification of common factors driving subjective survival probabilities and their allocation into homogeneous groups. The homogeneous groups are based on the subjective survival probabilities direction and consistency with actual mortality patterns.

Common factors increasing subjective and actual survival

- The first group consists of the characteristics of individuals whose subjective survival probabilities are higher than the associated objective survival probabilities but they are also likely to exhibit lower actual mortality.

Common factors decreasing subjective and actual survival

- The second group consists of the features of individuals whose subjective survival probabilities are lower than the associated objective survival probabilities but, they are also likely to exhibit worse actual mortality.

Factors related to divergence of subjective and actual survival

- The third group consists of the features which differentiate the patterns of actual mortality and subjective survival probabilities. These features affect the direction of subjective survival probabilities but not in a manner consistent with actual mortality.

This objective is examined in Chapter 3.

Objective 2: Gender and socio-demographic effects on updating subjective survival probabilities

The second objective is the investigation of the mechanism by which individuals update their subjective survival probabilities. More specifically, the aim is to assess whether subjective survival probabilities are updated consistently with actual mortality patterns, and whether there are variations by gender.

This objective is examined in Chapter 4.

Objective 3: The quantification of 'Self-perceived age'

The third objective is the development of a method for the quantification of 'Self-perceived age'. 'Self-perceived age' incorporates survival information from the general population life tables as well as the self-reported subjective survival probabilities. The validation of 'Self-perceived age' is carried out by examining associations with physical health, cognitive function as well as lifestyle and behavioral risk

factors. I validate further ‘Self-perceived age’ by comparing it with actual mortality patterns as well as expected patterns of ‘Biological Age’. These comparisons are based on the relevant literature.

This objective is examined in Chapter 5.

Objective 4: Assessing the predictive power of subjective survival metrics on mortality

The fourth objective relates to the assessment of the predictive power of subjective survival probabilities as well as of ‘Self-perceived age’. To achieve this objective, I introduce the concept of ‘Force of subjective mortality’ which allows the estimation of subjective survival probabilities which correspond to a short time horizon. On average, SHARE Wave 6 respondents report subjective survival probabilities which refer to the next 14 years (Figure 5.1).

On the other hand, SHARE Wave 7 data, including End of Life interviews which were collected in 2017, two years after SHARE Wave 6. So, there is a mismatch between the time interval of SHARE waves (about 2 years) and the time interval which corresponds to subjective survival probabilities (on average 14 years). The subjective survival probabilities which correspond to the next two years are estimated using the ‘Force of subjective mortality’

This objective is examined in Chapter 6.

1.3. Contribution of the study

This study adds to the existing literature across the following different dimensions.

First, the grouping of sociodemographic, health and lifestyle characteristics according to the subjective survival probabilities consistency with general population mortality and actual mortality patterns reveals the extent to which various features are incorporated in forming survival expectations. In other words, these comparisons provide indications of whether individuals are aware of the effect of a specific factor on mortality.

Second, this study shows that individuals update survival expectations after the occurrence of specific events, consistently with actual mortality patterns. Moreover, the degree of consistency varies by gender. In other words, individuals incorporate in their survival expectations the change in mortality risk, which is the result of the occurrence of specific events.

Third, the methodology of calculating 'Self-perceived age' (SPA) is novel. 'Self-perceived age' is expressed in years, and therefore is easier to understand and communicate than subjective survival probabilities, which are expressed in percentages. 'Self-perceived age' is linked to a population life table and can be compared directly to Chronological Age. The difference between 'self-perceived' and Chronological Age is communicated in years. Furthermore, the calculation of 'Self-perceived age' requires significantly less data compared to the calculation of 'Biological Age'.

Fourth, the introduction of the new concept of 'Force of subjective mortality' allows the calculation of subjective survival probabilities which correspond to shorter time horizons. For instance, the average SHARE Wave 6 respondent reports a subjective survival probability which corresponds to the next 14 years (Figure 6.1). However, in order to assess the explanatory power of subjective survival probabilities on actual mortality, the subjective survival probabilities for the next 2 years are required. These probabilities can be estimated using the 'Force of subjective mortality'.

A distinct advantage of this study is that it uses data from European and American surveys as well as life tables form the Human Mortality Database. The model estimates are based on a large number of respondents and the selection of predictors covers different domains of life. Overall, this study presents findings which are consistent with current literature, and it also includes an extensive analysis on findings and ideas which are unique.

1.4. Structure of the Thesis

The core element of this study is subjective survival probabilities (SSPs). These are self-reported quantities, expressed as probabilities, which reflect the chances of own future survival over a specified period of time.

A literature review of factors affecting the reporting of subjective survival probabilities is presented in Chapter 2. Sections 2.1 and 2.2 present a summary of the patterns of subjective survival probabilities and actual mortality. Sections 2.3 and 2.4 discuss the use of Biological Age and Subjective Age as aging indicators. Sections 2.5 describes the data sources whereas Section 2.6 provides an overview of the predictors and covariates used in the statistical models.

Chapter 3 is dedicated to the identification of common factors driving subjective survival probabilities and their allocation into homogeneous groups. The homogeneous groups are based on the subjective survival probabilities direction and consistency with actual mortality patterns. Section 3.1 presents the sample, the dependent variables and the predictors. The statistical modelling and the results are described in Sections 3.2 and 3.3. Finally, the results are compared to the actual mortality patterns in Section 3.4.

Chapter 4 is dedicated to the update of subjective survival probabilities and their allocation into homogeneous groups. In other words, the extent to which subjective survival probabilities are updated consistently with actual mortality patterns is examined. Section 4.1 presents the sample, the dependent variables and the predictors. The statistical modelling and the results are described in Sections 4.2 and 4.3. Finally, the results are compared to the actual mortality patterns in Section 4.4.

Chapter 5 focuses on the estimation of 'Self-perceived age'. 'Self-perceived age' is a quantity, expressed in years, which incorporates survival-related information from two sources, namely, population life tables and subjective survival probabilities. In Section 5.1 the research objective is explained in detail. Section 5.2 presents the sample, the dependent variables and the predictors. The statistical modelling and the results are described in Sections 5.3 and 5.4. Finally, the results are compared to the actual mortality patterns in Section 5.5.

Chapter 6 is dedicated to the assessment of the predictive power of Subjective Survival Probabilities and 'Self-perceived age' on actual mortality. In Section 6.1 the research objective is explained in detail. Section 6.2 presents the definition and estimation of 'Force of subjective mortality'. The concept is analogous to the 'Force of mortality' in Actuarial Mathematics. The sample, the dependent variables

and the predictors as also presented in this section. The statistical modelling and the results are described in Sections 6.3 and 6.4. Finally, the results are compared to the actual mortality patterns in Section 6.5.

The findings of the Thesis are discussed in Chapter 7. The procedure followed to compare the results with the relevant literature is presented in Section 7.1. The factors associated with higher subjective survival, lower aging and longevity are presented in Section 7.2. The factors associated with lower subjective survival, deteriorating aging and higher mortality are presented in Section 7.3. The factors exhibiting inconsistent associations with subjective survival, biological age and actual mortality are presented in Section 7.4. Section 7.5 outlines some limitation of this Thesis whereas Section 7.6 describes areas for further research. The conclusions of the Thesis are presented in Section 7.7. Finally, there is an Appendix which includes some supplementary results and material and programming code used as part of the analysis.

2. Chapter 2 – Subjective survival probabilities and other indicators of ageing

2.1. Factors affecting the reporting of subjective survival probabilities

Subjective survival probabilities are numbers reflecting individuals' view on likely future survival and they vary considerably (Hamermesh 1985). Past research provides a plethora of explanations to support the variation observed in subjective survival probabilities.

Variation of subjective survival probabilities according to socio-demographic characteristics

Gender is one of the main factors affecting survival expectations. Arpino et al. (2018), using objective survival probabilities estimated from the US Health and Retirement Study, concluded that males tend to report higher survival expectations than females. Another interesting finding is that older people tend to report higher survival expectations compared to younger people. Several researchers have estimated a positive association between age and subjective life expectancy (Griffin et al. 2013; Mirowsky 1999; Ross and Mirowsky 2002) and concluded that people tend to be more optimistic as they get older. Subjective survival probabilities vary also across countries. Rappange et al. (2016) using data from the Survey of Health and Retirement in Europe (SHARE) Wave 2 found a significant variation in subjective survival probabilities across 15 countries in Europe.

A larger number of children is associated with higher survival expectations. Several researchers noted that respondents who live with children report higher survival expectations (Liu et al. 2007; Mirowsky 1999; Ross and Mirowsky 2002). Parental longevity also affects the direction of survival expectations. Hurd and McGarry (1995) using data from the Health and Retirement Study found that subjective survival probabilities are higher if both parents are alive.

Marital status affects the direction of survival expectations but its impact is not always consistent. On the one hand, Balia (2014) noted that widows report higher subjective survival probabilities while Liu et al. (2007) and Rarrange et al. (2016) found that living alone is associated with lower subjective survival expectations. On the other hand, Van Solinge and Henkens (2018) reported that partnership status is not clearly associated with subjective life expectancy.

Higher education and higher income are associated with higher subjective survival probabilities. There is an overall agreement in the literature that individuals with higher socio-economic status tend to be

more optimistic regarding own future survival (Arpino et al. 2018; Rappange et al. 2016; Liu et al. 2007; Mirowsky 1999; Balia 2014).

Variation of subjective survival probabilities according to the status of physical and mental health

Poor physical health is clearly associated with lower survival expectations (Liu et al. 2007; Balia 2014; Hurd and McGarry 1995). In addition, Rappange et al. (2016) found that the number of chronic diseases is associated with lower subjective survival probabilities. Van Solinge and Henkens (2018) using data from the first wave of the Netherlands Interdisciplinary Demographic Institute (NIDI) Work and Retirement Panel, also noted that worse subjective health and the existence of chronical illnesses; cancer or heart related diseases, are associated with lower subjective life expectancy.

Cognitive function and mental health are factors which differentiate subjective survival probabilities. Griffin et al. (2013) using data from the Australian population-based "45 and Up Study", noted that distressed or individuals who have low social connectedness tend to underestimate subjective life expectancy. Furthermore, d'Uva et al. (2017) and Elder (2007) using data from the Health and Retirement Study, concluded that respondents with better numeracy as well as better total recall score tend to report accurate subjective survival probabilities.

Variation of subjective survival probabilities according to behavioural risk factors and social connectedness

Behavioural risk factors such as the frequency of physical activities, smoking status or obesity differentiate subjective survival probabilities. Rappange et al. (2016) and Liu et al. (2007) using data from Europe and Taiwan concluded that individuals who do physical activities and maintain a normal weight tend to report higher subjective survival probabilities.

The evidence on the effect of smoking status on subjective survival probabilities is not conclusive. Khwaja et al. (2007) using data from the Health and Retirement Study, noted that smokers are optimistic whereas never smokers are pessimistic in their survival predictions. Liu et al. (2007) concluded also that current smokers overestimate survival. Furthermore, Balia (2014) using data from SHARE noted that smoking cessation causes the overestimation of subjective survival probabilities. In contrast, other researchers found that smoking is associated with lower survival expectations (Hurd & McGarry 1995; Rarrange et al. 2016)

Dietary habits and social connectedness are differentiating factors of subjective survival expectations. Griffin et al. (2013), comparing subjective to actuarial life expectancy, found that the number of servings of fruits or vegetables consumed per day, is positively associated with subjective life expectancy. Furthermore, Mirowsky (1999) using data form the survey of Aging, Status and the Sense of Control in the United States, estimated a positive association between social connectedness, family support and subjective life expectancy. A positive association between life satisfaction and survival expectations has been estimated in past research (Van Solinge and Henkens 2018, Ross and Mirowsky 2002).

2.2. Actual mortality patterns

Life tables, provided by the Human Mortality Database, reflect the average mortality across all sub-groups in each population. Life tables do vary by country and sex, but they don't differentiate by socio-economic or health status. Nevertheless, past research provides evidence on the impact of a wide range of factors on actual mortality patterns.

Actual mortality patterns based on socio-demographic features

It is well known that females live more than males. On this topic Austad (2006) argues that although women live longer than men, they also tend to report worse health. Socio-economic status also differentiates actual mortality patterns (Antonovsky 1967). Gregg et al. (2003) using data from the Established Populations for Epidemiologic Studies of the Elderly in the United States concluded that higher socio-economic status, represented by education and income, is associated with lower actual mortality. Country of residence is another differentiating factor of actual mortality. White and Holmes (2006) using data from the World Health Organisation Statistical Information Services Mortality Database concluded that mortality rates vary across countries as well as between males and females. Regarding parents' lifespan, Ikeda et al. (2006) using data from the Japan Collaborative Cohort Study for Evaluation of Cancer Risk, concluded that parental longevity is a predictor of mortality for Japanese men and women. The number of children is also a factor which increases the longevity of parents (McArdle et al. 2006).

Marital status does not have a homogeneous effect on mortality. For example, increased mortality for widows is observed during the first year following bereavement (Kaprio et al. 1987) but the increase is smoothed over the following years. Moreover, Ebrahim et al. (1995) using data from the British Regional Heart Study concluded that men who divorced had higher mortality. However, for men who became widowed there was no change in mortality risk.

Actual mortality patterns based on the status of physical and mental health

Self-rated health, disability and chronic diseases are key predictors of mortality and their predictive power is not affected by the incorporation of other covariates known to be associated with mortality (Idler and Benyamini 1997; Verropoulou 2014). The status of mental health and cognitive function differentiates mortality. Depression and poor cognitive function are associated with a shorter lifespan (Wulsin et al. 1999; Smits et al. 1999). Furthermore, Park et al. (2013) using data from a longitudinal study of Korean elderly individuals, concluded that cognitive dysfunction and orientation in time is

associated with higher mortality. The existence of mobility difficulties is also related to mortality. Olaya et al. (2017) using data from the Collaborative Research on Ageing in Europe concluded that more mobility difficulties are associated with higher mortality.

Actual mortality patterns based on lifestyle, behavioural risk factors and social support

Lifestyle and behavioral risk factors such as smoking, Body Mass Index as well as physical exercise are factors which differentiate actual mortality. In particular, smokers and obese as well as underweight are likely to have shorter lifespans (Doll et al. 1994; Solomon and Manson 1997) whereas individuals who do vigorous exercise are likely to live longer (Gregg et al. 2003). Furthermore, Knoops et al. (2004) noted that individuals, who adopted Mediterranean diet and a healthy lifestyle, are more likely to increase their lifespan. Social support or connectedness as well as life satisfaction differentiate actual mortality. Lyyra and Heikkinen (2006) using data from Finland concluded that perceived social support is associated with lower mortality. Furthermore, Netuveli et al. (2012) using data from the British Household Panel Survey noted that quality of life is a predictor of mortality and that improving quality of life is associated with lower mortality. On this topic Buono et al. (1998) reported that greater satisfaction with life and better social and family relations are linked to longevity.

2.3.Biological Age

'Biological Age' is a point-in-tine indicator of an individual's accumulated ageing, and it is estimated from a set of independent biomarkers (Belsky et al. 2015). Individuals, who have the same chronological age but older Biological Age tend to age more rapidly, have more difficulties with physical functioning and have a declining cognitive function. Moreover, Belsky et al. (2015) noted that individuals with older Biological Age also look older. Physical exercise differentiates the estimated Biological Age. Nakamura et al. (1989) noted that elderly Japanese men who regularly did physical exercise had younger biological ages. Furthermore, Jylhävä et al. (2017) concluded that 'Biological Age' is associated with morbidity and mortality.

There are two main challenges related to the quantification of Biological Age; namely data requirements and the calculation algorithm. Several researchers report on the specific data required for the quantification of Biological Age. Johnson (2006) outlines the features biomarkers should have in order to be appropriate for predicting aging. Sebastiani et al. (2017) using data from the Long Life Family Study, examined the correlation of a large set of biomarkers with signatures of aging such as the incident of cancer, cardiovascular disease, diabetes, and mortality. They concluded that a smaller set of 19 independent biomarkers has association with signatures of mortality and morbidity. The second challenge is the development of the algorithm used to estimate Biological Age. Levine (2013) using data from National Health and Nutrition Examination Survey, attempted to compare the credibility of three alternative methodologies; principal components analysis; multiple linear regressions and Klemera and Doubal's (2006) method. She concluded that the Klemera and Doubal's method is the most reliable for predicting mortality.

2.4.Subjective Age

'Subjective Age' indicates how old a person feels and it could diverge from chronological age (Kleinspehn-Ammerlahn et al. 2008). Rubin and Berntsen (2006) noted that 'Subjective Age' varies across individuals' lifespan. In particular, younger respondents feel older whereas older respondents feel younger. The cross-over age is estimated around 25 years old. 'Denial of aging' is a theory that could explain why older individuals feel younger (Barak et al. 2001). According to this theory older adults tend to maintain young subjective age identities. Women feel younger than men; aging is viewed as a particularly negative experience for women in terms of both attractiveness and reproductive potential (Barrett 2005). In addition, better health of parents and having more children are associated with younger 'Subjective Age'. Lower SES is associated with older 'Subjective Age' and this can be partially attributed to less optimistic views about one's own health (Barrett 2003).

Better self-rated health and better cognitive function are related to younger 'Subjective Age' (Stephan et al. 2012; Stephan et al. 2016). Furthermore, Thyagarajan et al. (2019) note that people who feel younger have favorable biomarker profiles and therefore exhibit a lower prevalence of age-related diseases. They also suggest that 'Subjective Age' could act as an early indicator to detect high risk individuals for several age-related diseases. Finally, Levy et al. (2002) noted that negative self-perceptions of aging could reduce actual life expectancy whereas the opposite holds for positive self-perceptions of aging.

2.5.Data

This study uses four main data sources in order to achieve the objectives, described below.

Table 2.1 List of datasets used in relation to the objectives of the study.

Objective No.	Description of objective	Dataset	Use of Human Mortality Database life tables
1	The identification of common factors driving subjective survival probabilities	A cross-sectional dataset from SHARE Wave 6.	Yes
2	How individuals update their subjective survival probabilities	A combined longitudinal dataset from SHARE Wave 6 and Wave 7.	No
3	The quantification of 'Self-perceived age'	A harmonized cross-sectional dataset from HRS Wave 12 and SHARE Wave 6.	Yes
4	Assessing the predictive power of subjective survival metrics on mortality	A combined longitudinal dataset from SHARE Wave 6 and Wave 7.	No

2.5.1. Description of data sources

Survey of Health, Ageing and Retirement in Europe – Wave 6

The first source of data is from wave 6 and wave 7 of the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE (Börsch-Supan et al. 2013) is a cross-national and multidisciplinary panel database with information on health, socio-economic status, and social and family networks. SHARE has been mainly funded by the European Commission and is coordinated centrally at the Mannheim Research Institute for the Economics of Ageing. The data collection of the 6th Wave was completed in November 2015 (Börsch-Supan 2017) and the sampling was carried out in 18 countries. The response rates vary by country; for instance, the household cooperation rates for the wave 6 sample vary from 54% in France to 88% in Croatia. More documentation and information on SHARE can be found at http://www.share-project.org.

The original Wave 6 sample covered 67346 individuals aged 50 or higher. Due to SHARE rules, information about subjective survival probabilities was not collected for 2906 individuals (4.3%), for whom proxy interviews were conducted (Table A1 in the Appendix). These individuals tend to be on average older while they also report poorer self-rated health, more chronic diseases, lower quality of life and relatively high depression levels. In addition, there were 596 individuals with missing values

in the variables of interest (less than 1%). Hence, the Wave 6 sample used in the analysis includes 63844 individuals.

Survey of Health, Ageing and Retirement in Europe – Wave 7

The SHARE Wave 7 dataset became available in April 2019 (Börsch-Supan 2019), the data collection took place in 28 countries in 2017 and covers 76520 individuals.

Survey of Health, Ageing and Retirement – Wave 6 & 7 combined sample

The combined Wave 6 and Wave 7 longitudinal sample covers 51849 individuals. It is notable that 14397 individuals who participated in wave 6 do not participate in wave 7. Furthermore 1985 respondents who participated in W6 but died before W7. A supplementary analysis of the characteristics of these respondents as well as the impact on our results is carried out in this study (Table A3 in the Appendix). The combined Wave 6 and Wave 7 longitudinal sample consists of 51245 individuals aged 50 or older. Due to SHARE rules, information about subjective survival probabilities was not collected for 1764 individuals aged 50 or older (3.4%), for whom proxy interviews were conducted. Hence, the combined Wave 6 and Wave 7 longitudinal sample reduces down to 49505 individuals. The countries included in SHARE waves 6 and 7 are presented in Table 1.2.

Table 2.2 Countries included in SHARE waves 6 and 7.

Country	Wave 6	Wave 7
Austria	✓	✓
Belgium	✓	✓
Bulgaria	×	✓
Croatia	✓	✓
Cyprus	×	✓
Czech republic	✓	✓
Denmark	✓	✓
Estonia	✓	✓
Finland	×	✓
France	✓	✓
Germany	✓	✓
Greece	✓	✓
Hungary	×	✓
Israel	×	✓
Italy	✓	✓
Latvia	✓	✓
Lithuania	×	✓
Luxemburg	✓	✓
Malta	×	✓
Netherlands	×	✓
Poland	✓	✓
Portugal	✓	✓
Romania	×	✓
Slovakia	×	✓
Slovenia	✓	✓
Spain	✓	✓
Sweden	✓	✓
Switzerland	✓	✓

Harmonised Health and Retirement Study - Wave 12 and Harmonised Survey of Health, Ageing and Retirement in Europe – Wave 6

The Health and Retirement Study (HRS) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. It is an age-cohort-based longitudinal panel survey of persons aged 50 and older in the USA. The data collection of the 12th Wave was completed in 2014. The harmonised version of the longitudinal studies, RAND HRS and RAND SHARE, provided by the Gateway to Global Aging Data, was used in order to produce a consistent combined dataset. The original combined SHARE Wave 6 and HRS Wave 12 sample covered 76252 individuals aged 50 or higher. However, as information about future life expectancy is not collected in SHARE for proxy interviews (2906 individuals) and there were 1967 individuals with missing values in the variables of interest, the combined SHARE Wave 6 and HRS Wave 12 sample includes 71379 individuals.

Human Mortality Database

The objective survival probabilities were calculated from the Human Mortality Database (HMD) population life tables. HMD provides both period and cohort life tables. However, cohort life tables are incomplete for individuals aged 50 or above and have to be forecasted (Peracchi and Perotti 2010). Moreover, Peracchi and Perotti (2010) noted that the forecasted cohort life tables may underestimate actual mortality for certain countries. Therefore, in this study we use period life tables by country and sex which refer to the 5-year period 2010-2014.

2.5.2. The calculation of subjective survival probabilities

In the 'Expectations' module of the SHARE questionnaire, in section I of the 'Retirement Plans, Expectations' module of the RAND HRS dataset and in Section I of the 'Retirement and Expectations' module of the RAND SHARE dataset , respondents were asked to state their survival expectations on a scale from 0 to 100 as follows:

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 at the interview; it is set at age 75 for respondents aged 50–65, at age 80 for respondents aged 66–70, at age 85 for respondents aged 71–75, at age 90 for respondents aged 76–80, at age 95 for respondents aged 81–85, at age 100 for respondents aged 86–95, at age 105 for respondents aged 96–100 and at age 110 for respondents aged 101 or higher. The difference between the respondents' actual age and his/her target age is the prediction interval 'N', in years. For example, a respondent aged 65 is asked to report his chances of surviving up to age 75; in this case, the target age is 75 years and the prediction interval is 10 years.

The reported survival expectations were divided by 100 in order to derive the individual subjective survival probabilities (SSPs).

$$Subjective\ survival\ probability = \frac{What\ are\ the\ chances\ that\ you\ will\ live\ to\ be\ age\ [T] or\ more?}{100}$$

Note that SSPs correspond to time horizons between 5 and 25 years.

2.5.3. The calculation of objective survival probabilities

The life table survival probabilities (OSPs) are calculated from the corresponding country and sex specific complete period HMD life tables. The reported SSPs correspond to a specific prediction interval, starting from current age up to the target age. Therefore, the OSPs should cover the same time horizon (Peracchi and Perotti 2010). Hence,

$$OSP_{x,N} = \prod_{t=1}^{N}(1 - q_{x+t})$$

where 'x' is the chronological age of the respondent, 'N' is the prediction interval and q_{x+t} is the probability of dying between ages $x+t$ and one year later.

2.6.Predictors and covariates

A wide range of explanatory variables, allocated in groups, covering many aspects of respondents' lives have been incorporated in this study. Table 1.2 summarises which variables are used in relation to each objective.

Demographic characteristics

In relation to the first objective, this group includes chronological age (in years), gender, marital status (widowed, divorced, never married, separated, married, in partnership), the number of parents of the respondent still alive at the time of the survey, the number of children of the respondent as well as country of residence.

In relation to the second objective, this group includes chronological age (in years), the increase in chronological age between SHARE waves 6 and 7, gender and the change in marital status. The change in marital status is re-coded in 4 levels (no change in marital status between waves 6 and 7, widowed, divorced and other changes in marital status). The aim of this variable is to isolate the impact of widowhood and divorce on updating SSPs. Country of residence is used as a control variable.

In relation to the third objective, this group includes chronological age (in years), chronological age squared /100, gender, marital status, the number of children of the respondent, the number of parents of the respondent still alive at the time of the survey as well as country of residence; all these are used as control variables.

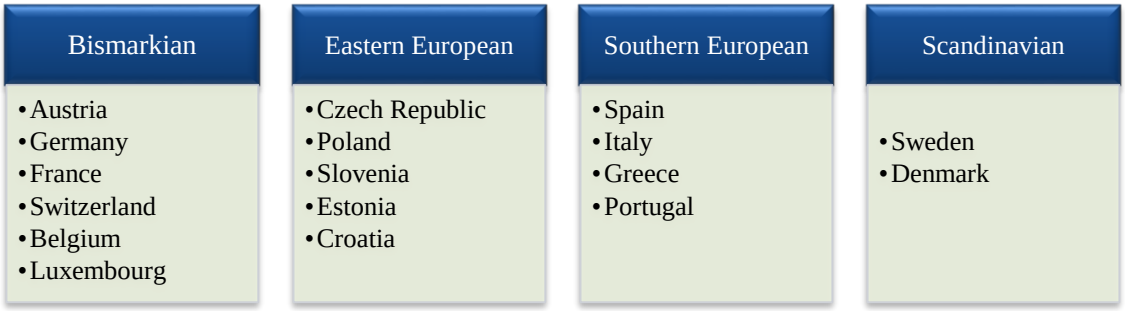
Socioeconomic status

In relation to the first and second objectives, socioeconomic status is represented by the "equivalised" individual income in quartiles. The equivalised income per individual was calculated using the reported household income and the OECD-modified equivalence scale. This scale, first proposed by Haagenars et al. (1994), assigns a value of 1 to the household head, of 0.5 to each additional adult member and of 0.3 to each child. Socioeconomic status is also represented by educational attainment, considered in 4 levels, based on the ISCED-97 classification, Primary (code 1), Lower Secondary (code 2), Upper Secondary (codes 3 & 4) and Tertiary (codes 5 & 6).

In relation to the third objective, total household income in quartiles and total wealth in quartiles are calculated separately for each welfare state (Eikemo et al. 2008). In Europe, the Bismarkian welfare state includes Austria, Germany , France , Switzerland, Belgium and Luxembourg; the Eastern

European welfare state includes Czech Republic , Poland, Slovenia, Estonia and Croatia; the Southern European welfare state includes Spain, Italy, Greece and Portugal whereas the Scandinavian welfare state includes Sweden and Denmark (Figure 1.1). USA is treated as a separate welfare state. Educational level is considered in 3 categories (lower than Upper Secondary, Upper Secondary and Tertiary).

Figure 2.1 Allocation of European countries in welfare states.



Physical & Mental Health

In relation to the first objective, physical health includes the number of chronic conditions (out of a list of 13), the number of limitations in Activities of Daily Living (out of a list of 6 basic, everyday tasks) and self-rated health (ranging from 1=excellent to 5=poor). Furthermore, mental health includes EURO-D depression scale (ranging from 0 to 12 symptoms) and cognitive function based on the score of a memory test (1=excellent to 5=poor), the score of a numeracy test (1=bad to 5=good), the score of self-rated writing skills (1=excellent to 5=poor) and the score of orientation in time test (0=bad to 4=good).

In relation to the second objective, the change in the number of chronic conditions, the change in the number of limitations in Activities of Daily Living and the change in self-rated health are calculated. In particular, the change in self-rated health (SRH) is re-coded in 5 levels (no change in SRH between waves 6 and 7, improvement by 1 level, improvement by 2 or more levels, deterioration by 1 level and deterioration by 2 or more levels). The aim of these variables is to isolate the impact of health improvement or deterioration on updating SSPs.

In relation to the third objective, physical health includes a mobility index (ranging from 0 to 5), which indicates in how many activities the respondent experiences difficulties. The list of activities are: walking one block, walking several blocks, walking across a room, climbing one flight of stairs, and climbing several flights of stairs. Cognitive function is represented by the respondents' scores in two

tests. In the 'serial 7s' test the respondent is asked to subtract 7 from a given number; the score provides the number of correct subtractions, taking values from 0 to 5. Next, the total word recall test score is the sum of the immediate and of the delayed (about 5 minutes) words recalled correctly and ranges from 0 to 20.

Lifestyle & Behavioural risk factors

In relation to the first objective, this set of variables includes the BMI in four categories (underweight, normal weight, overweight and obese), whether the respondent does vigorous or moderate physical activities, whether the respondent ever smoked daily and the frequency of eating meat or chicken, fruits or vegetables, legumes and eggs and dairy products (1 = almost daily to 5 = less than once a week).

In relation to the second objective, the change in BMI category of a respondent between SHARE waves 6 and 7 is calculated. This variable is re-coded in nine levels (no change in BMI group between waves 6 and 7, change from underweight to normal, change from underweight to overweight or obese, change from normal to underweight, change from normal to overweight or obese, change from overweight to obese, change from overweight to normal or underweight, change from obese to overweight and change form obese to normal or underweight). The aim is to investigate how SSPs are updated following a change in BMI category.

In relation to the third objective, the behavioural risk factors include the frequency of vigorous physical activities within a week (1 = every day to 5 = never) and whether the respondent ever smoked daily. BMI is also used in continuous form as a control variable.

Quality of life & Social support

In relation to the first objective, this group of variables includes the quality of life CASP index (covers the domains Control, Autonomy, Self-realisation and Pleasure in life; range: 12 to 48), the life satisfaction score (ranging from 0 to 10) and the number of times the respondent received help from others (ranging from 0 to 3).

In relation to the second objective, the change in the life satisfaction score between SHARE waves 6 and 7 is calculated. The aim of this variable is to investigate how SSPs are updated as life satisfaction varies during the respondents' lifespans.

Finally, the prediction interval in years is also included as explanatory variable.

Table 2.3 Summary of explanatory variables in relation to the objectives of this study.

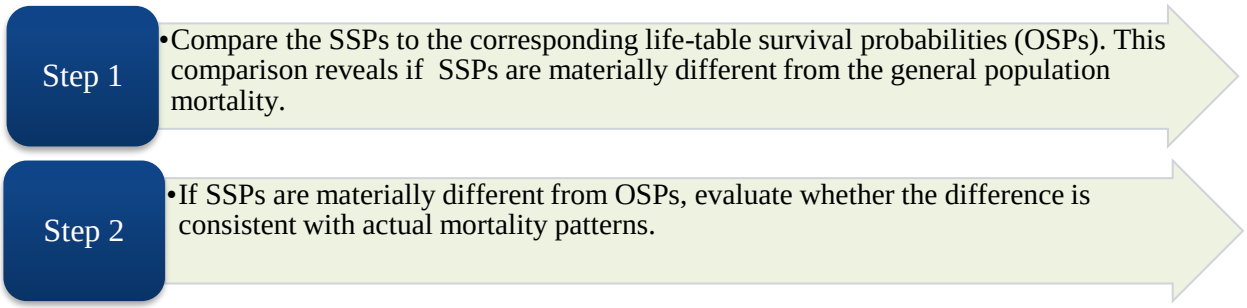
Explanatory variable	Objective 1 : The identification of common factors driving subjective survival probabilities	Objective 2 : How individuals update their subjective survival probabilities	Objective 3 : The quantification of 'Self-perceived age'	Objective 4: Assessing the predictive power of subjective survival metrics on mortality
Demographic characteristics				
Chronological age	Included without adjustments	Included as increase in chronological age	Included as control variable	Included without adjustments
(Chronological age)^2 /100	Not Included	Not Included	Included as control variable	Not Included
Gender	Included without adjustments	Used as control variable	Included as control variable	Included without adjustments
Marital status	Included without adjustments	Included as change in marital status	Included as control variable	Not Included
Number of parents alive	Not Included	Not Included	Included as control variable	Not Included
Number of children	Included without adjustments	Not Included	Not Included	Included without adjustments
Country of residence	Included without adjustments	Used as control variable	Included as control variable	Included as control variable
Socio-economic status				
Total household income	The "equivalised" individual income in quartiles is included	The change in "equivalised" individual income in quartiles is included.	Total household income in quartiles, calculated for each welfare state	The "equivalised" individual income in quartiles is included
Total wealth	Not used	Not used	Total wealth in quartiles, calculated for each welfare state	Not Included
Educational attainment	Included in 4 levels	Included as control variable	Included in 3 levels	Included without adjustments
Physical Health, Mental Health and Cognitive function				
Number of chronic conditions	Included without adjustments	Included as change in number of chronic conditions	Not Included	Not Included
Limitations in Activities of Daily Living	Included without adjustments	Included as change in the number of ADLs	Not Included	Included without adjustments
Self-rated health	Included without adjustments	Included as change in self-rated health	Not Included	Included without adjustments
Grip strength	Included without adjustments	Not Included	Not Included	Not Included
Peak flow	Included without adjustments	Not Included	Not Included	Not Included
EURO-D depression scale	Included without adjustments	Not Included	Not Included	Not Included
Memory test score	Included without adjustments	Not Included	Not Included	Not Included
Numeracy test score	Included without adjustments	Not Included	Not Included	Included without adjustments
Self-rated writing skills	Included without adjustments	Not Included	Not Included	Not Included

Orientation in time score	Included without adjustments	Not Included	Not Included	Included without adjustments
Mobility index score	Not Included	Not Included	Included without adjustments	Not Included
'Serial 7s' test score	Not Included	Not Included	Included without adjustments	Not Included
Total word recall score	Not Included	Not Included	Included without adjustments	Not Included
Lifestyle & Behavioural risk factors				
BMI	Included in 4 groups	Included as change in BMI group	Included as control variable	Included without adjustments
Frequency of physical activities	Included without adjustments	Not Included	Included without adjustments	Included without adjustments
Ever smoked daily	Included without adjustments	Included without adjustments	Included without adjustments	Included without adjustments
Frequency of eating meat or chicken, fruits or vegetables, legumes and eggs and dairy products	Included without adjustments	Not Included	Not Included	Not Included
Quality of life & Social support				
CASP index	Included without adjustments	Not Included	Not Included	Not Included
Life satisfaction score	Included without adjustments	Included as change in Life satisfaction score	Not Included	Included without adjustments
Number of times the respondent received help from others	Included without adjustments	Not Included	Not Included	Not Included
Prediction Interval	Included without adjustments	Not Included	Included as control variable	Not Included

Notes: The table above includes all variables used in this study in relation to each objective. Any adjustments or modifications of the original variables are also briefly outlined.

3. Chapter 3 - Identifying common factors affecting subjective survival probabilities

This first objective of this study is to identify which factors affect in similar manner subjective survival probabilities (SSPs) and whether these factors are also related to actual mortality. In other words, the question I attempt to answer is whether individuals take into account factors affecting actual mortality, in forming survival expectations. To achieve the aforementioned objective, the following two-step process is adopted.



To achieve the objective, data from the SHARE wave 6 were used. The data and statistical modelling are described in the following sections. All data sources are also described in Section 2.5.

3.1.Methods and sample description

In this section the dependent and all explanatory variables are presented in detail.

Dependent variables

The self-reported subjective survival probabilities (SSPs) reflect own survival expectations (Figure 3.1). There is a concentration on rounded values (e.g. 50%). A possible interpretation of this phenomenon is that people respond depending to whether they are confident, not at all confident, or uncertain about own future survival (Hurd and McGarry, 1995). Another interpretation is that it is more straightforward and easier to report a round number, ending on the digit 0 or 5. The distribution of objective survival probabilities (OSPs), calculated from population life tables, is concentrated on values above 60% (Figure 3.2). The difference of SSPs and OSPs is calculated as follows:

$$Survival\ expectation\ difference_{x,N} = SSP_{x,N} - OSP_{x,N}$$

where 'x' is the age of the respondent and 'N' is the prediction interval.

Figure 3.1 Distribution of subjective survival probabilities.

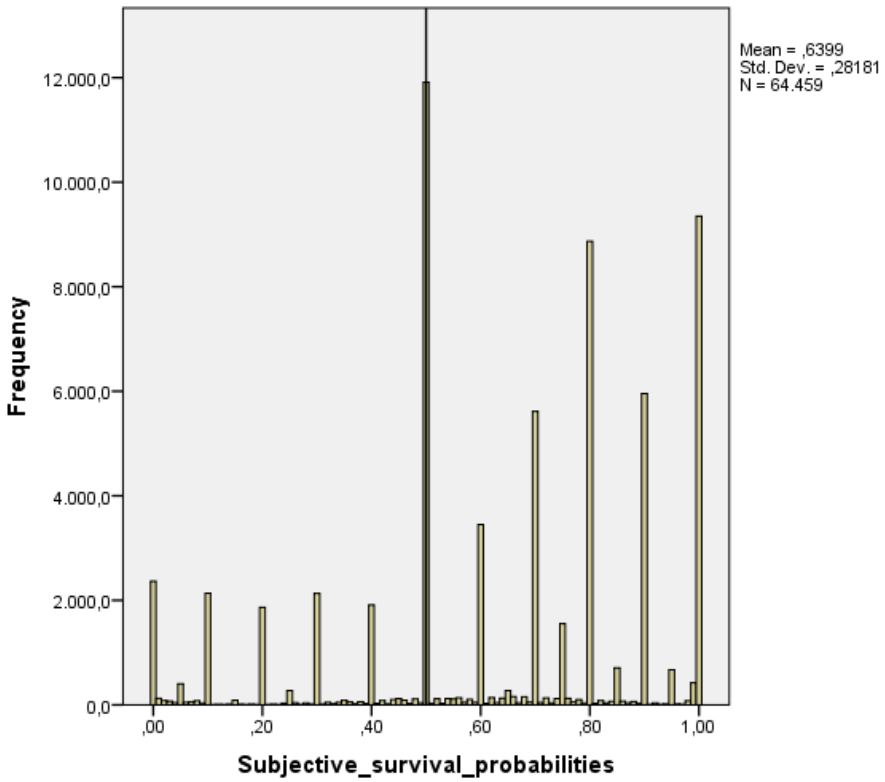


Figure 3.2 Distribution of objective survival probabilities.

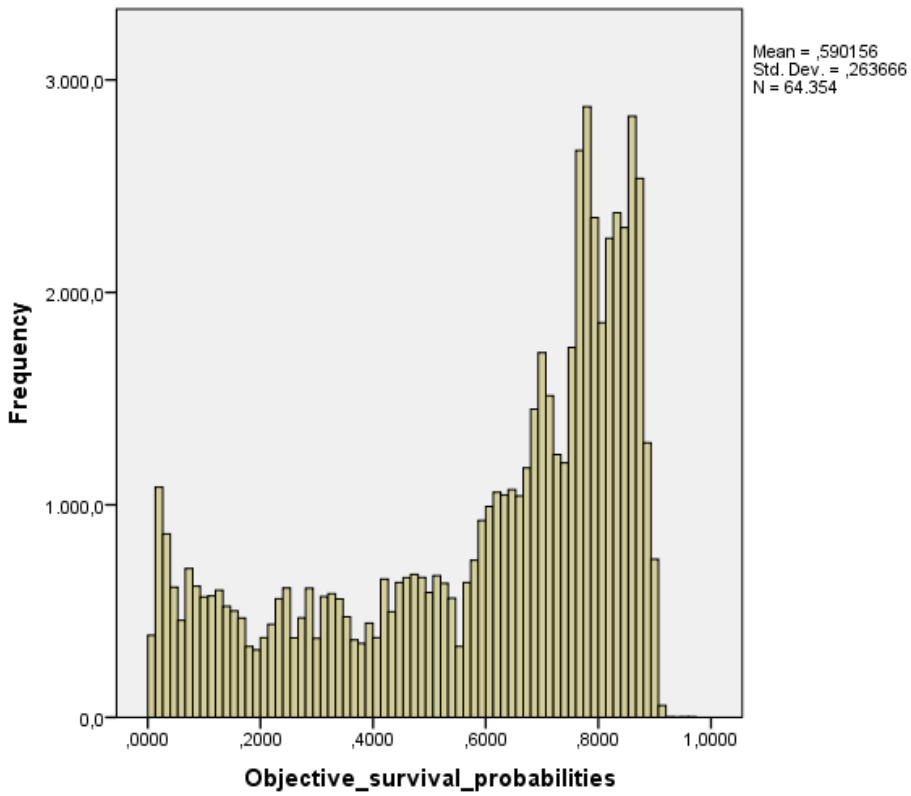
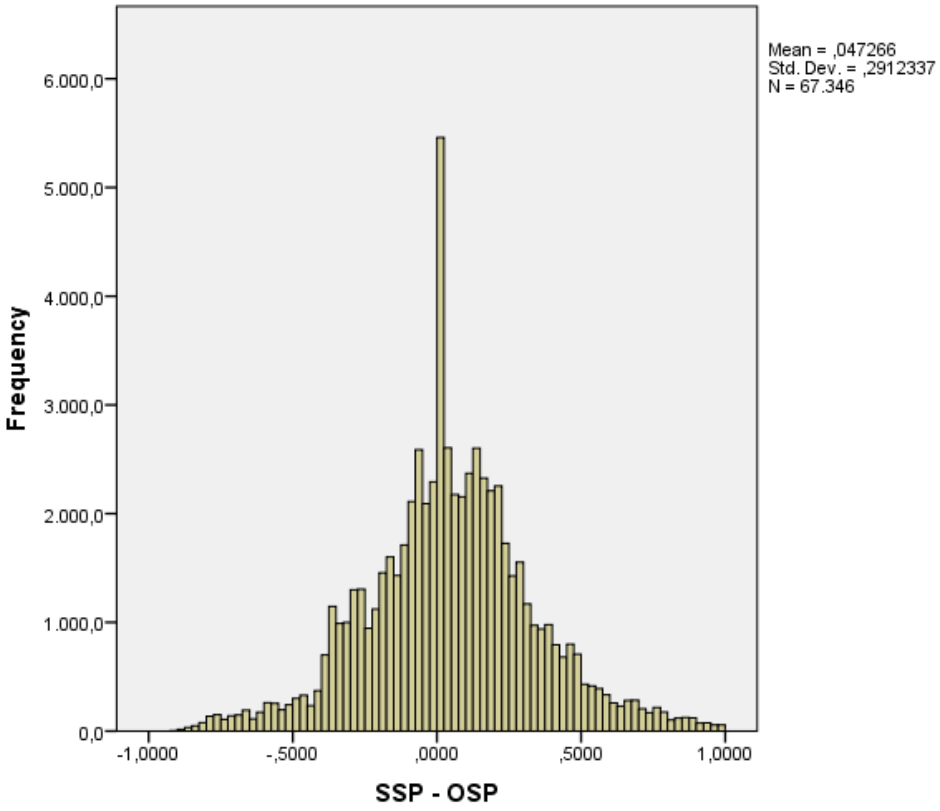


Figure 3.3 Distribution of $Survival\ expectation\ difference_{x,N}$



The distribution of the difference of SSPs and OSPs is bell-shaped (Figure 3.3). However, the high concentration around the center of the distribution implies that the distribution is not similar to a normal distribution (Kolmogorov-Smirnov Z = 9.669, p<1%). Therefore, the difference in survival probabilities is transformed to a categorical variable (Table 2.1).

The categorical variable, 'Survival expectation groups' is defined as follows.

$$Survival\ expectation\ groups_{x,N} = \begin{Bmatrix} 1,\ if\ |Survival\ expectation\ difference_{x,N}| \leq tolerance\ (\%) \\ Else\ if\ \ |Survival\ expectation\ difference_{x,N}| > tolerance\ (\%): \\ 2,\ if\ \ Survival\ expectation\ difference_{x,N} > 0 \\ 3,\ if\ \ \ Survival\ expectation\ difference_{x,N} < 0 \end{Bmatrix}$$

where 'x' is the age of the respondent and 'N' is the prediction interval.

The aim of 'tolerance' is to capture various levels of difference between SSPs and OSPs. The tolerance level is set to 10%, 15% and 20%. The different tolerance levels are used for the sensitivity analysis of regression coefficients as well. The interpretation of the nominal dependent variable is: 1 indicates that that the respondent estimates that his/her future survival will be similar to that of the general population (i.e. the difference lies within the tolerance level); 2 implies that future survival will exceed

while 3 means that future survival will be lower than the general population. The categorical variable clearly differentiates subjective and objective survival expectations, as shown below.

Table 3.1 Categorical variables for the difference of SSPs and OSPs.

Comparison of SSPs to OSPs		Mean of SSPs	Mean of OSPs
SSPs greater than …	20% of OSPs	80%	40%
	15% of OSPs	80%	45%
	10% of OSPs	79%	49%
SSPs close to OSPs …	At 20% tolerance	66%	64%
	At 15% tolerance	65%	64%
	At 10% tolerance	63%	63%
SSPs lower than …	20% of OSPs	35%	73%
	15% of OSPs	39%	72%
	10% of OSPs	40%	70%

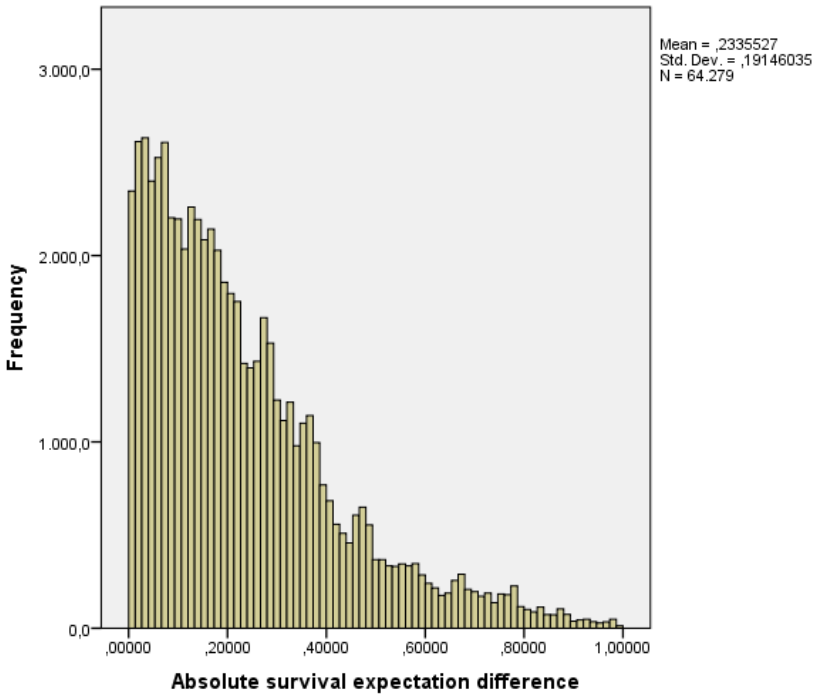
Another variable which will be used to evaluate the variation of SSPs is the following.

$$Absolute\ survival\ expectation\ difference_{x,N} = |SSP_{x,N} - OSP_{x,N}|$$

where 'x' is the age of the respondent and 'N' is the prediction interval.

The 'Absolute survival expectation difference' is in continuous form and provides an alternative estimate of the distance between subjective and objective survival probabilities.

Figure 3.4 Distribution of $Absolute\ survival\ expectation\ difference_{x,N}$.



The distribution of $'Absolute\ survival\ expectation\ difference_{x,N}'$ is right skewed.

Explanatory variables

The explanatory variables are presented in homogeneous groups.

Demographic characteristics

The demographic group includes chronological age (in years), gender, marital status (widowed, divorced, never married, separated, married, in partnership), the number of children of the respondent as well as country of residence (Table 3.2).

Table 3.2 Descriptive statistics of demographic variables.

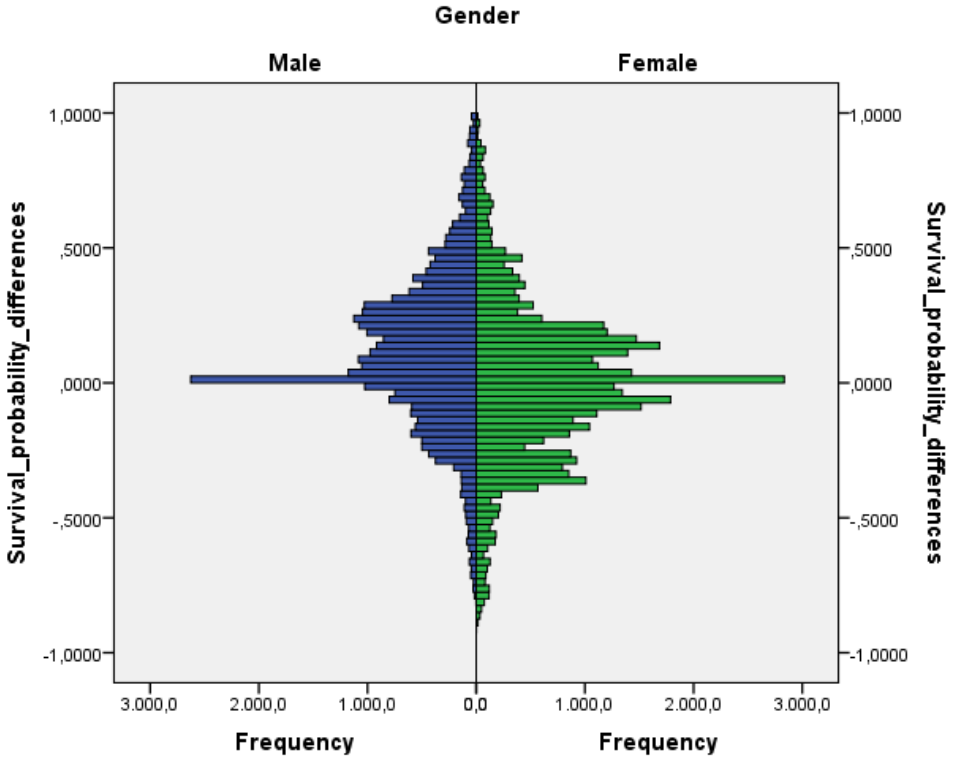
Demographic Characteristics	Descriptive Measures
Age (mean [SD])	68 [10]
Male	44%
Marital Status	
Widowed	15%
Divorced	9%
Never married	6%
Separated	1%
Partnership	2%
Married	68%
Country of residence	
Austria	5%
Belgium	9%
Croatia	4%
Czech Republic	7%
Denmark	6%
Estonia	8%
France	6%
Germany	7%
Greece	7%
Israel	3%
Italy	8%
Luxembourg	2%
Poland	3%
Portugal	2%
Slovenia	6%
Spain	8%
Sweden	6%
Switzerland	4%
Number of children (mean [SD])	2.1 [1.3]

The average respondent is 68 years old; males represent 44% of the SHARE Wave 6 sample and 68% of the respondents are married. Luxemburg represents the lowest proportion whereas Belgium

represents the highest proportion of the sample. The proportion of older individuals aged 90 or higher in SHARE Wave 6 sample is on average below 2% (Table A2 in the Appendix). These figures indicate that for females, there is an under-sampling of older individuals in certain countries (e.g. Italy, Germany).

It is worth highlighting the impact of gender on subjective survival expectations (Figure 3.5). The distribution of the difference in survival probabilities is positively skewed for males but negatively skewed for females. In other words, males are more optimistic and females are rather pessimistic less optimistic about future own survival. Similar results are found in past research (Arpino et al. 2018; Liu et al. 2007; Mirowsky 1999).

Figure 3.5 Impact of gender on the difference of SSPs and OSPs.



Socioeconomic status

Socioeconomic status is represented by the "equivalised" individual income in quartiles and educational attainment, considered in 4 levels, based on the ISCED-97 classification, Primary (code 1), Lower Secondary (code 2), Upper Secondary (codes 3 & 4) and Tertiary (codes 5 & 6). A summary of these measures is presented below.

Table 3.3 Descriptive statistics of socioeconomic variables.

Comparison of SSPs to OSPs		Mean / Proportion in sample	Mean of SSPs	Mean of OSPs
Equivalised Income	Q1	€ 4,700	57%	53%
	Q2	€ 23,932	61%	55%
	Q3	€ 25,464	65%	60%
	Q4	€ 74,060	72%	67%
Education level	ISCED-97 code 0 & 1 (Primary)	22%	57%	50%
	ISCED-97 code 2 (Lower secondary)	18%	61%	59%
	ISCED-97 codes 3 & 4 (Upper secondary)	37%	66%	62%
	ISCED-97 codes 5 & 6 (Tertiary)	23%	69%	63%

The majority of the respondents have completed post-secondary education and the average equivalised income for the first three quartiles increases smoothly. The average value of subjective survival probabilities increases in line with socioeconomic status for the respondents in SHARE Wave 6. Several researchers noted that higher socioeconomic status is associated with higher subjective survival expectations (Arpino et al. 2018; Rappange et al. 2016; Liu et al. 2007; Mirowsky 1999; Balia 2014).

Physical & Mental Health

Physical health includes the number of chronic conditions, the number of limitations in Activities of Daily Living and self-rated health. Mental health is represented by the EURO-D depression scale and cognitive function by the score of a memory test, the score of a numeracy test, the score of self-rated writing skills and the score of orientation in time test. A summary of the descriptive statistics for these variables is presented in Table 3.4.

Table 3.4 Descriptive statistics of physical and mental health variables.

Comparison of SSPs to OSPs			Proportion in sample	Mean of SSPs	Mean of OSPs
Physical health	**Self-rated health**	Excellent	7%	81%	68%
		Very good	18%	76%	67%
		Good	36%	67%	61%
		Fair	29%	57%	53%
		Poor	10%	42%	47%
	Number of chronic diseases	0	22%	74%	68%
		1	28%	67%	62%
		2	21%	62%	57%
		3	14%	57%	53%
		4+	14%	45%	45%
	Number of ADLs	0	88%	66%	61%
		1+	12%	43%	40%
Mental health	**EURO-D depression scale** *(ranging from 0 to 12 symptoms)*	0	21%	72%	63%
		1	20%	69%	61%
		2	17%	65%	59%
		3	14%	62%	57%
		4+	28%	43%	55%
	Score of memory test *(1=excellent to 5=poor)*	1	7%	73%	66%
		2	21%	71%	66%
		3	45%	64%	60%
		4	23%	57%	51%
		5	5%	49%	44%
	Score of numeracy test *(1=bad to 5=good)*	1	4%	53%	46%
		2	15%	59%	54%
		3	33%	63%	59%
		4	33%	65%	61%
		5	16%	69%	63%
	Score of self-rated writing skills *(1=excellent to 5=poor)*	1	27%	69%	64%
		2	26%	66%	62%
		3	30%	61%	57%
		4	12%	57%	51%
		5	5%	51%	44%
	Score of orientation in time test *(0=bad to 4=good)*	0 - 3	12%	48%	34%
		4	88%	65%	61%

Approximately six out of ten respondents rate their health as fair or good. The average respondent reports 1.8 chronic diseases, 0.3 ADL limitations and 2.4 depressive symptoms. Poor physical and mental health is associated with lower subjective survival probabilities. This association is in line with the results of past research (Liu et al. 2007; Balia 2014; Hurd and McGarry 1995; Rappange et al. 2016) and it is further supported by the correlation coefficients presented in Table 3.5.

Table 3.5 Descriptive statistics of health-related independent variables and correlation coefficients with subjective survival probabilities.

Health-related variables	Mean [St.dev.]	Subjective survival probabilities	
		Pearson	Spearman's rho
Number of chronic diseases	1.8 [1.6]	-2.8%**	-2%**
Number of ADLs	0.3 [0.9]	-1.95%**	-1.9%**
Depression scale	2.4 [2.2]	-27%**	-24%**
Memory score	3 [0.95]	-21.5%**	-21%**
Numeracy score	3.4 [1.03]	12.6%**	11.6%**
Self-rated writing skills score	2.4 [1.14]	-16.9%**	-6.3%**
Orientation in time score	3.8 [0.51]	13%**	11.1%**

** p<5% , ** p<1%*

Lifestyle & Behavioural risk factors

This group of variables includes the BMI in four categories, whether the respondent does vigorous or moderate physical activities, whether the respondent ever smoked daily and the frequency of eating meat or chicken, fruits or vegetables, legumes and eggs and dairy products. A summary of the descriptive statistics for these variables is presented in Table 3.6.

Table 3.6 Descriptive statistics of lifestyle-related variables.

Comparison of SSPs to OSPs		Proportion in sample	Mean of SSPs	Mean of OSPs
BMI	Underweight	1%	58%	55%
	Normal	30%	65%	60%
	Overweight	45%	64%	58%
	Obese	24%	62%	60%
Physical activity	Do physical vigorous or moderate physical activity	88%	66%	61%
	Never do vigorous nor moderate physical activity	12%	48%	43%
Smoking status	Never smoked daily	55%	63%	57%
	Smoked daily	45%	65%	61%
Frequency of eating meat or chicken	Every day	36%	65%	59%
	3-6 times a week	43%	64%	59%
	Twice a week	15%	62%	57%
	Once a week	4%	61%	59%
	Less than once a week	2%	63%	63%
Frequency of eating fruits or vegetables	Every day	77%	65%	59%
	3-6 times a week	17%	63%	58%
	Twice a week	4%	60%	58%
	Once a week	1%	56%	57%
	Less than once a week	1%	57%	59%
Frequency of eating legumes and eggs	Every day	10%	67%	59%
	3-6 times a week	27%	66%	60%
	Twice a week	31%	64%	59%
	Once a week	23%	62%	59%
	Less than once a week	10%	60%	58%
Frequency of eating dairy products	Every day	65%	65%	59%
	3-6 times a week	20%	63%	59%
	Twice a week	8%	63%	60%

	Once a week	3%	62%	60%
	Less than once a week	4%	62%	59%

Seven out of ten respondents are overweight or obese and nine out of ten do some sort of physical activity. Furthermore, the majority of individuals consume fruits or vegetables and dairy products on a daily basis. However, the frequency of consumption is lower for meat or chicken and eggs or legumes. In particular, respondents eat meat or chicken 3 – 6 times per week and eggs or legumes twice a week. Physical activity and normal BMI are linked to higher SSPs (Figure 3.6). On the other hand, underweight and obese tend to report lower SSPs. The impact of diet habits on SSPs will be further explored using correlation coefficients (Table 2.7). In particular, less frequent consumption is associated with lower SSPs.

Figure 3.6 Impact of BMI group on subjective survival probabilities

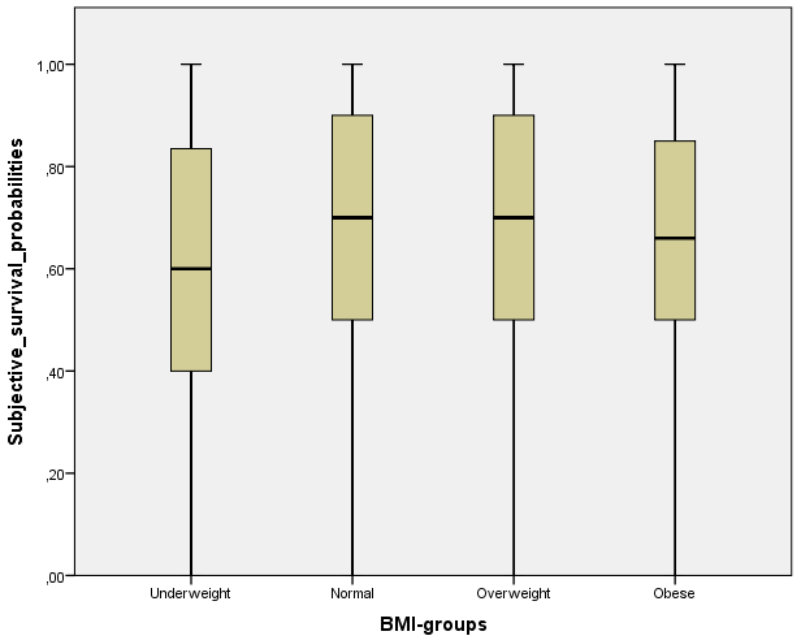


Table 3.7 Pearson and Spearman correlation coefficients of the subjective survival probabilities with dietary habits (frequency of eating different products).

Dietary habits (frequency)	Subjective survival probabilities	
	Pearson	Spearman's rho
Dairy products	-3.3%**	-3.7%**
Meat or chicken	-6.8%**	-4.4%**
Eggs or legumes	-4.1%**	-7.1%**
Fruits or vegetables	-6%**	-5.3%**

** p<5% , ** p<1%*

Note: Frequency ranges from 1 (every day) to 5 (less than once a week).

Quality of life & Social support

This group of variables includes the quality of life CASP index, the life satisfaction score and the number of times the respondent received help from others. As noted below, high life satisfaction score, high social support and good quality of life are associated with higher SSPs (Figure 3.7 and Table 3.8). Similar results are presented by several researchers (Van Solinge and Henkens 2018, Ross and Mirowsky 2002).

Figure 3.7 Impact of life satisfaction and social support on subjective survival probabilities.

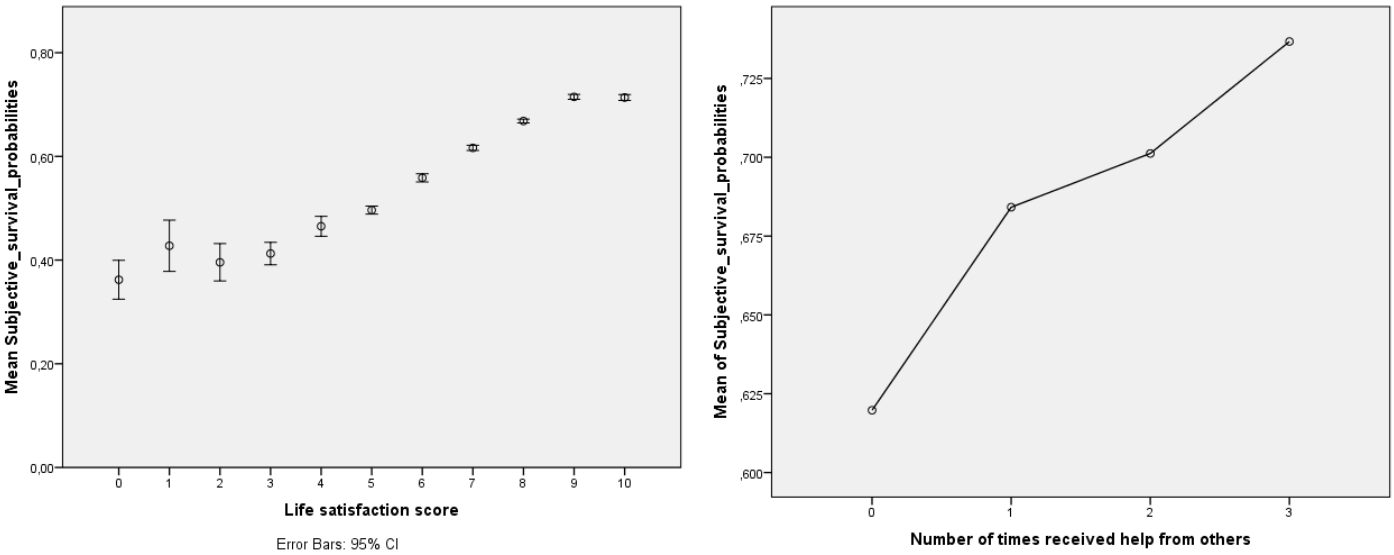


Table 3.8 Descriptive statistics of variables related to quality of life and social support and their correlation coefficients with subjective survival probabilities.

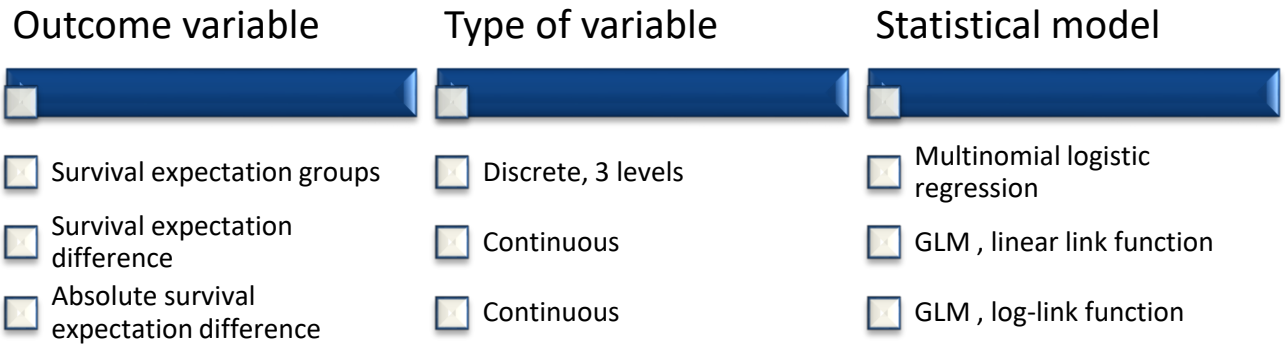
Variables related to diet habits	Mean [St.dev.]	Subjective survival probabilities	
		Pearson	Spearman's rho
Quality of life CASP index	34 [11.1]	29.3%**	34.7%**
Life satisfaction score	7.6 [1.7]	27.7%**	26.4%**
Number of times received help	0.37 [0.7]	11.5%**	11.5%**

** p<5% , ** p<1%*

In the next section I discuss the statistical modelling.

3.2.Statistical modelling

In the analysis multinomial and generalised linear models (GLMs) are used, dependent upon the type of the outcome variable. SPSS 21 is used for the analyses.



The ‘Survival expectation difference’ variable measures the difference of subjective and objective survival probabilities, whereas the ‘Absolute survival expectation difference’ variable reflects the absolute distance of the two probabilities. The ‘Survival expectation groups’ is a discrete variable which aims to categorize the difference between SSPs and OSPs in three groups. A tolerance level is used to allocate individuals to groups depending on the value of the difference of subjective and objective survival probabilities in comparison to the value of the tolerance level. The tolerance level is set to 10%, 15% and 20% which means that three separate multinomial models are run.

The value of regression coefficients in conjunction with the actual mortality patterns will be used to allocate the explanatory variables in three groups, namely, better chances of survival, worse chances of survival and diverging survival patterns. An additional multinomial model including only sociodemographic variables is run to investigate the sensitivity of model estimates. The results of the analyses are summarised in the next section.

3.3.Results

The results of the statistical models are presented in tables 3.9 – 3.11. In particular,

- The Relative Risk Ratios for the 'socio-demographic' and the 'full' multinomial models are presented in Table 3.9. The dependent variable is 'Survival expectation groups' using a tolerance level of 10%.
- The Relative Risk Ratios for two additional 'full' multinomial models are presented in Table 3.10. The dependent variable is 'Survival expectation groups' using a tolerance level of 15% and 20%, respectively.
- The coefficients based on 'full' generalised linear and log-linear models are presented in Table 3.11. The dependent variables are 'Survival expectation difference' and 'Absolute survival expectation difference'.

The results are discussed in detail following presentation of the tables.

Table 3.9 Relative Risk Ratios (RRRs) based on Multinomial Logistic regression. Tolerance level is 10%.

Independent variables	Higher vs Close to population survival predictions [a]		Lower vs Close to population survival predictions [a]	
	Socio-demographic model Tolerance 10%	Full model Tolerance 10%	Demographic & socioeconomic model Tolerance 10%	Full model Tolerance 10%
Prediction Interval	1.098**	1.116**	0.940**	0.917**
Demographic Characteristics				
Age	1.072**	1.088**	0.921**	0.917**
Gender (reference: Female)				
Male (reference: Female)	1.754**	1.793**	0.669**	0.680**
Country of residence (reference: Austria)				
France	0.573**	0.639**	1.563**	1.363**
Switzerland	0.709**	0.692**	1.144	1.290**
Sweden	0.738**	0.685**	1.007	1.157*
Belgium	0.593**	0.627**	1.304**	1.142*
Germany	0.760**	0.867*	1.321**	1.147*
Czech Republic	0.759**	0.895	1.212**	1.156*
Luxembourg	1.085	1.126	1.476**	1.243*
Israel	0.721**	0.852*	0.856	0.780**
Spain	0.765**	0.829**	0.896	0.882
Greece	0.688**	0.832**	1.104	1.032
Croatia	0.969	1.015	1.289**	1.162
Poland	0.895	1.088	1.384**	1.087
Italy	0.962	1.096	1.174*	1.037
Slovenia	1.181**	1.331**	1.395**	1.210*
Portugal	0.949	1.206*	1.297*	0.872
Denmark	1.822**	1.656**	0.790**	0.839**
Estonia	1.144**	1.572**	1.074	0.785**
Marital status (reference: Married)				

Widowed	1.017	1.072*	1.058	1.015
Divorced	1.037	1.115**	1.011	0.926
Never married	1.014	1.112*	1.038	0.951
Separated	0.966	1.022	1.073	0.9699
Partnership	1.054	1.096	1.015	0.978
Number of children	1.034**	1.029**	1.007	0.997
Socio-economic Status				
Education (reference: Tertiary education)				
Primary	1.048	1.204**	1.553**	1.098*
Lower Secondary	1.014	1.122**	1.446**	1.119**
Upper secondary	1.076**	1.138**	1.301**	1.132**
Equivalised Income quartiles (reference: Q4)				
Equivalised Income Q1	0.929	1.046	1.557**	1.165**
Equivalised Income Q2	1.021	1.086	1.491**	1.220**
Equivalised Income Q3	0.999	1.026	1.223**	1.098**
Physical & Mental Health				
Chronic diseases		0.994		1.064**
ADLs		0.970		0.996
Self-rated health (reference: Poor)				
Excellent		2.467**		0.500**
Very good		1.705**		0.491**
Good		1.458**		0.672**
Fair		1.324**		0.859**
Depression		1.002		1.063**
Numeracy		0.965**		0.982
Writing skills		0.984		1.015
Orientation		0.985		1.041
Memory		0.950**		1.027
Lifestyle & Behavioural Risk Factors				
BMI (reference: Obese)				
Underweight		0.858		0.925
Normal		0.883**		0.966
Overweight		0.956		0.973
Physical activity (reference: Physically Inactive)				
Physically Active		1.050		0.943
Ever smoked daily (reference: Yes)				
Never smoked daily		1.029		0.943**
Dietary habits				
Meat or chicken		0.992		0.980
Dairy		1.031**		1.028*
Egg or legumes		0.961**		1.009
Fruit or vegetables		0.970*		1.036*
Quality of life and Social Support				
Quality of Life		1.018**		0.990**
Times received help		1.029		0.935**
Life satisfaction		1.106**		0.912**
Pseudo R^2	0.168	0.278		

[a] The dependent variable is an unordered categorical variable reflecting the deviation of subjective survival probabilities compared to objective survival probabilities. The tolerance level is 10%. * p<5%, ** p<1%. The value of RRRs in

conjunction with the actual mortality patterns will be used to allocate the explanatory variables in the groups: better chances of survival; worse chances of survival; and diverging survival patterns.

Table 3.10 Relative Risk Ratios (RRRs) based on Multinomial Logistic regression. Tolerance level is 15% and 20%.

Independent variables	Higher vs Close to population survival predictions[a]		Lower vs Close to population survival predictions [a]	
	Tolerance 15%	Tolerance 20%	Tolerance 15%	Tolerance 20%
Prediction Interval	1.174**	1.208**	0.910**	0.898**
Demographic characteristics				
Age	1.128**	1.154**	0.906**	0.894**
Gender (reference: Female**)**				
Male	2.143**	2.592**	0.716**	0.644**
Country of residence (reference: Austria)				
France	0.561**	0.693**	1.661**	1.635**
Switzerland	0.553**	0.669**	1.554**	1.609**
Sweden	0.565**	0.656**	1.308**	1.276*
Belgium	0.661**	0.716**	1.246**	1.266**
Germany	0.956	0.964	1.260**	1.278**
Czech Republic	1.025	1.177*	1.378**	1.211*
Luxembourg	1.223	1.144	1.317**	1.444**
Israel	0.798**	1.007	0.986	0.966
Spain	0.764**	1.000	0.912	0.787**
Greece	0.855*	0.975	1.189*	1.076
Croatia	1.106	1.428**	1.304*	1.277**
Poland	1.407**	1.569**	1.312*	1.274*
Italy	0.926	1.298**	1.119	1.051
Slovenia	1.276**	1.236**	1.291**	1.197*
Portugal	1.137	1.492**	0.920	0.943
Denmark	2.059**	1.970**	0.954	0.840
Estonia	2.222**	2.513**	0.921	0.893
Marital status (reference: Married)				
Widowed	1.068	1.089*	1.032	1.043
Divorced	1.201**	1.269**	0.930	0.971
Never married	1.114*	1.119*	0.932	0.948
Separated	0.985	1.075	1.009	0.985
Partnership	1.049	0.992	0.981	0.9686
Number of children	1.022*	1.023*	0.989	1.003
Socio-economic status				
Education (reference: Tertiary education)				
Primary	1.300**	1.292**	1.138**	1.126**
Lower Secondary	1.155**	1.157**	1.131**	1.113**
Upper secondary	1.126**	1.121**	1.106**	1.073*
Equivalised Income quartiles (reference: Q4)				
Equivalised Income Q1	1.083	1.080	1.178**	1.196**
Equivalised Income Q2	1.113**	1.012**	1.202**	1.205**
Equivalised Income Q3	1.050	1.060	1.088*	1.098**
Physical & Mental Health				
Chronic diseases	0.991	0.991	1.064**	1.069**
ADLs	0.949**	0.978	0.996	1.007
Self-rated health (reference: Poor)				
Excellent	2.602**	2.883**	0.447**	0.407**
Very good	1.930**	2.158**	0.475**	0.426**
Good	1.591**	1.686**	0.660**	0.615**

Fair	1.375**	1.377**	0.848**	0.810**
Depression	1.002	0.997	1.064**	1.068**
Numeracy	0.943**	0.934**	0.971*	0.969*
Writing skills	0.983	0.983	1.005	1.000
Orientation	0.963	0.947*	1.037	0.991
Memory	0.935**	0.932**	1.037**	1.041**
Lifestyle & Behavioural Risk Factors				
BMI (reference: Obese)				
Underweight	0.825*	0.749	0.913	0.864
Normal	0.862**	0.855**	0.950	0.949
Overweight	0.956	0.958	0.992	0.991
Physical activity (reference: Physically Inactive)				
Physically Active	1.068	1.009*	0.972	0.910*
Ever smoked daily (reference: Yes)				
Never smoked daily	1.016	1.023	0.906**	0.895**
Dietary habits				
Meat or chicken	1.009	1.006	1.003	1.009
Dairy	1.024*	1.022	1.028*	1.023
Egg or legumes	0.950**	0.958**	1.015	1.015
Fruit or vegetables	0.959*	0.974	1.028	1.047**
Quality of life and Social Support				
Quality of Life	1.022**	1.022**	0.987**	0.986**
Times received help	1.003	1.003	0.913**	0.921**
Life satisfaction	1.108**	1.113**	0.899**	0.891**
Pseudo R^2	0.290	0.309		

[a] The dependent variable is an unordered categorical variable reflecting the deviation of subjective survival probabilities compared to objective survival probabilities. The tolerance level is 15% and 20%. * p<5%, ** p<1%. The value of RRRs in conjunction with the actual mortality patterns will be used to allocate the explanatory variables in the groups: better chances of survival; worse chances of survival; and diverging survival patterns.

Table 3.11 Coefficients based on generalised linear and log-linear models. The dependent variable is the difference between subjective and objective survival probabilities.

Independent variables	'Survival expectation difference' linear model	'Absolute survival expectation difference' log-linear model
Constant	-2.034**	-4.332**
Prediction Interval	0.025**	0.037**
Demographic characteristics		
Age	0.021**	0.034**
Gender (reference: Female)		
Male	0.112**	0.115**
Country of residence (reference: Austria)		
France	-0.067**	-0.026
Switzerland	-0.060**	-0.003
Sweden	-0.055**	-0.110**
Belgium	-0.038**	-0.020
Germany	-0.014*	0.007
Czech Republic	-0.009	-0.017**
Luxembourg	-0.003	0.132**
Israel	0.005	-0.065*
Spain	0.001	-0.044
Greece	0.006	-0.010**
Croatia	0.009	0.068
Poland	0.015*	0.013
Italy	0.013*	0.115**
Slovenia	0.015*	0.045*
Portugal	0.047**	0.112**
Denmark	0.061**	0.123**
Estonia	0.096**	0.099**
Marital status (reference: Married)		
Widowed	0.004	-0.029**
Divorced	0.020**	0.033**
Never married	0.013**	0.007
Separated	-0.001	0.028
Married/Partnership	0.004	0.004
Number of children	0.002*	0.010**
Socio-economic status		
Education (reference: Tertiary education)		
Primary	0.011**	0.081**
Lower Secondary	0.001	0.064**
Upper secondary	0.003	0.059**
Income quartiles (reference: Q4)		
Equivalised Income Q1	-0.013**	0.112**
Equivalised Income Q2	-0.012**	0.109**
Equivalised Income Q3	-0.007	0.067**
Physical & Mental Health		
Chronic diseases	-0.007**	0.005*
ADLs	-0.001	-0.012**
Self-rated health (reference: Poor)		
Excellent	0.153**	0.123**
Very good	0.129**	0.053**
Good	0.095**	0.041**
Fair	0.057**	0.028*
Depression	-0.007**	0.014**

Numeracy	-0.003**	-0.022**
Writing skills	-0.002*	-0.008*
Orientation	-0.006**	-0.005
Memory	-0.007**	-0.008*
Lifestyle & Behavioural Risk Factors		
BMI (reference: Obese)		
Underweight	-0.01	-0.043
Normal	-0.007*	-0.024**
Overweight	-0.002	-0.031**
Physical activity (reference: Physically Inactive)		
Physically Active	0.017**	0.015
Ever smoked daily (reference: Yes)		
Never smoked daily	0.007**	-0.034**
Dietary habits		
Meat or chicken	0.000	0.004
Dairy	0.000	0.013**
Egg or legumes	-0.005**	-0.017**
Fruit or vegetables	-0.007**	0.011**
Quality of life and Social Support		
Quality of Life	0.003**	0.003**
Times received help	0.008**	-0.001
Life satisfaction	0.020**	0.009**

* p<5%, ** p<1%. The sign and value of regression coefficients in conjunction with the actual mortality patterns will be used to allocate the explanatory variables in the groups: better chances of survival; worse chances of survival; and diverging survival patterns. The sings of the "Absolute survival expectation difference" model coefficients show whether predictions are close to general population. Positive coefficients indicate divergence whereas negative coefficients indicate convergence of subjective and objective survival probabilities.

Demographic characteristics

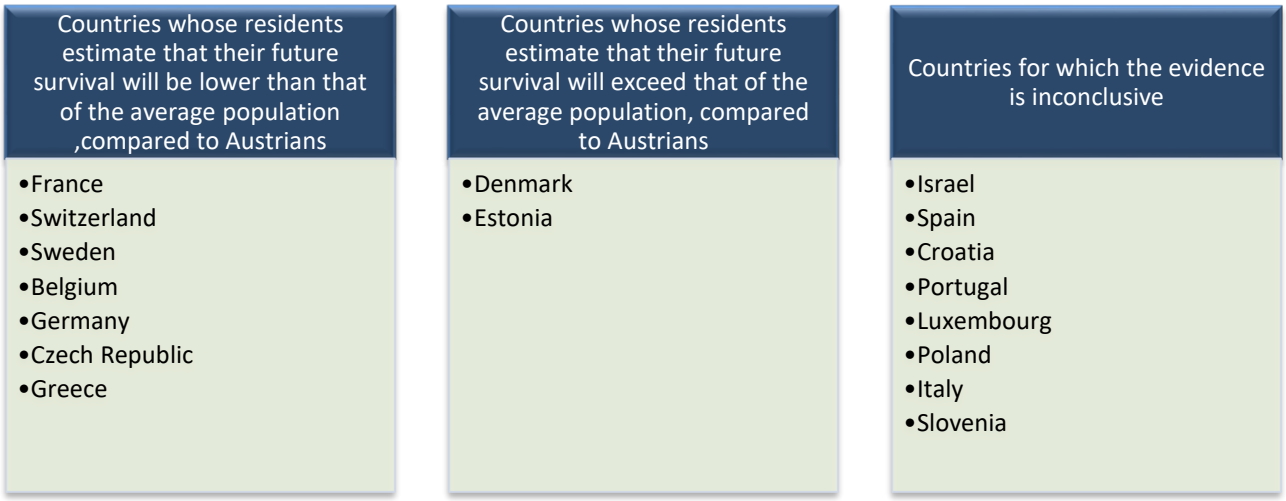
Older individuals tend to report higher survival probabilities. All multinomial regression models (Tables 3.9 and 3.10) show that older individuals tend to estimate higher survival probabilities compared to the general population. This is confirmed by the results of two generalised linear models (Table 3.11). The regression coefficients are positive for both models which imply that older persons report higher survival probabilities compared to the objective survival probabilities.

Males report higher survival probabilities. The Relative Risk Ratios (RRRs) of the Socio-demographic model show clearly that males predict that their future survival will be higher than the general population, compared to females (RRR = 1.754). This is further confirmed by the results of two generalised linear models. The coefficients are positive for both models which imply that males report higher survival probabilities compared to females (b=0.112).

Country of residence is another differentiating factor of SSPs. Countries are grouped into three buckets, based on whether the RRRs of the Socio-demographic model are above or below unity or they do not differentiate compared to Austria (Figure 3.8). More specifically, residents of North-Central European countries, Czechs and Greeks tend to estimate lower future survival than that of the

average population compared to the Austrians. It is worth noting that the significant variation of household cooperation rates by country, from 54% in France to 88% in Croatia (Bergmann et al., 2017). This variation may have affected the representation of high, middle and low socioeconomic classes and therefore impact our estimations.

Figure 3.8 Classification of countries in groups compared to Austria.



Note: Austria is used as reference category. The purpose of this figure is to present the variation of SSPs across countries in Europe.

Widowed, divorced and never married persons tend to estimate that their future survival will exceed that of the general population compared to married. For divorced and never married the findings are confirmed by the results of two generalised linear models. The regression coefficients are positive for both models which imply that divorced and never married report higher survival probabilities compared to married. However, for widowed persons the coefficient of the 'Absolute survival expectation difference' log-linear model (b=-0.029) implies that they tend to report SSPs close to general population OSPs in comparison to married respondents.

The number of respondent's children is positively correlated with SSPs. Our multinomial regression results show that individuals with more children are optimistic about their future survival. This is further confirmed by the results of the two generalised linear models. The coefficients show that individuals with more children are optimistic (b=0.002) and their SSPs diverge from the general population OSPs (b=0.010).

Socioeconomic status

Respondents who earn more income report higher SSPs. Our regression results suggest that low and medium income earners estimate lower future survival than the general population, compared to high

income earners. This is further confirmed by the results of two generalised linear models. The coefficients show that low and medium income earners report low SSPs (b=-0.013) and their forecasts diverge from the general population OSPs (b=0.112). The results of the Socio-demographic model indicate that those with fewer qualifications tend to estimate that their future survival is either lower or higher than that of the general population, compared to those who have completed tertiary education. This implies that individuals in the latter category predict survival closely to the average population.

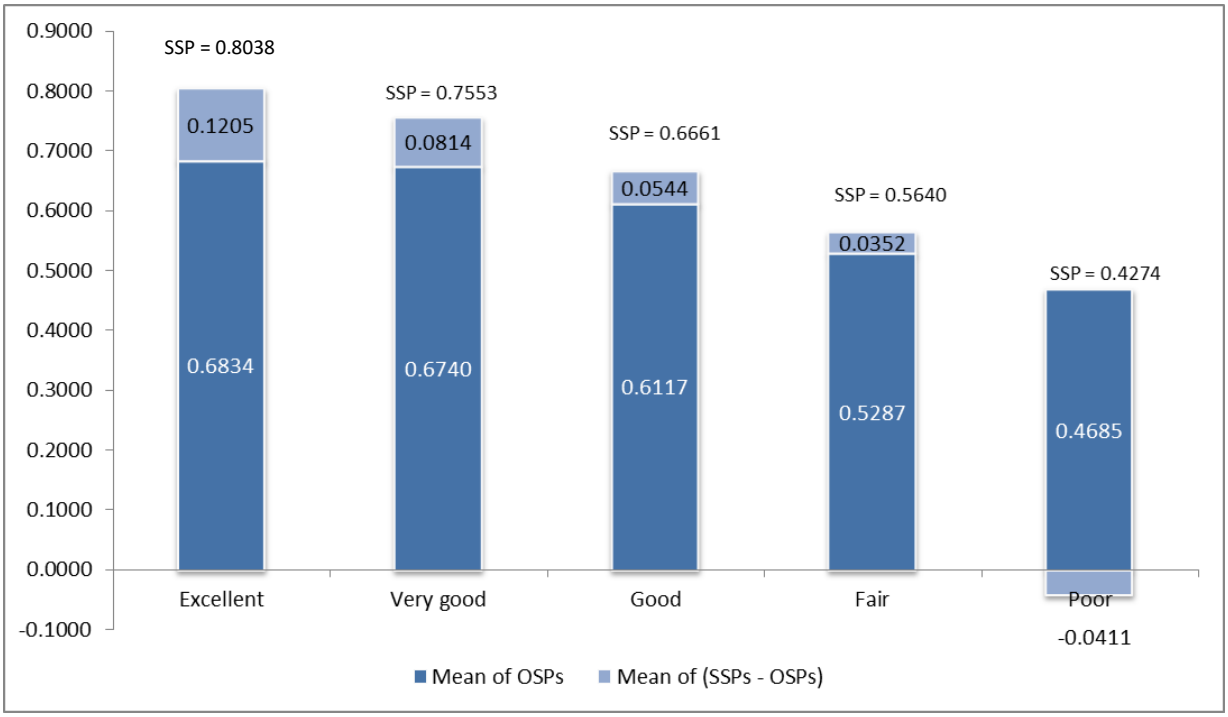
The results of the generalised linear models are similar. For example individuals who completed primary education tend to report high SSPs (b=0.011) and their forecasts diverge from the general population OSPs (b=0.081). In other words, low and medium educational attainment is linked to overly optimistic survival forecasts, whereas high educational attainment is linked to the alignment of survival forecasts with general population mortality experience.

It should be noted that the incorporation of physical and mental health, lifestyle, behavioural risk factors and quality of life in the multinomial regression 'full' model , using the same tolerance level 10%, do not materially differentiate the findings of the socio-demographic model.

Physical and Mental Health

Poor physical health is associated with lower SSPs. In particular, the RRRs of all 'Full' multinomial models show that an additional chronic disease increases chances of estimating that future survival will be lower than that of the general population by 6.4% to 6.9%. Moreover, the coefficient of 'Survival expectation difference' linear model is negative (b=-0.007) which confirms the association. Respondents who report excellent self-rated health also report the highest SSPs (Figure 3.9). The RRRs of all 'Full' multinomial models show that those with excellent, very good, good and fair health estimate that their future survival will exceed that of the general population, compared to those in poor health. The findings are confirmed by the coefficients of the 'Survival expectation difference' linear model.

Figure 3.9 Decomposition of the mean of subjective survival probabilities across self-rated health status.



Note: The average subjective survival probability for each SRH status consists of the average objective survival probability plus the average difference of SSPs and OSPs.

The number of ADLs is negatively correlated with SSPs. The multinomial regression results show that individuals with more ADLs tend to estimate a lower future survival than that of the general population. For example, the coefficient of 'Survival expectation difference' linear model is negative (b=-0.001), which implies that more ADLs are linked to lower SSPs. Depression is negatively associated with SSPs (Figure 3.10 and Table 3.5). All 'Full' multinomial models indicate that an additional symptom of depression increases chances of estimating that future survival will be lower than that of the general population by 6.3% to 6.8%. Similar results are estimated from the 'Survival expectation difference' linear model (b=-0.007). The score of memory test is negatively associated with SSPs (Figure 3.11 and Table 3.5). The regression results show that the poorer a respondent's memory is, the higher are the chances of reporting low SSPs. These results are confirmed by the 'Survival expectation difference' linear model (b=-0.007). Furthermore, the coefficient of the 'Absolute survival expectation difference' log-linear model indicates that respondents with poor memory report SSPs closer to the general population OSPs (b=-0.008).

Figure 3.10 Variation of the mean of subjective survival probabilities with respect to the number of depression symptoms.

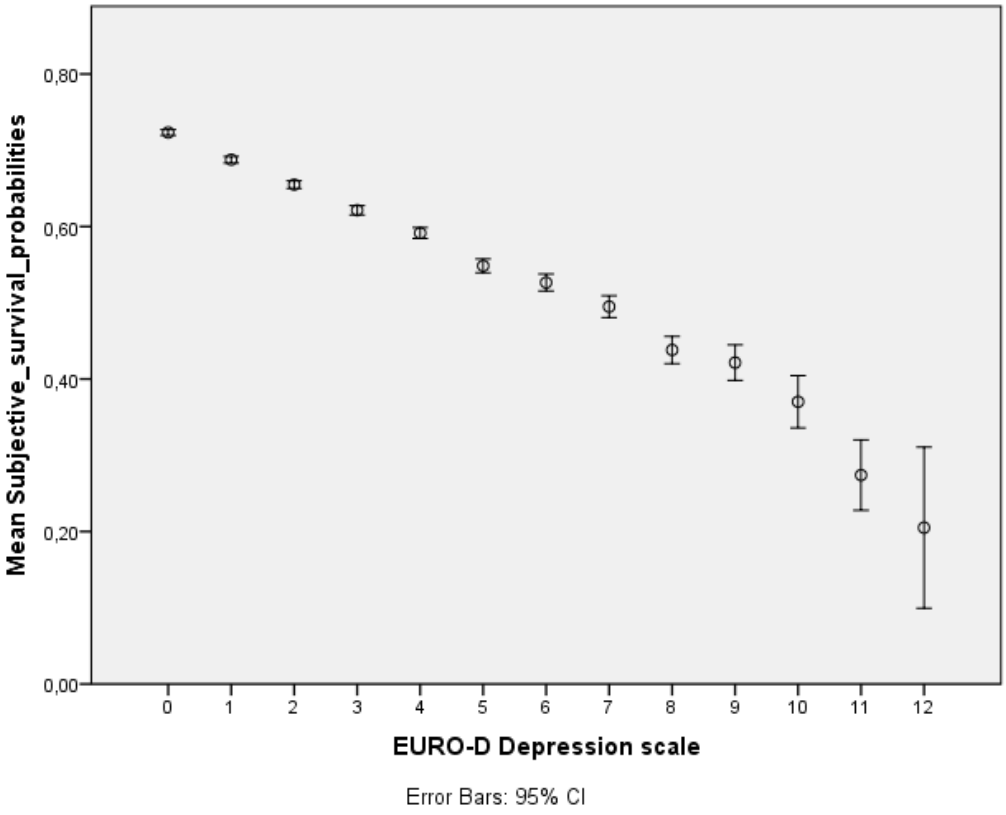
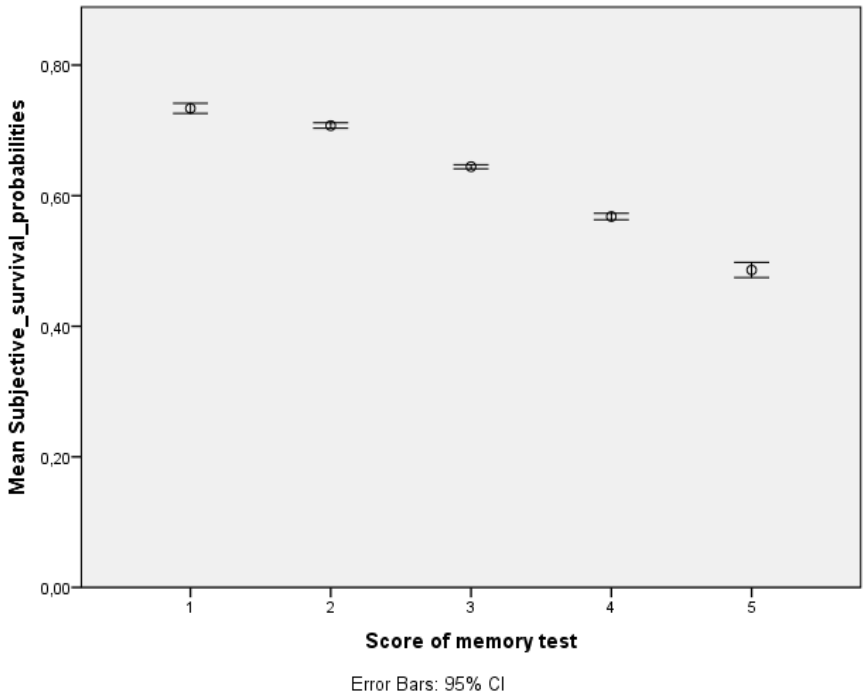


Figure 3.11 Variation of the mean of subjective survival probabilities with respect to the score of memory test.



The score of numeracy test and the self-reported writing skills are positively associated with SSPs (Figure 3.12 and Table 3.5). The RRRs of the multinomial models show that respondents with better numeracy skills have lower chances to report SSPs either lower or higher than the general population. In other words, better numeracy skills are linked to survival predictions closely to the general population average.

Figure 3.12 Variation of the mean of subjective survival probabilities with respect to the score of numeracy test.

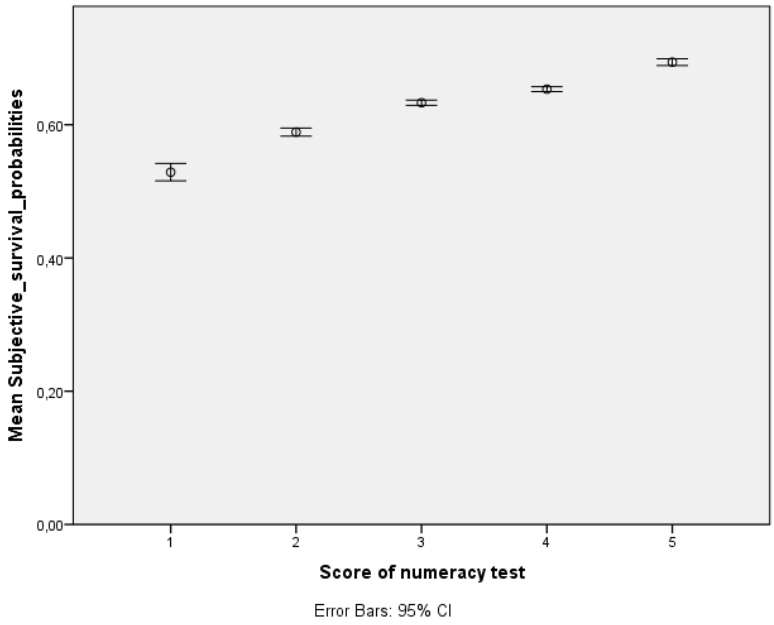
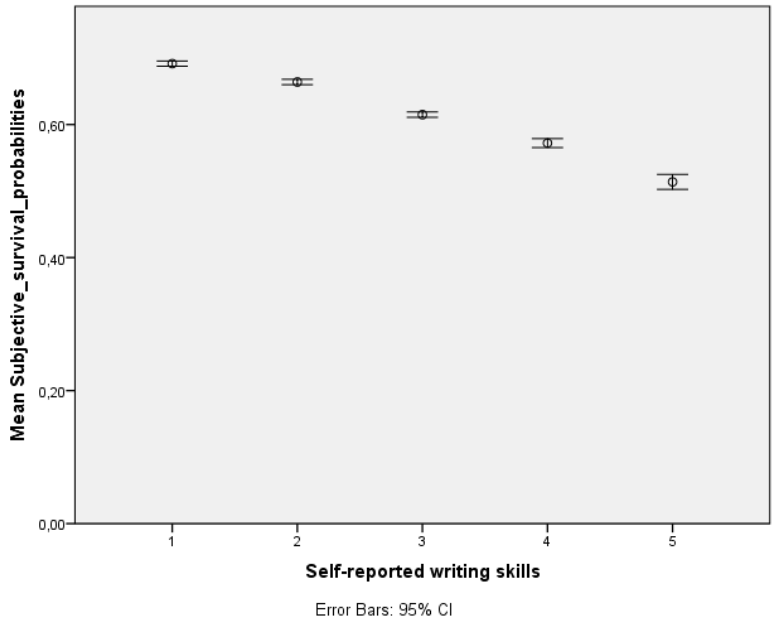


Figure 3.13 Variation of the mean of subjective survival probabilities with respect to the score self-reported writing skills.

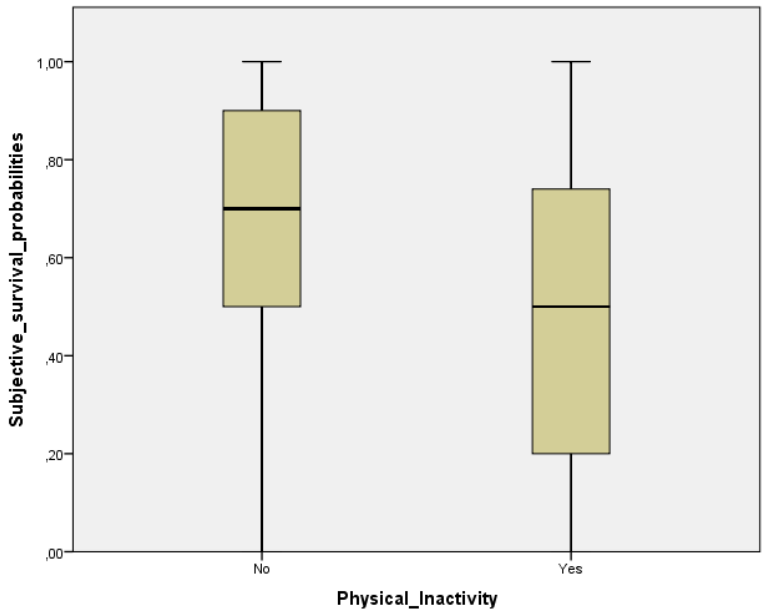


Better writing skills are related to a higher future survival compared to the general population (b=-0.002). Furthermore, the coefficient of the 'Absolute survival expectation difference' log-linear model indicates that respondents with better numeracy and better writing skills report SSPs closer to the general population OSPs (b=-0.022 and b=-0.008). Orientation in time is positively associated with SSPs (Table 3.5). The multinomial regression results show that respondents who have better orientation in time score are more likely to report a lower SSPs compared to the general population. These results are confirmed by the 'Survival expectation difference' linear model (b=-0.006).

Lifestyle and Behavioural Risk Factors

BMI differentiates the value of subjective survival probabilities. The RRRs of all 'Full' multinomial models indicate that individuals of normal weight are 11.7% - 14.5% less likely to estimate that future survival will be higher than that of the general population, compared to the obese. Similar results hold for underweight and overweight persons. In other words, obese persons estimate that their future survival will be higher than that of the general population.

Figure 3.14 Impact of physical inactivity on subjective survival probabilities.



Respondents who are physically active report higher SSPs (Figure 3.14).The multinomial regression results show that physical activity increases the chances of estimating that future survival will be higher than that of the general population by 5% to 9%. These results are confirmed by the 'Survival expectation difference' linear model (b=-0.017). Regarding smoking status, the regression results

indicate that non-smokers are more likely to report that future survival will exceed that of the general population, compared to smokers.

Overall, the frequency of consumption of dairy products, meat or chicken, eggs or legumes and fruits or vegetables is negatively correlated with the SSPs (Table 3.7). However the impact depends on the type of food (Table 3.12).

Table 3.12 Impact of consumption frequency on SSPs based on RRRs.

Lower consumption frequency implies…	…reporting higher SSPs compared to OSPs	…reporting lower SSPs compared to OSPs	…reporting SSPs close to OSPs
Meat or chicken	×	×	✓
Dairy	✓	✓	×
Egg or legumes	×	✓	×
Fruit or vegetables	×	✓	×

The multinomial model results suggest that lower consumption frequency of fruits or vegetables and eggs or legumes tend is linked to lower SSPs compared to the general population OSPs. On the other hand, respondents who consume dairy products less frequently tend to estimate that their future survival will be either lower or higher compared to the general population. Finally, those consuming meat or chicken less frequently report SSPs close to the general population OSPs.

3.4. Comparison of SSPs and mortality patterns

In this final section the patterns of subjective survival probabilities are compared to the patterns of actual mortality. Each factor is allocated in one of the following homogeneous groups (Table 3.13).

i. Factors linked to better actual mortality and higher subjective survival probabilities
ii. Factors linked to worse actual mortality and lower subjective survival probabilities
iii. Factors related to the divergence of actual mortality and subjective survival probabilities patterns

Table 3.13 Grouping of factors based on the subjective survival probabilities consistency to actual mortality patterns.

Factors	Factors linked to better actual mortality and higher SSPs	Factors linked to worse actual mortality and lower SSPs	Factors exhibiting inconsistent associations with SSPs and actual mortality
Demographic			
Age			✗
Gender			✗
Marital status			✗
Number of children	✓		
Socio-economic Status			
Higher Education	✓		
Better Income	✓		
Physical & Mental Health			
Chronic diseases		✓	
ADLs		✓	
Poor Self-rated health		✓	
Depression		✓	
Better Numeracy			✗
Better Writing skills	✓		
Better Orientation in time			✗
Poor Memory		✓	
Lifestyle & Behavioural Risk Factors			
BMI			✗
Frequent Physical activity	✓		
Smoking status		✓	
Egg or legumes	✓		
Fruit or vegetables	✓		
Quality of life and Social Support			
Better Quality of Life	✓		
Times received help	✓		
Greater Life satisfaction	✓		

The factors included in the first homogeneous group are associated with better actual mortality and higher subjective survival probabilities. These factors are:

- Better socioeconomic status and education
- Healthier dietary habits and physical exercise
- Better life satisfaction, quality of life and social support
- More children and better writing skills

The factors included in the second homogeneous group are associated with worse actual mortality and lower subjective survival probabilities. These factors are:

- Worse self-rated health, more chronic conditions and ADLs
- Smoking
- Depression and poor memory

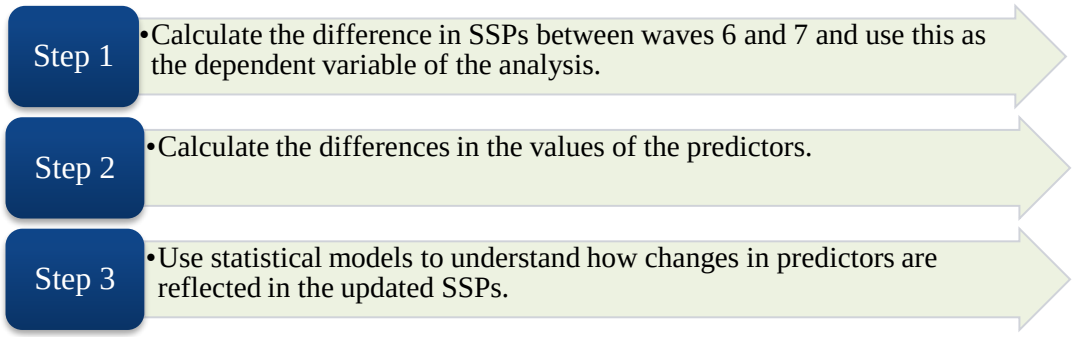
The factors included in the third homogeneous group exhibit inconsistent associations with subjective survival probabilities and actual mortality. These factors are:

- Age, gender and marital status
- Better numeracy and orientation in time
- Body Mass Index

The results are discussed in more detail in Section 7.

4. Chapter 4 - Gender and socio-demographic effects on updating subjective survival probabilities

The second objective of this study is to investigate whether individuals incorporate factors affecting actual mortality when they update their subjective survival probabilities. In other words, the aim is to assess whether subjective survival probabilities are updated consistently with actual mortality patterns, and whether there are variations by gender. To achieve the aforementioned objective, the following process is adopted.



To achieve the objective data from the SHARE wave 6 and wave 7 were used. The data and statistical modelling are described in the following sections. All data sources are also described in Section 1.7. Upward revisions of SSPs are expected after the occurrence of events such as improvement in health, increase in income and increase in life satisfaction. On the other hand, downward revisions of SSPs are expected after the occurrence of events such as an increase in the number of chronic conditions, transitions out of marriage through widowhood or divorce and transitions of BMI from normal to other categories.

4.1.Methods and sample description

The combined SHARE Wave 6 and Wave 7 consist of 49505 respondents. More details are presented in Section 2.5.

Analysis of non-respondents

A comparison of the characteristics between respondents who participated in both waves and those who participated only in Wave 6 is presented in Table 4.1. On average, respondents who participated only in Wave 6 but died before Wave 7, report lower SSPs, are older, less educated, have more chronic conditions and ADLs, report poorer self-rated health, are more frequently underweight and are less satisfied with their lives compare to respondents who participated in both waves. On the other hand, respondents who participated in both waves are more frequently married and have a lower frequency of widowhood. The 'equivalised' income is not materially different between respondents who participated in both waves and those who participated only in Wave 6. Overall, the combined longitudinal sample for both waves includes a lower proportion of older, less educated, less satisfied and less healthy individuals. Males have also lower representation in the sample.

Table 4.1 Characteristics of respondents who participated in both waves compared to those who participated only in Wave 6.

Descriptive measures	Respondents who participated in Waves 6 & 7	Respondents who participated in Wave 6 but not in Wave 7	Respondents who participated in Wave 6 but died before W7
Number of respondents	51,849	14,397	1,985
Subjective survival probabilities - Wave 6 (mean)	64.88%	63.48%	44.25%
Demographic Characteristics			
Gender			
Males	42.98%	44.40%	53.15%
Females	57.02%	55.60%	46.85%
Chronological age (mean)	67.4 years	67 years	79.4 years
Marital Status			
Married	68.39%	68.95%	51.94%
Partnership	1.75%	2.09%	1.21%
Separated	1.35%	1.44%	1.66%
Never married	5.61%	5.30%	5.84%
Divorced	8.46%	7.63%	5.24%
Widowed	14.45%	14.59%	34.11%
Socio-economic Status			
'Equivalised' Income (mean value by quartile)			
1st Quartile (mean in €)	3,880	3,681	3,997
2nd Quartile (mean in €)	10,144	10,115	9,895
3rd Quartile (mean in €)	19,679	19,477	18,971
4th Quartile (mean in €)	57,518	57,011	51,922
Education level			
ISCED-97 code 0 & 1 (Primary)	21.36%	27.73%	40.20%

ISCED-97 code 2 (Lower secondary)	17.54%	16.94%	18.99%
ISCED-97 codes 3 & 4 (Upper secondary)	38.19%	34.26%	29.22%
ISCED-97 codes 5 & 6 (Tertiary)	22.90%	21.06%	11.59%
Physical Health			
Number of chronic conditions (mean)	1.77	1.79	2.70
Number of ADLs (mean)	0.22	0.35	1.34
Self-rated health			
Excellent	7.16%	6.81%	1.36%
Very good	18.29%	17.87%	4.48%
Good	36.74%	33.92%	18.04%
Fair	28.09%	28.05%	38.64%
Poor	9.72%	13.34%	37.48%
Lifestyle and behavioral risk factors			
BMI			
Underweight	1.00%	1.35%	3.43%
Normal	39.20%	41.29%	43.32%
Overweight	40.60%	40.02%	36.32%
Obese	19.19%	17.35%	16.93%
Quality of Life			
Life satisfaction (mean)	7.71	7.47	7.03

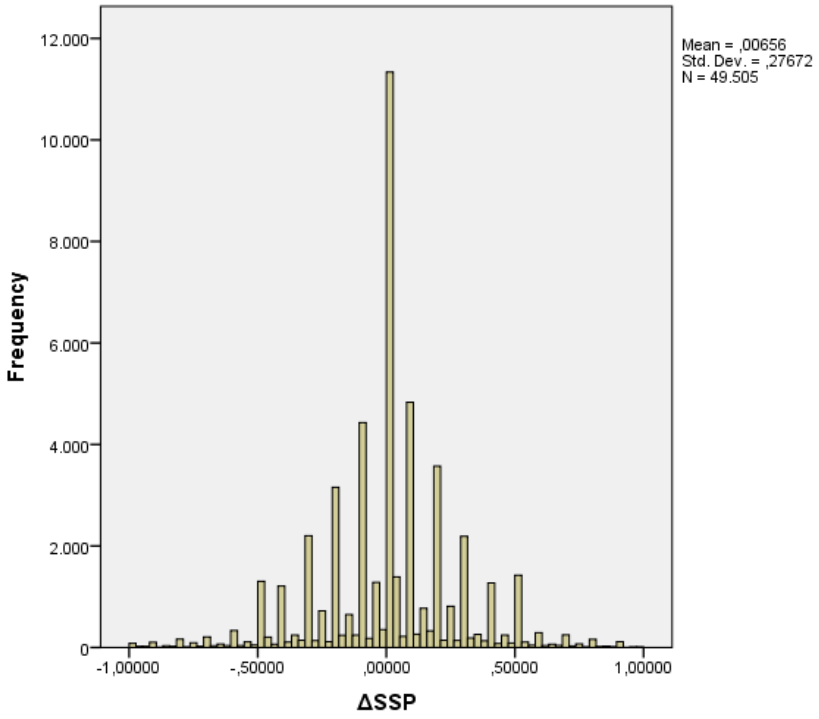
Males, those with lower socio-economic status and lower life satisfaction as well as older respondents demonstrate higher mortality compared to individuals who did not respond for reasons other than death. It is worth noting that the average SSP for respondents who died is 44%; considerably lower than the other two groups. Furthermore, poor self-rated health, more chronic conditions, more functional limitations and underweight BMI category demonstrate a strong association with in-sample mortality.

Dependent variable

The dependent variable is calculated as the difference in the SSPs, reported by the same individual, between SHARE waves 6 and 7.

$$\Delta SSP = SSP_{x,N}^{Wave\ 7} - SSP_{x,N}^{Wave\ 6}$$

Figure 4.1 Distribution of the difference in subjective survival probabilities; *'ΔSSP'*

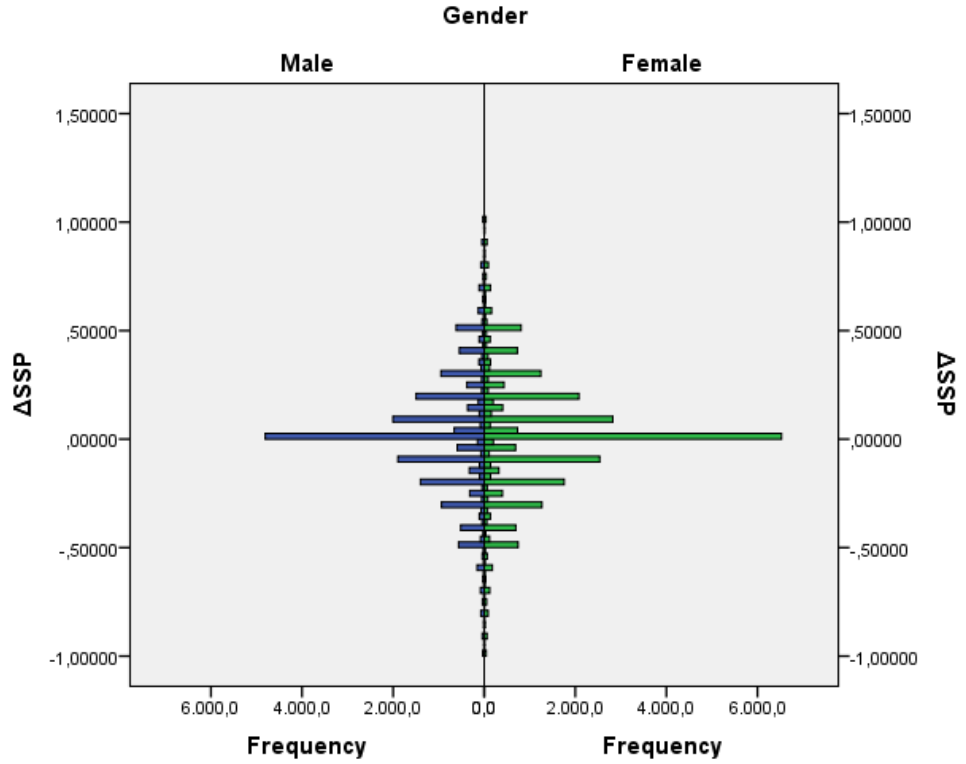


The distribution of the difference in SSPs 'ΔSSP' is concentrated in the range of +/- 5%. Furthermore, the distribution of 'ΔSSP' does not differ by gender (Mann-Whitney U, p=34.5%). On average subjective survival probabilities between waves are marginally increased by 0.66% (Figure 4.2). The average revision of SSPs is higher for females (Table 4.2). Males represent 44% of the sample.

Table 4.2 Average of SSPs and ΔSSP across SHARE waves, by gender.

Gender	Number of respondents	Mean of SSPs Wave 6	Mean of SSPs Wave 7	Mean of ΔSSP
Males	21326 (44%)	65.05%	65.64%	0.58%
Females	28179 (56%)	64.72%	65.43%	0.71%
Total	**49505**	**64.86%**	**65.52%**	**0.66%**

Figure 4.2 Distribution of the difference in subjective survival probabilities; *'ΔSSP'*, by gender.



Explanatory variables

Demographic characteristics

This group of variables includes chronological age at baseline (in years), the increase in chronological age between waves 6 and 7, gender and the change in marital status. The marital status change is re-coded in 4 categories (no change in marital status between waves 6 and 7, becoming widowed, becoming divorced and other changes in marital status). Furthermore, country of residence is used as a control variable since it does not vary between waves. The average respondent who reports no change in marital status is 67 years old.

Table 4.3 Impact of changes in marital status on SSPs' revision.

Marital status change	Mean Age (Wave 6)		Participation in the sample		Mean 'ΔSSP'	
	Males	Females	Males	Females	Males	Females
No change in marital status	67.6	67.2	97.51%	96.16%	0.619%	0.833%
Widowed	75.2	72.9	1.23%	2.69%	-2.910%	-3.840%
Divorced	61.3	63.2	0.30%	0.23%	5.855%	1.456%
Other changes in marital status	66.8	63.6	0.96%	0.92%	-0.329%	1.155%

Age is negatively correlated to 'ΔSSP' (Pearson = -6%, p<1%). The majority of respondents reported no change in marital status, and this is associated with a small positive SSP revision (Table 4.3). The experience of widowhood is associated with negative SSPs revisions whereas the decision of divorce is associated with positive SSPs revisions. Approximately three out of a hundred women experience widowhood in the period during the waves. The frequency of divorces is marginally higher for males.

Socio-economic factors

This group of variables includes the change in the "equivalised" individual income quartiles occurring between waves 6 and 7. Moreover, the change in "equivalised" individual income quartiles is re-coded in three groups: no change in income quartiles between waves; a fall of income to lower quartiles, a rise of income to higher quartiles. The "equivalised" income per individual was calculated using the reported household income and the OECD-modified equivalence scale. This scale, first proposed by Haagenars et al. (1994), assigns a value of 1 to the household head, of 0.5 to each additional adult member and of 0.3 to each child. The aim of this variable is to isolate the impact of income changes on updating SSPs.

Table 4.4 Impact of changes in "equivalised" income quartiles on SSPs' revision.

"Equivalised" income quartiles change	**Mean Age (Wave 6)**		**Participation in the sample**		**Mean 'ΔSSP'**	
	Males	**Females**	**Males**	**Females**	**Males**	**Females**
Fall of income to lower quartiles	68.4	67.4	17.3%	16.0%	-0.32%	1.59%
No change in income quartiles between waves	67.7	67.5	62.3%	62.0%	0.73%	0.35%
Rise of income to higher quartiles	66.7	66.9	20.4%	22.1%	0.89%	1.08%

Approximately six out of ten respondents have no significant change in "equivalised" income quartiles between waves. This category is related to a slight upward revision of SSPs; on average, males revise upwards their SSPs more than females. Males whose "equivalised" income falls to lower quartiles revise downwards their SSPs. In contrast, females whose "equivalised" income falls to lower quartiles revise upwards their SSPs. Both genders revise their SSPs upwards after an increase in "equivalised" income to higher quartiles.

The educational level at baseline in 4 categories, based on the ISCED-97 classification, includes Primary (code 1), Lower Secondary (code 2), Upper Secondary (codes 3 & 4) Tertiary (codes 5 & 6), and is used as a control variable.

Table 4.5 Impact of educational level on SSPs' revision.

Educational level	Mean Age (Wave 6)		Participation in the sample		Mean 'ΔSSP'	
	Males	Females	Males	Females	Males	Females
Primary	72.0	72.3	18.2%	22.8%	-0.16%	-0.58%
Lower Secondary	67.0	67.4	16.6%	18.3%	0.34%	0.93%
Upper secondary	66.3	65.7	39.6%	37.4%	1.26%	1.38%
Tertiary	67.0	64.8	25.6%	21.5%	0.21%	0.74%

Older respondents who have primary education revise downwards their SSPs. In contrast, respondents who have secondary and tertiary education revise upwards their SSPs (Table 4.5).

Physical health

Physical health includes the change in the number of chronic conditions (computed as 'number of chronic conditions at wave 7 minus the respective number at wave 6'), the change in the number of limitations in Activities of Daily Living (computed in the same manner as for chronic conditions) and the change in self-rated health. The change in self-rated health (SRH) is re-coded in 5 levels, namely, no change in SRH between waves 6 and 7, improvement by 1 level, improvement by 2 or more levels, deterioration by 1 level and deterioration by 2 or more levels. The increase in the number of chronic conditions (Pearson = -5%) as well as the increase in the number of ADLs (Pearson = -4.3%) are negatively associated with SSPs' revisions.

Figure 4.3 Distribution of change in SRH level by gender.

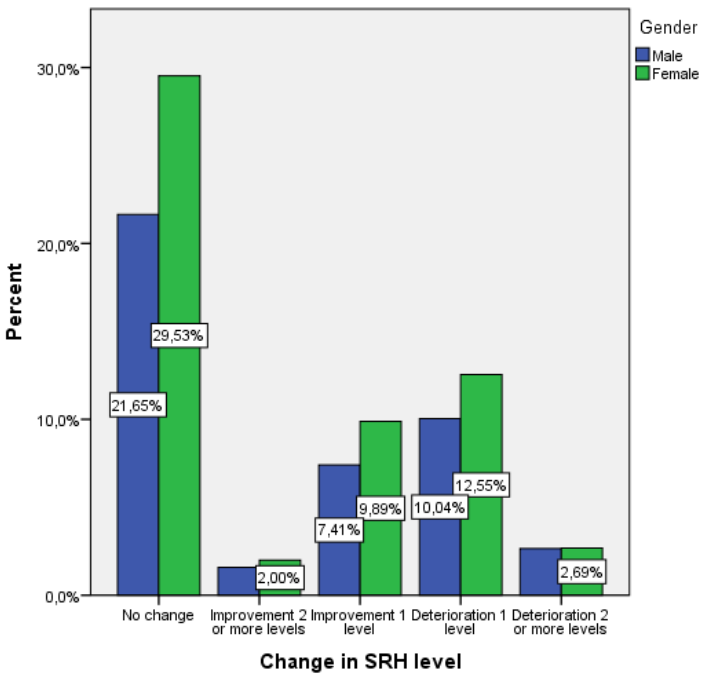


Table 4.6 Impact of SRH change on SSPs' revisions.

Average SSPs revision based on SRH changes	Males	Females	Total
No change in SHR level between waves	1.04%	0.56%	0.76%
Improvement by 2 or more levels	7.68%	8.66%	8.23%
Improvement by 1 or more level	3.67%	3.73%	3.71%
Deterioration by 1 or more level	-2.14%	-1.57%	-1.82%
Deterioration by 2 or more levels	-5.61%	-3.90%	-4.75%

Both genders are more likely to report no change in self-rated health between SHARE waves (Figure 4.3). Improvement of self-rated health is associated with upwards SSP revisions whereas deterioration of self-rated health is associated with downward SSP revisions (Table 4.6). Overall, respondents who report no change in SRH revise their SSPs upwards by 0.76% on average. However, the positive revision of SSPs is driven by females only. After a deterioration of SRH, both genders revise negatively theirs SSPs. On average males report a bigger increase in the number of chronic conditions and ADLs compared to females (Table 4.7).

Table 4.7 Average increase in chronic conditions and ADLs by gender

	Average of increase in the number of chronic conditions	Average of increase in the number of ADLs
Males	0.17	0.06
Females	0.13	0.05

Lifestyle & Behavioral risk factors

This group includes changes in the BMI category of a respondent between the waves. This variable is re-coded in nine levels, namely, no change in BMI group between waves 6 and 7, change from underweight to normal, change from underweight to overweight or obese, change from normal to underweight, change from normal to overweight or obese, change from overweight to obese, change from overweight to normal or underweight, change from obese to overweight and change from obese to normal or underweight (Table 3.8).

Table 3.8 Grouping of BMI transitions in a single variable.

BMI Wave 6	BMI Wave 7			
	Underweight	Normal	Overweight	Obese
Underweight	1	2	3	3
Normal	4	1	5	5
Overweight	7	7	1	6
Obese	9	9	8	1

Males and females revise negatively their SSPs after a reduction of BMI from normal to underweight (Table 4.9). Only males revise downwards their SSPs after a change of BMI from normal to overweight. Moreover, only females revise downwards their SSPs after a change of BMI from overweight to obese.

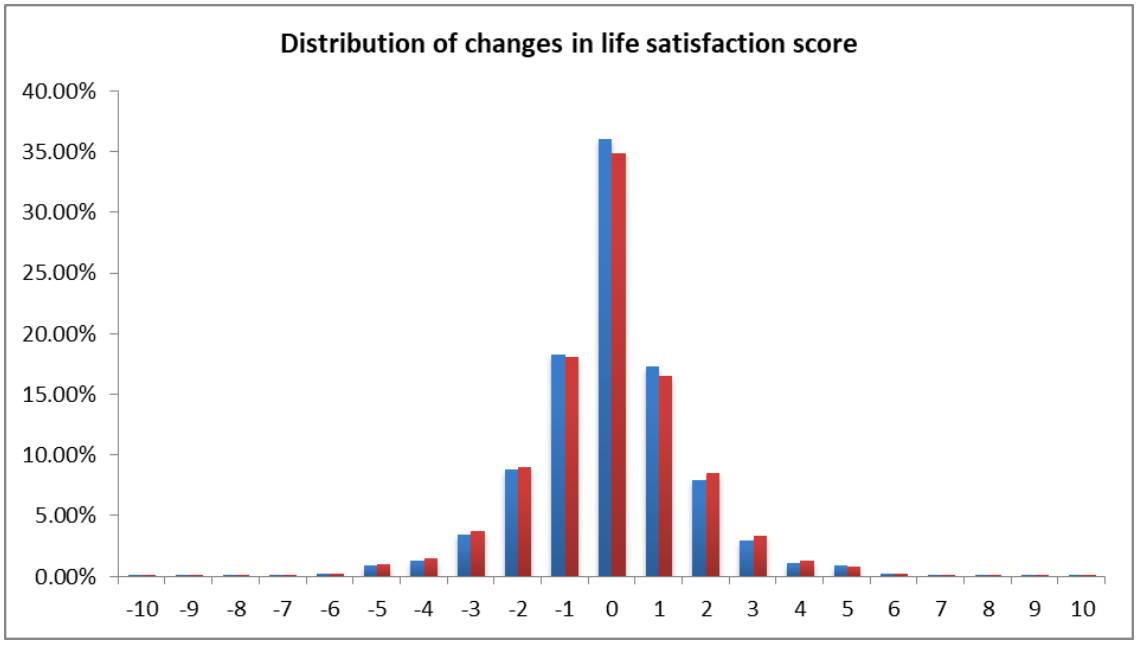
Table 4.9 Frequencies of BMI transitions and average SSP revision.

Change in BMI category	Males		Females	
	Frequencies	Mean of 'ΔSSP'	Frequencies	Mean of 'ΔSSP'
No change in BMI category between waves (1)	79.69%	0.68%	79.40%	0.75%
Change from underweight to normal (2)	0.14%	9.17%	0.49%	0.80%
Change from underweight to overweight or obese (3)	0.01%	9.93%	0.07%	3.52%
Change from normal to underweight (4)	0.25%	-11.30%	0.64%	-0.42%
Change from normal to overweight or obese (5)	6.11%	1.33%	6.24%	1.39%
Change from overweight to obese (6)	3.40%	0.94%	3.63%	-0.10%
Change from overweight to normal or underweight (7)	6.39%	-0.12%	5.40%	0.21%
Change from obese to overweight (8)	3.62%	-1.35%	3.50%	0.52%
Change from obese to normal or underweight (9)	0.38%	-0.46%	0.63%	-0.39%

Quality of life

This group includes the change in the life satisfaction score. Improvement in life satisfaction score is associated with upward revisions of SSPs (Pearson = 13.2%, p<1%). The distribution of changes in life satisfaction score is similar for both genders (Figure 4.4).

Figure 4.4 Distribution of changes in life satisfaction score by gender.



4.2.Statistical modelling

In the analysis we use two Generalised Linear Models (GLMs) with a linear link function as well as ordinary least squares regressions. SPSS 21 is used for the analyses. The dependent variable is *'ΔSSP'* and is assumed to be normally distributed. The distribution of 'ΔSSP' is bell shaped (Figure 4.1) but it does not satisfy the strict criteria of normality (KS- test, p<1%). The models are estimated for males and females separately.

4.3.Results

The sings of the models' coefficients indicate the influence of predictors on the revision of the SSPs. Positive coefficients indicate an upward revision whereas negative coefficients indicate a downward revision (Table 4.10).Chronological age is negatively associated with the difference of SSPs (b= -0.001). Individuals tend to revise downwards their SSPs as they become older. However, the increase in chronological age is weakly associated with a positive revision of SSPs for males (b=0.001) and a negative revision of SSPs for females (b= -0.009). Widowhood is associated with negative SSPs revisions for both genders. This association is stronger for females (b= -0.025). Getting a divorce is associated with positive SSPs revisions for males (b= 0.048) but not for females (b= -0.002).

Table 4.10 Coefficients based on generalised linear models and ordinary least squares regressions.

Independent Variables	Males		Females	
	GLM	**OLS**	**GLM**	**OLS**
Intercept	0.105**	0.104**	0.104**	0.117**
Demographic Characteristics				
Chronological age at Wave 6	-0.001**	-0.001**	-0.001**	-0.001**
Increase in chronological age between waves	0.001	0.000	-0.009	-0.009
Change in Marital Status (reference : No change in marital status)				
Widowed	-0.015	-0.015	-0.025*	-0.025
Divorced	0.048	0.047	-0.002	-0.004
Other change in marital status	-0.007	-0.008	-0.002	-0.005
Socio-economic Status				
Change in 'Equivalised' Income quartiles (reference : No change in income quartiles)				
Income falls to lower quartiles	-0.007	-0.008	0.013*	0.012**
Income rises to higher quartiles	0.002	0.001	0.006	0.005
Physical Health				
Change in the number of chronic conditions	-0.005**	-0.005**	-0.007**	-0.007**
Change in the number of ADLs	-0.012**	-0.012**	-0.006**	-0.006**
Change in self-rated health (reference : No change in self-rated health)				
Improvement by 2 or more scales	0.051**	0.049**	0.060**	0.060**
Improvement by 1 scale	0.019**	0.019**	0.023**	0.023**
Deterioration by 1 scale	-0.026**	-0.027**	-0.016**	-0.017**
Deterioration by 2 or more scales	-0.055**	-0.056**	-0.034**	-0.034**
Lifestyle and behavioral risk factors				
Change in BMI category (reference : No change in BMI category)				
Change from underweight to normal	0.103*	0.103*	-0.007	-0.006

Change from underweight to overweight or obese	0.153	0.141	0.060	0.058
Change from normal to underweight	-0.111**	-0.112**	-0.006	-0.008
Change from normal to overweight or obese	0.009	0.006	0.009	0.006
Change from overweight to obese	0.006	0.006	-0.002	-0.003
Change from overweight to normal or underweight	0.002	0.000	0.002	0.001
Change from obese to overweight	-0.018	-0.018	0.002	0.001
Change from obese to normal or underweight	0.012	0.009	-0.007	-0.009
Quality of Life				
Change in life satisfaction	0.019**	0.019**	0.019**	0.018**

**p<1% , *p<5% . The dependent variable the difference in the subjective survival probabilities 'ΔSSP'. Controlling for country of residence and educational attainment.

Males and females who experienced increases in 'equivalised' income tend to revise upwards their SSPs (b=0.002 and b=0.006). In contrast, following an income drop males revise downwards their SSPs (b= -0.007) whereas females revise upwards their SSPs (b= 0.013). The association of changes in health status with the revision of SSPs is consistent for both genders. More specifically, more chronic conditions and ADLs are strongly associated with negative SSPs revisions. Furthermore, an improvement of self-rated health is associated with positive SSPs revisions whereas a deterioration of self-rated health is associated with negative SSPs revisions.

Males who increase their BMI from underweight to normal or from underweight to overweight or obese, tend to revise upwards their SSPs (b=0.103 and b=0.153). Furthermore, males and females who reduce their BMI from normal to underweight tend to revise downwards their SSPs (b= -0.111 and b= -0.006). A change of BMI group from obese to overweight is associated with a decrease in SSPs for males (b= -0.018) and an increase in SSPs for females (b= 0.002). An increase in life satisfaction is associated with positive SSPs revisions for both genders. Overall, the results of generalised linear models and ordinary least squares regressions are consistent.

4.4. Comparison of SSPs' revisions and mortality patterns

In this final section the patterns of subjective survival probabilities' revisions are compared to the patterns of actual mortality. Each factor is allocated in one of the following homogeneous groups (Table 4.11).

i. Factors linked to better actual mortality and upward revision of subjective survival probabilities
ii. Factors linked to worse actual mortality and downward revision of subjective survival probabilities
iii. Factors exhibiting inconsistent associations with subjective survival probabilities ' revisions and actual mortality

Table 4.11 Grouping of factors based on the subjective survival probabilities consistency to actual mortality patterns.

Factors	Factors linked to better actual mortality and upward SSPs' revisions	Factors linked to worse actual mortality and downward SSPs' revisions	Factors exhibiting inconsistent associations with SSPs' revisions and actual mortality
Demographic			
Age		✓	
Change in Marital Status: Widowed		✓	
Change in Marital Status: Divorced		✓ *(for females)*	✗ *(for males)*
Socio-economic Status			
Increase in Income	✓		
Decrease in Income		✓ *(for males)*	✗ *(for females)*
Physical & Mental Health			
Increase in Chronic diseases		✓	
Increase in ADLs		✓	
Improvement in Self-rated health	✓		
Lifestyle & Behavioural Risk Factors			
BMI : Change from normal to underweight		✓	
BMI : Change from normal to overweight or obese			✗
BMI : Change from underweight to normal	✓ *(for males)*		✗ *(for females)*
Quality of life and Social Support			
Increase in life satisfaction	✓		

The factors included in the first homogeneous group are associated with better actual mortality and upward revision of subjective survival probabilities. These factors are:

- Increase in income and life satisfaction
- Improvement in Self-rated health
- BMI change from underweight to normal for males

The factors included in the second homogeneous group are associated with worse actual mortality and downward revision subjective survival probabilities. These factors are:

- Getting older
- Widowhood
- The transition out of marriage though a divorce for females
- Increase in the number of chronic conditions and ADLs
- BMI change from normal to nor underweight

The factors included in the third homogeneous group exhibit inconsistent associations with the direction of subjective survival probabilities' revisions and actual mortality patterns. These factors are:

- The transition out of marriage though a divorce for males
- Increase in income for females
- BMI change from underweight to normal for females

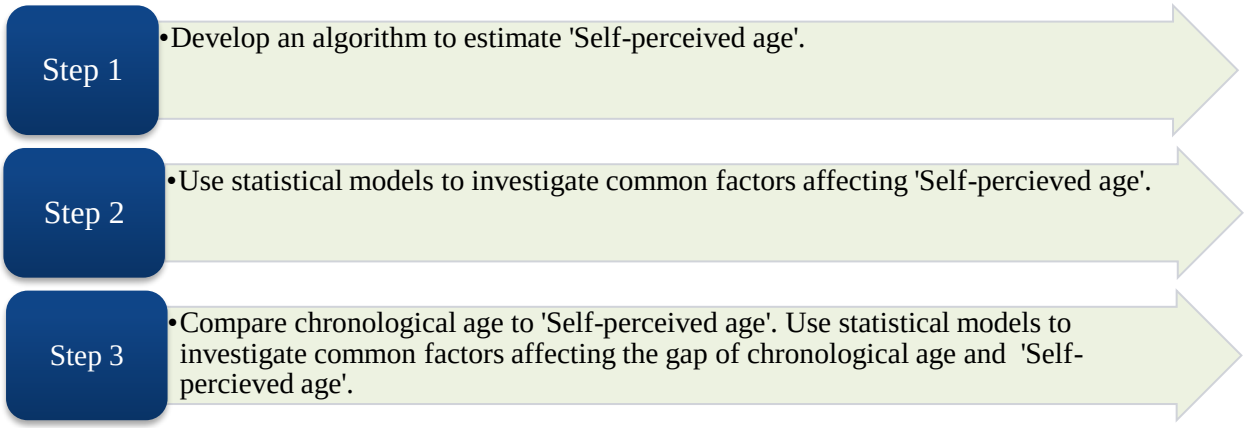
The results, including relevant literature, are discussed in more detail in Section 7.

5. Chapter 5 – ‘Self-perceived age’

‘Self-perceived age’ is a quantity, expressed in years, which incorporates survival-related information from two sources, namely, population life tables and subjective survival probabilities. The concept, the calculation and the patterns of ‘Self-perceived age’ are investigated in this chapter.

5.1. Research objective

Chronological age, gender and population life tables are commonly used to estimate an individual’s average life expectancy. However there is a range of factors such as socio-economic status, self-rated health and smoking which affect future life expectancy (Antonovsky 1967; Idler and Benyamini 1997; Verropoulou 2014; Doll et al. 1994) which are not accounted for in life tables. Subjective survival probabilities incorporate information on future survival and vary across socio-demographic factors (Griffin et al. 2013; Mirowsky 1999; Ross and Mirowsky 2002; Rarrange et al. 2016; Balia 2014; Hurd and McGarry 1995). Therefore, a more flexible definition of ‘age’ is required in order to incorporate additional important information about mortality. The third objective of this study is the development of a method for the quantification of ‘Self-perceived age’. The aim of ‘Self-perceived age’ is to incorporate information from the general population life tables as well as the self-reported subjective survival probabilities. To achieve the aforementioned objective, the following process is adopted.



To achieve the objective, data from the harmonized version of SHARE wave 6 (EU) and HRS wave 12 (USA) were used. The data and statistical modelling are described in the following sections. All data sources are also described in Section 2.5.

Concept of 'Self-perceived age'

'Self-perceived age' is a quantity, expressed in years, which incorporates information from population life tables and subjective survival probabilities. Biological Age is a similar quantity, but its estimation requires a large dataset of biomarkers (Klemera and Doubal 2006). The main advantages of 'Self-perceived age' are:

- It is expressed in years and can, directly, be compared to chronological age
- It incorporates information about general population mortality as well as subjective survival probabilities
- The estimation process of 'Self-perceived age' requires less data compared to 'Biological Age'.

It is worth mentioning that 'Self-perceived age' reflects expectations specifically about one's own future survival as opposed to a general feeling related to one's own aging process, which is reflected in 'Subjective Age'.

5.2.Methods and sample description

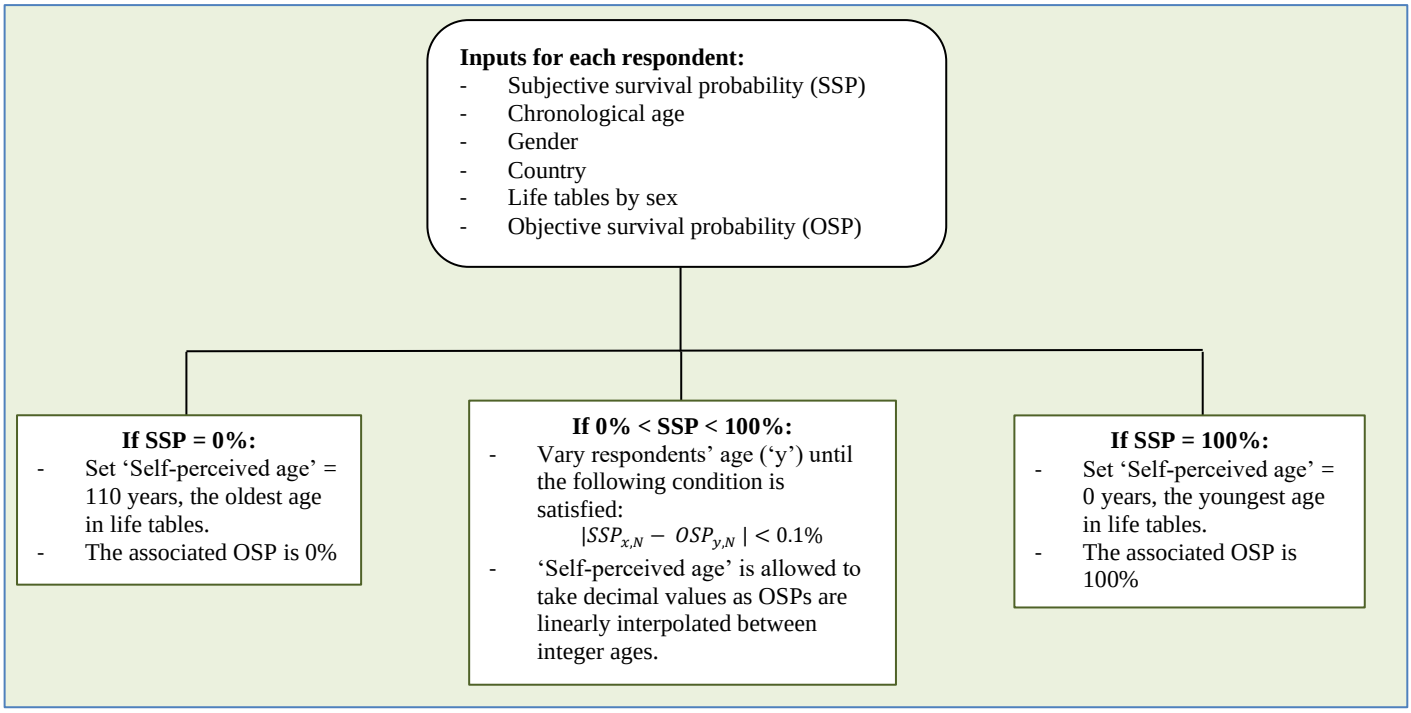
The first step is the development of an algorithm to estimate 'Self-perceived age', which is the dependent variable. The algorithm is coded and executed in Visual Basic for Applications (see Appendix).

Dependent variables

Calculation of 'Self-perceived age'

The inputs required for the calculation of 'Self-perceived age' for each individual include the self-reported SSP, the chronological age, gender, country of residence and the population life tables provided by HMD (Figure 5.1). In addition, the cumulative or objective survival probabilities calculated from population life tables are calculated.

Figure 5.1 Structure of the algorithm to estimate 'Self-perceived age'.



Next, the age of the respondent is changed in order to minimize the difference of SSP and OSP. Finally, boundary conditions are applied if an individual reports SSP = 0% or SSP = 100%. In particular, if the reported SSP = 100% then 'Self-perceived age' is set to 0 years, which is the youngest age in the population life table. In contrast, if SSP = 0% then 'Self-perceived age' is set to 110, which is the oldest age in the population life tables.

For example, let us suppose that an American male aged 74 reported a Subjective Survival Probability of 50% to be alive over the next 11 years.

- The first step is to calculate the Objective Survival Probabilities (OSP) by reference to period life tables by country and sex. OSP is 48.8%, which differs from the self-reported Subjective Survival Probability (SSP). The fact that OSP is close to SSP means that he thinks that his future life expectancy is close to general population.
- The second step is to calculate how many years younger he thinks he is, based on the SSP and the life table. More specifically, we change his Chronological Age until the calculated OSP is equal to 50%. The result is 74.66 years old.
- Hence, we conclude that his 'Self-perceived age' is 74.66 years whereas his Chronological Age is 74 years old. In this example, 'Self-perceived age' is similar to Chronological Age.

The distribution of ‘Self-perceived age’ is centered at age 58 for males and 61 for females (Figure 5.2 and 5.3). The concertation at ages 0 and 111 is notable. The features of respondents who report SSP = 0% and SSP = 100% are presented in Table 5.1.

Figure 5.2 Distribution of ‘Self-perceived age’ for the total sample.

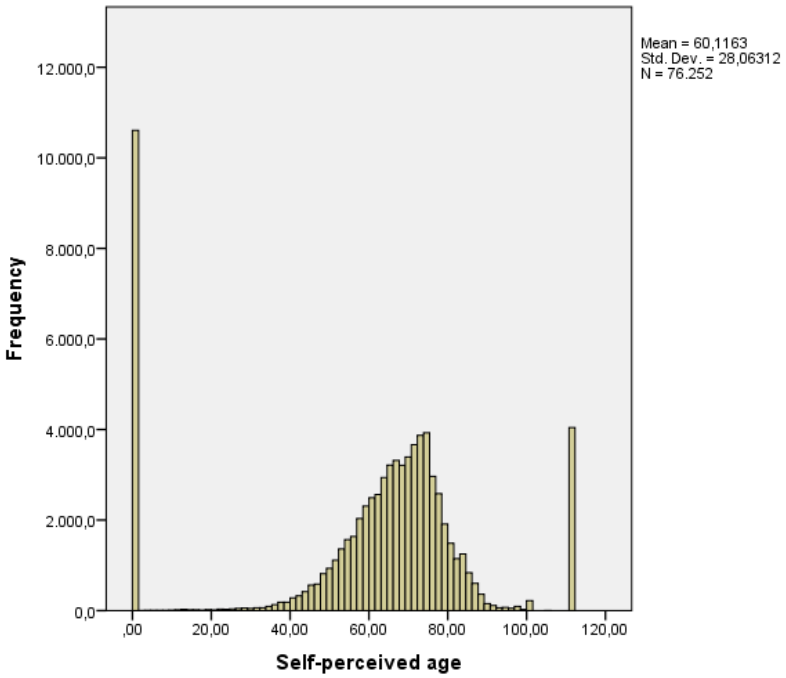


Figure 5.3 Distribution of ‘Self-perceived age’ by gender.

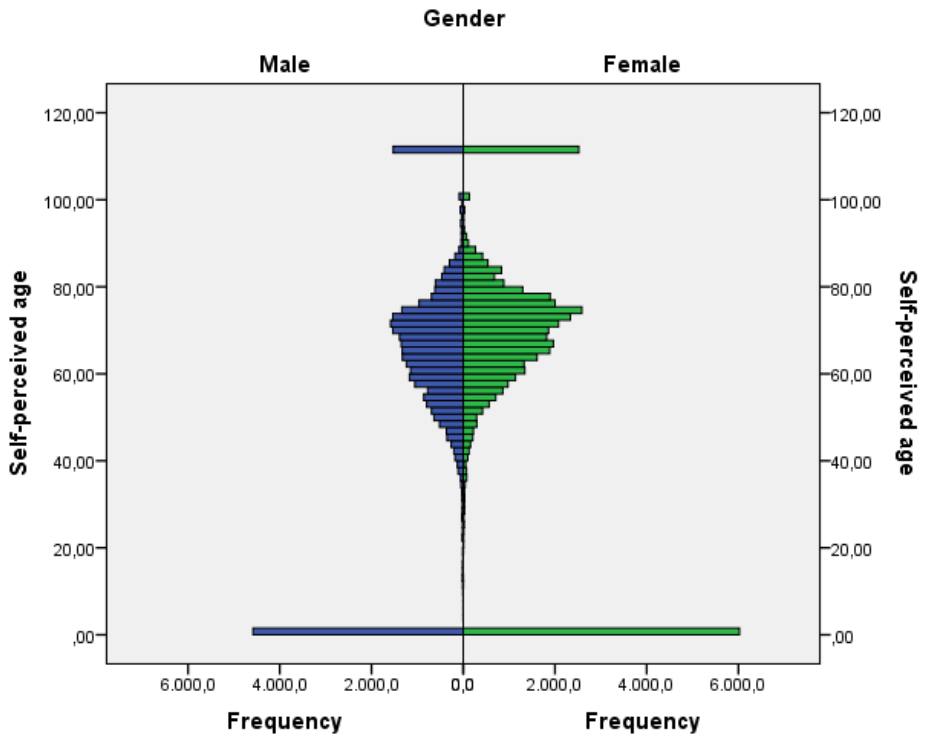


Table 5.1 Features of respondents who report SSP = 0% and SSP = 100%.

Socio-demographic factor	Respondents who report SSP = 0%		Respondents who report SSP = 100%	
	Males	Females	Males	Females
Number of respondents	1,774	2,596	4,580	6,019
Mean chronological age	72.6	74.0	64.8	63.8
Mean self-rated health	3.9	3.9	2.8	2.8
Mean income quartile	2.2	1.7	2.7	2.5
Mean of vigorous activities frequency	4.1	4.5	3.2	3.4
Mean number of parents alive	0.1	0.1	0.4	0.4
Mean score of 'serial 7s' test	3.7	3.2	4.3	4.1
Mean score of word recall test	7.3	7.6	9.4	10.4

Respondents who report SSP = 0% tend to be older, have poorer health, avoid physical exercise and have poorer cognitive function. In contrast, respondents who report SSP = 100% tend to be wealthier, younger, have better health and have on average more parents alive.

Calculation of gap between 'Self-perceived age' and chronological age

The difference between Self-perceived age' (SPA) and chronological age (CA) is calculated in discrete form and it serves as an additional dependent variable. This difference was compared to a tolerance level, in order to capture different levels of closeness between 'Self-perceived' and Chronological Age. The outcome variable is defined below.

$$SPA - CA\ gap = \begin{Bmatrix} 1,\ \ if\ |SPA - CA| \leq tolerance\ (in\ years) \\ Else\ if\ \ |SPA - CA| > tolerance\ (in\ years): \\ 2,\ \ if\ SPA - CA < 0 \\ 3,\ \ if SPA - CA > 0 \end{Bmatrix}$$

The tolerance level is set to 3, 5 and 7 years and the interpretation is: 1 indicates that 'Self-perceived age' is close to Chronological Age; 2 implies that 'Self-perceived age' is younger than Chronological Age, while 3 means that 'Self-perceived age' is older. As the tolerance level increases, the difference between 'Self-perceived age' and Chronological Age increases (Table 5.2).

Table 5.2 'Self-perceived age' and chronological age across tolerance levels.

Tolerance level	Gender	'Self-perceived age' is close to Chronological Age		Self-perceived age' is younger than Chronological Age		'Self-perceived age' is older than Chronological Age	
		Mean Chronological age'	Mean 'Self-perceived age'	Mean Chronological age'	Mean 'Self-perceived age'	Mean Chronological age'	Mean 'Self-perceived age'
3 years	Males	67.2	67	68.9	42.7	65.3	78.9
	Females	68.2	68.2	68.6	38.2	64.5	77.4
5 years	Males	66.8	66.6	69.2	39.4	65.4	81.3
	Females	67.8	68.2	68.4	34.3	64.2	79.2
7 years	Males	67.1	67.0	68.9	36.2	65.4	84.0
	Females	67.4	68.0	68.1	29.5	64.3	81.2

Explanatory variables

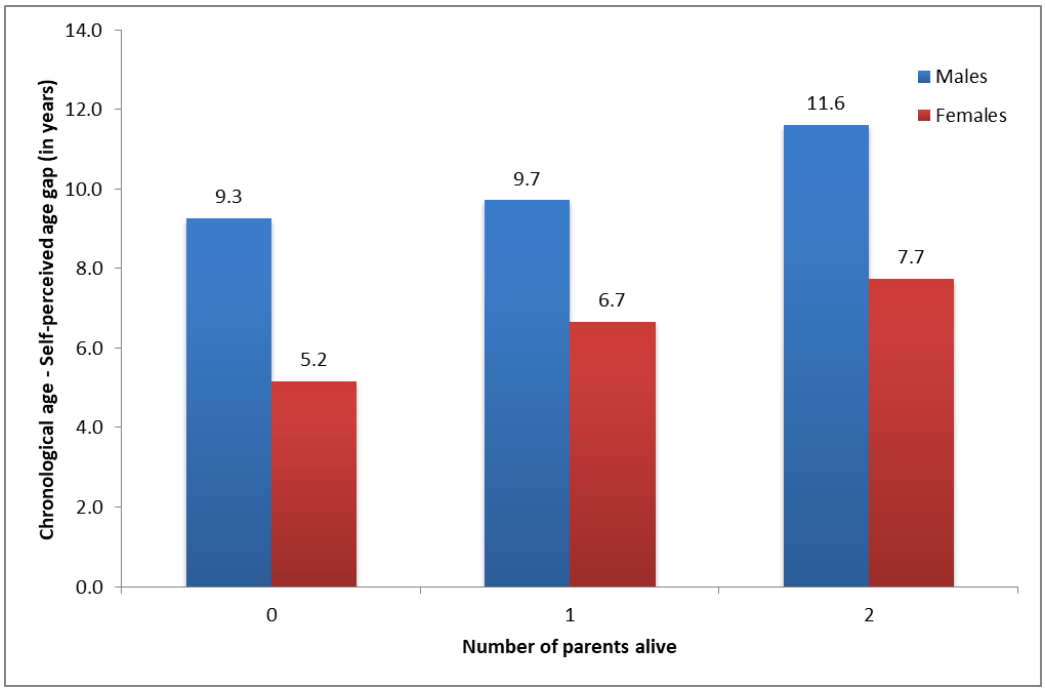
Demographic characteristics

This group of variables includes gender, the number of parents alive and chronological age. Chronological age is strongly associated with 'Self-perceived age' (Pearson = 33.1% , p<1%). Males represent 43.5% of the sample and they have younger 'Self-perceived age' compared to females (Table 5.3). Three out of ten males and four out of ten females have no parents alive. Parental survival is associated with younger chronological and 'Self-perceived age'. The gap of chronological and 'Self-perceived age' increases in line with the number of parents alive (Figure 5.4). In other words, individuals perceive themselves to be younger if their parents are alive. This trend is consistent for both genders.

Table 5.3 Descriptive measures for demographic variables.

Number of parents alive	Males			Females		
	Sample participation	Mean 'Self-perceived age'	Mean chronological age	Sample participation	Mean 'Self-perceived age'	Mean chronological age
0	33.40%	60.9	70.2	42.54%	64.9	70.1
1	7.72%	50.1	59.9	10.63%	52.6	59.2
2	2.36%	45.5	57.1	3.36%	48.6	56.3

Figure 5.4 Difference of chronological and 'Self-perceived age, by number of parents alive.



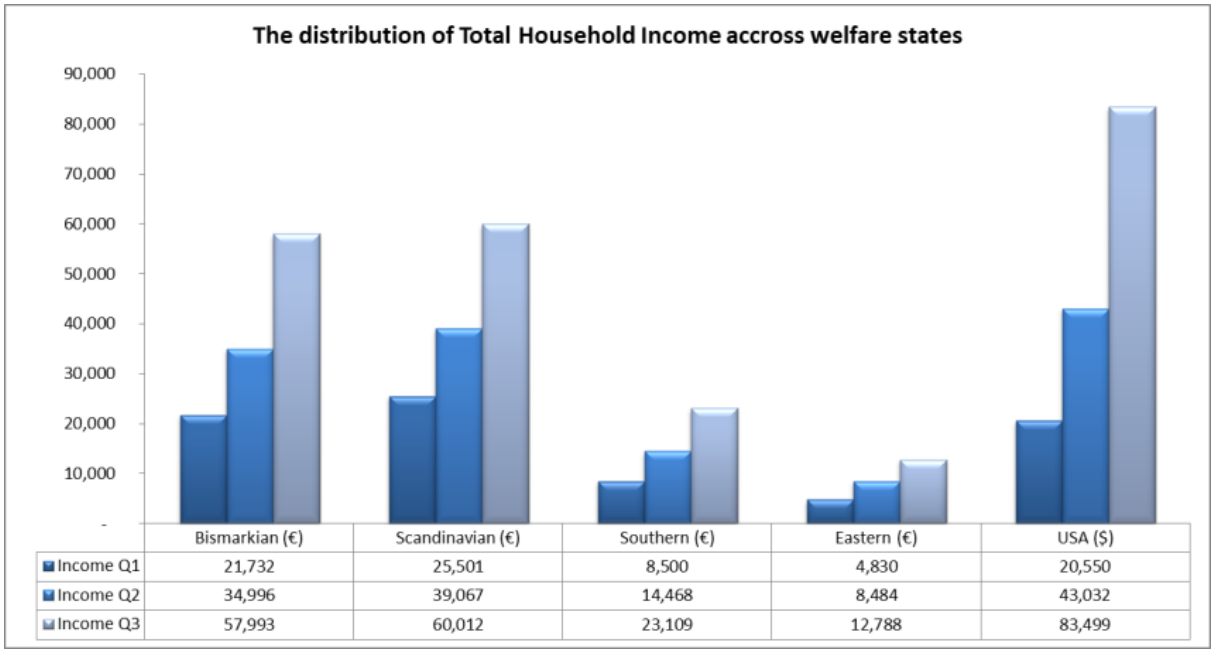
Socio-economic status

This group of variables includes total household income; total wealth and educational level in 3 categories (Lower than upper Secondary, Upper Secondary and Tertiary). The quartiles of total household income and total wealth are calculated for each welfare state and then included in the models (Eikemo et al. 2008). USA is treated as a separate welfare state (Figure 5.5). Southern and Eastern Europeans have the lowest income and wealth whereas Bismarkians the highest (Figure 5.6 and 5.7). Americans have the largest income and wealth inequality (i.e. difference of Income Q3 from Q1 as a percentage of Q2). Furthermore, respondents who completed tertiary education have younger 'Self-perceived age' (Figure 5.8). Overall better socio-economic status is associated with younger 'Self-perceived age'.

Figure 5.5 Allocation of European countries in welfare states.

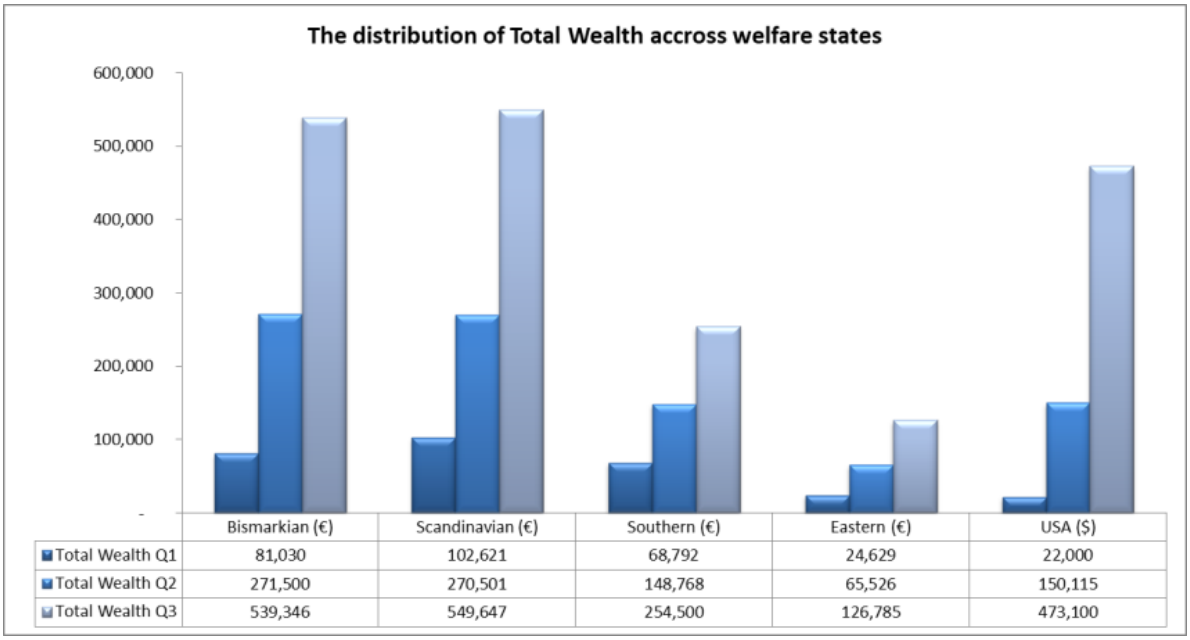
Bismarkian	Eastern European	Southern European	Scandinavian
• Austria • Germany • France • Switzerland • Belgium • Luxembourg	• Czech Republic • Poland • Slovenia • Estonia • Croatia	• Spain • Italy • Greece • Portugal	• Sweden • Denmark

Figure 5.6 Total household Income quartiles across welfare states.



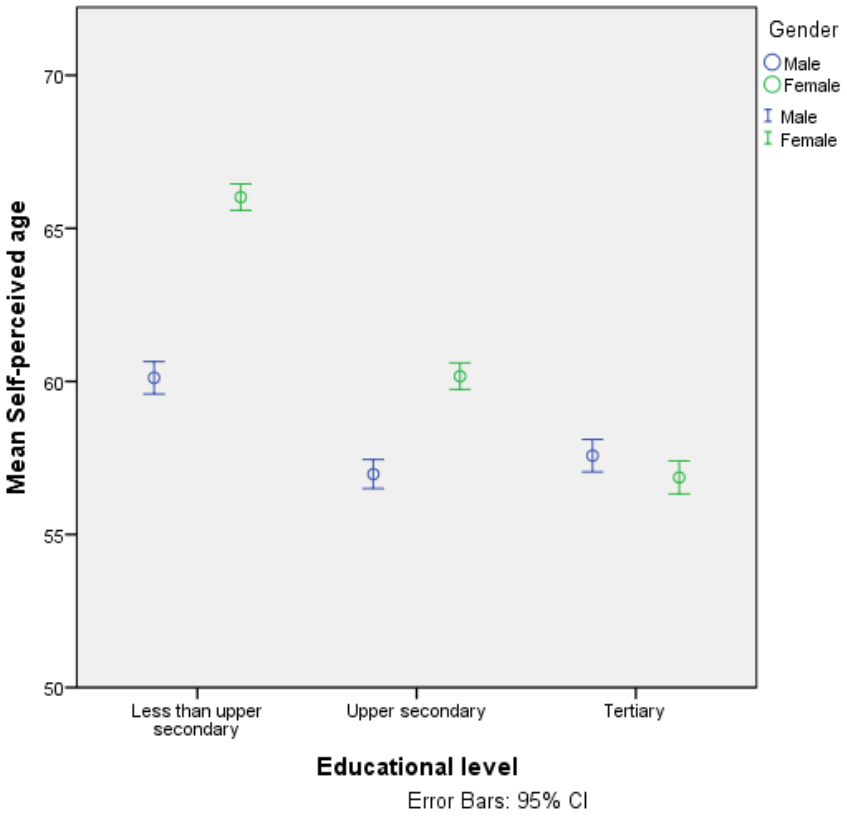
	Bismarkian (€)	Scandinavian (€)	Southern (€)	Eastern (€)	USA ($)
Income Q1	21,732	25,501	8,500	4,830	20,550
Income Q2	34,996	39,067	14,468	8,484	43,032
Income Q3	57,993	60,012	23,109	12,788	83,499

Figure 5.7 Total Wealth quartiles across welfare states.



	Bismarkian (€)	Scandinavian (€)	Southern (€)	Eastern (€)	USA ($)
Total Wealth Q1	81,030	102,621	68,792	24,629	22,000
Total Wealth Q2	271,500	270,501	148,768	65,526	150,115
Total Wealth Q3	539,346	549,647	254,500	126,785	473,100

Figure 5.8 Variation of 'Self-perceived age by educational level and gender.



Physical Health

This group of variables includes the number of limitations in Activities of Daily Living (out of a list of 5 basic, everyday tasks), self-rated health (ranging from 1=excellent to 5=poor) and the number of mobility difficulties (ranging from 0 to 5). On average, Europeans report fewer ADL limitations and mobility difficulties compared to Americans. However, self-rated health is on average better for Americans (Table 5.4).

Table 5.4 Descriptive measures for health variables

Variable	European males	European females	American males	American females
Number of ADLs (mean [SD])	0.15 [0.57]	0.18 [0.6]	0.28 [0.8]	0.35 [0.9]
Self-rated health (mean [SD])	3.08 [1.05]	3.16 [1.04]	2.88 [1.05]	2.91 [1.05]
Mobility Index (mean [SD])	0.40 [0.8]	0.56 [0.9]	0.95 [1.39]	1.31 [1.52]

Worse health is associated with older 'Self-perceived age' (Table 5.5, Figure 5.8). In addition, those in excellent health have younger 'Self-perceived age' whereas those in poor health have older 'Self-perceived age' compared to their chronological age (Figures 5.9 and 5.10).

Table 5.5 Correlation of 'Self-perceived age' with physical health variables.

Health variable	Pearson	Spearman's rho
Number of limitations in Activities of Daily Living	26.1%**	32.3%**
Self-rated health	14.8%**	19.3%**
Number of mobility difficulties	24.1%**	30.8%**

*** p<1%*

Figure 5.9 Association of ADLs and mobility difficulties with 'Self-perceived age'.

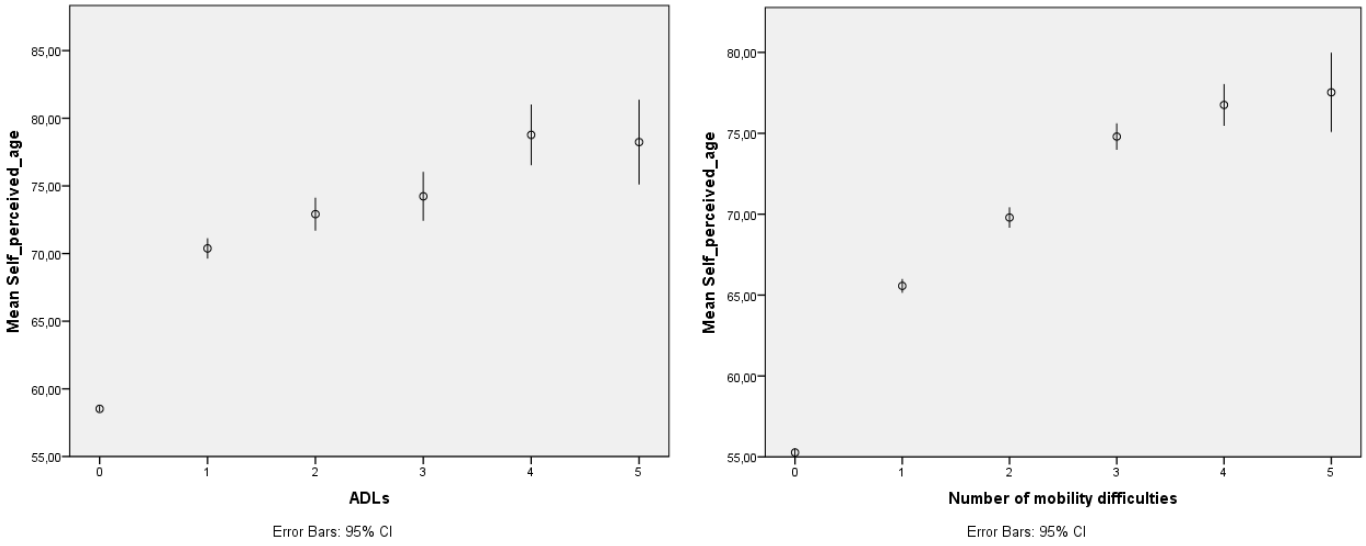


Figure 5.10 Relationship between 'Self-perceived age' and Chronological Age as self-rated health deteriorates.

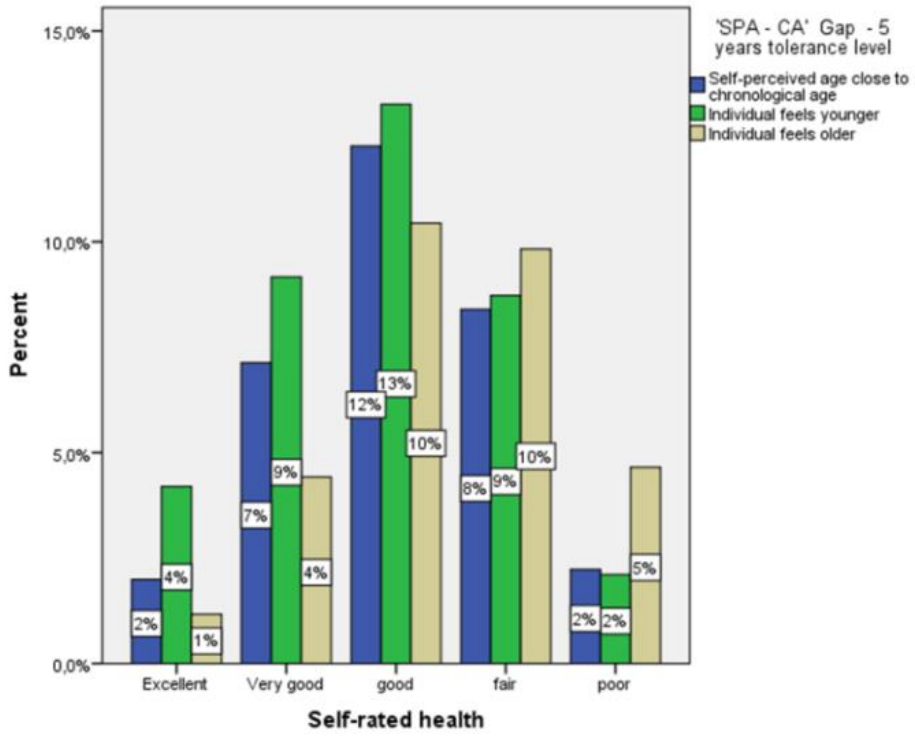
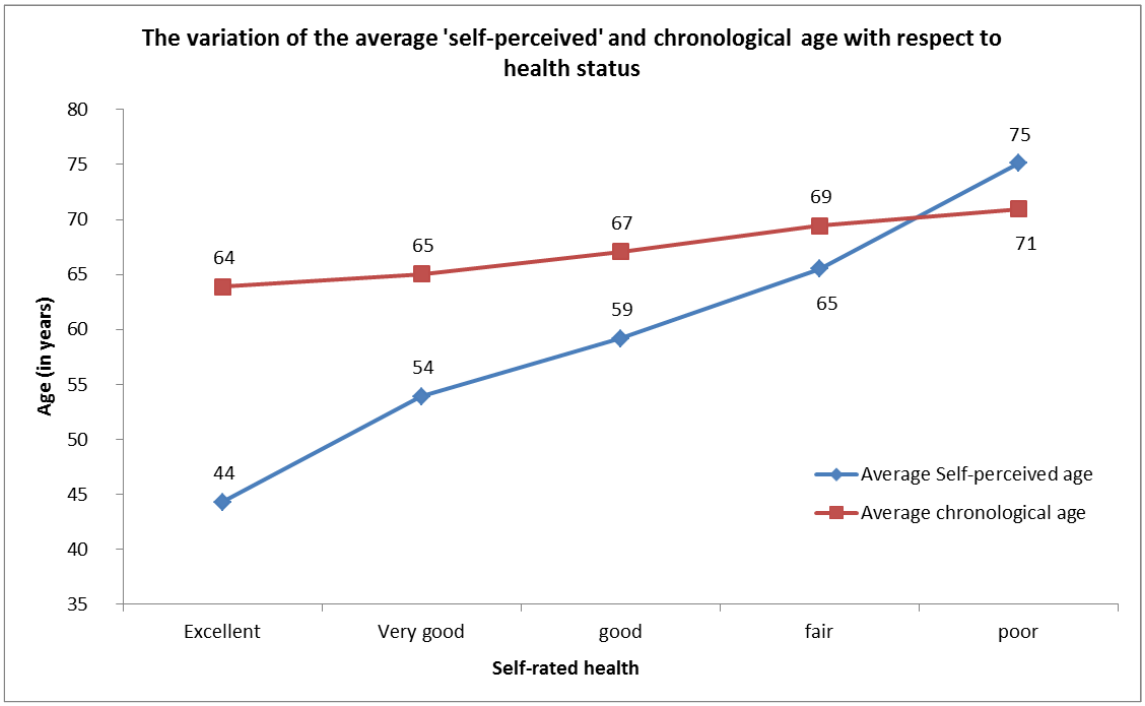


Figure 5.11 Comparison of 'Self-perceived age' and Chronological Age by self-rated health level.



Individuals who rated their health as excellent, very good, good or fair have self-perceived age younger than their chronological age. On the other hand, respondents who rate their health as poor have self-

perceived age older than their chronological age. It is worth noting that, 'Self-perceived age' crosses Chronological age from below at age 70, as self-rated health deteriorates (Figure 5.11).

Cognitive function

Cognitive function is represented by the respondents' scores in the serial 7s and the word recall tests. In the serial 7s test the respondent is asked to subtract 7 from a given number and the score provides the number of correct subtractions. The score ranges from 0 to 5. The total word recall test includes the sum of the immediate and of the delayed word recall scores. The score takes values from 0 to 20. The word recall score counts the number of words recalled correctly immediately and after a delay of 5 minutes. Cognitive function is better for Europeans regarding the serial 7s test while Americans score slightly higher in the word recall test. Furthermore, European and American females exhibit better word recall compared to males but the opposite holds regarding the serial 7s test (Table 5.6).

Table 5.6 Descriptive measures for cognitive function variables

Variable	European males	European females	American males	American females
Serial 7s test (mean [SD])	4.40 [1.09]	4.19 [1.28]	3.74 [1.52]	3.33 [1.7]
Word Recall score (mean [SD])	9.11 [3.4]	9.80 [3.6]	9.28 [3.3]	10.20 [3.4]

Better cognitive function is strongly associated with younger 'Self-perceived age' (Table 5.7). Furthermore, the difference of chronological age and 'Self-perceived age' increases as cognitive function improves (Figures 5.12 and 5.13). Males have a wider age gap than females.

Table 5.7 Correlation of 'Self-perceived age' with cognitive function variables.

Cognitive function variable	Pearson	Spearman's rho
Serial 7s test	-17.5%**	-9.8%**
Total word recall test	-23.9%**	-12.8%**

** *p<1%*

Figure 5.12 Variation of the difference between Chronological Age and 'Self-perceived age', by serial 7 test score and gender.

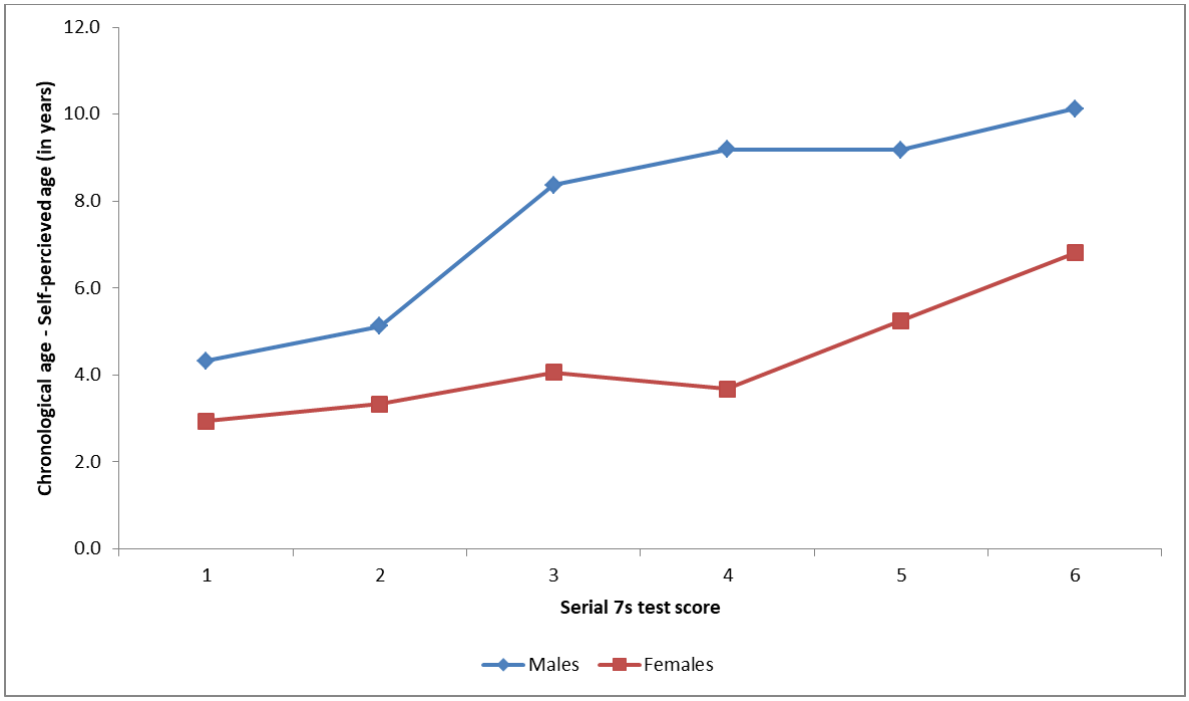
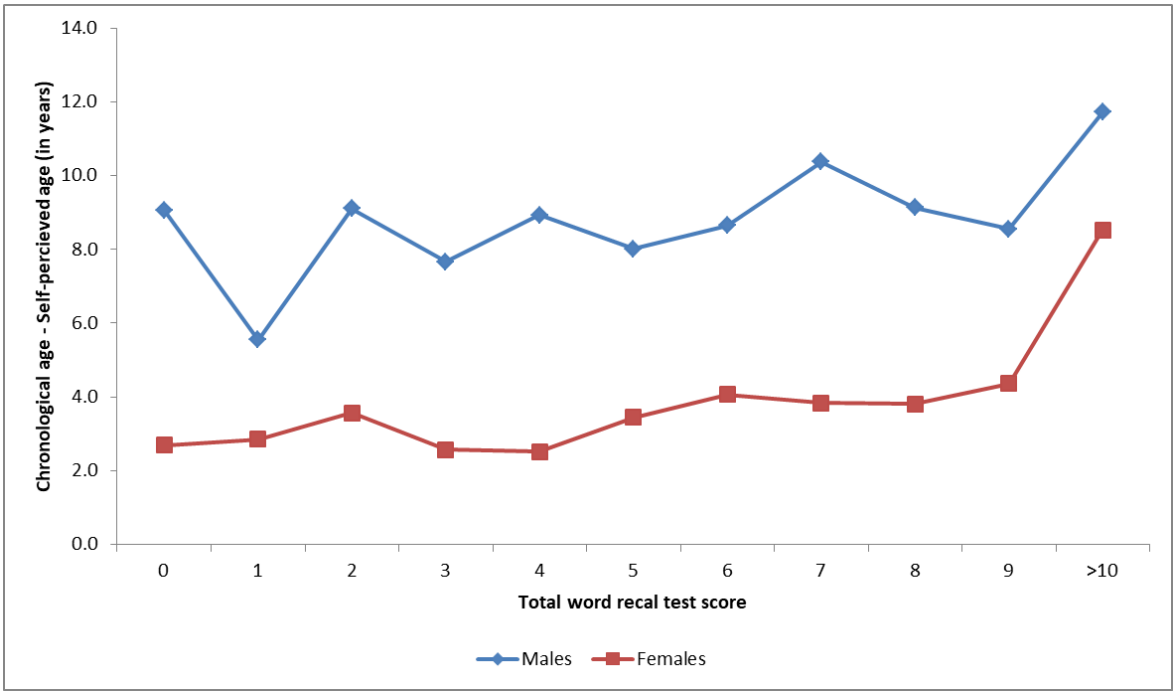


Figure 5.13 Variation of the difference between Chronological Age and 'Self-perceived age', by total word recall test score and gender.



It is worth noting that as cognitive function improves the standard deviation of the difference between chronological and self-perceived age deceases for both genders. This is a result of the decreasing standard deviation of SSPs (Table 5.8).

Table 5.8 Correlation of 'Self-perceived age' with cognitive function variables.

Serial 7s score	St. deviation of the difference in Chronological age and Self-perceived age (in years)		St. deviation of SSPs	
	Males	Females	Males	Females
0	33.1	33.0	36%	37%
1	28.4	28.9	33%	33%
2	28.9	28.2	33%	33%
3	27.4	26.1	32%	31%
4	26.7	26.2	31%	31%
5	25.1	26.0	29%	28%

Lifestyle & Behavioral risk factors

This group includes the frequency of vigorous physical activities within a week (1 = every day to 5 = never) and whether the respondent ever smoked daily. European males do vigorous activities more frequently compared to European females and Americans (Table 459). Regarding smoking habits, American males and females smoked daily more frequently compared to their European counterparts.

Table 5.9 Descriptive measures of Lifestyle & Behavioral risk variables.

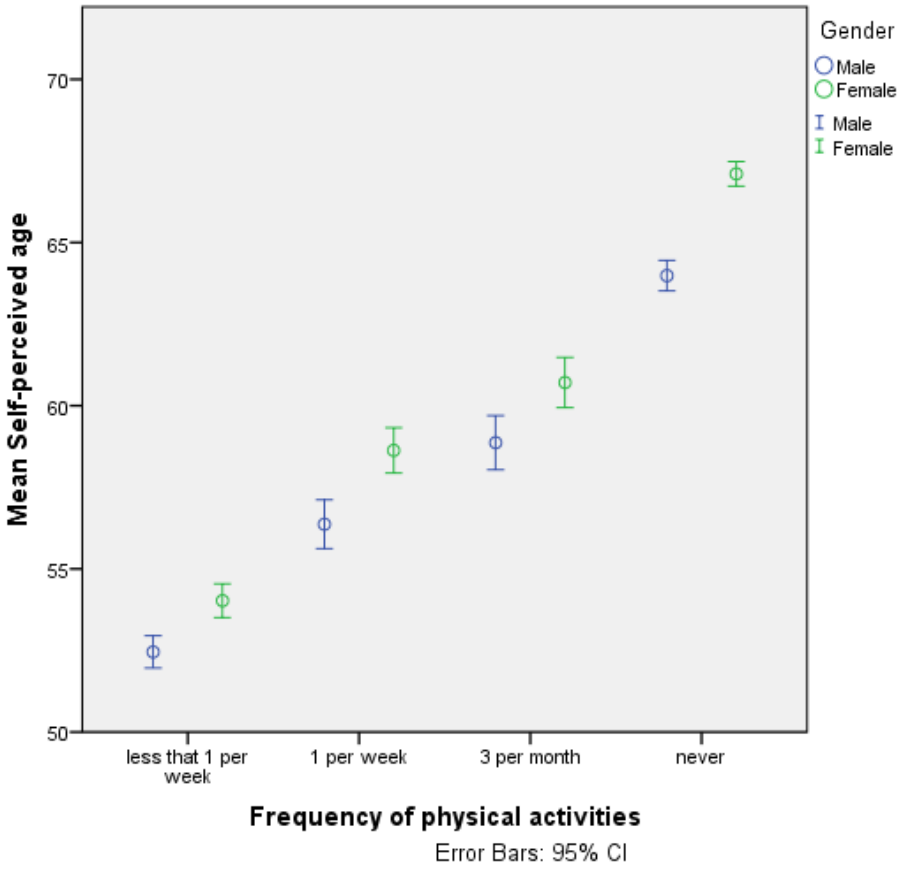
Lifestyle and behavioral risk factors	European males	European females	American males	American females
Frequency of vigorous activities (mean [SD])	3.46 [1.3]	3.67 [1.31]	3.70 [1.305]	4.07 [1.2]
Ever smoked daily	60%	36.2%	65.6%	49.8%

Smoking is a differentiating factor of 'Self-perceived age'. Male smokers have older 'Self-perceived age' than male non-smokers (t-test, t= -6.469). Furthermore, the difference of chronological and self-perceived age is larger for smokers (Table 5.10). Less frequent physical exercise is associated with older 'Self-perceived age' (Figure 5.14).

Table 5.10 Average difference of chronological and self-perceived age by smoking status (in years).

Average gap of chronological and self-perceived age		
Ever smoked daily?	Males	Females
Yes	10.7	6.0
No	8.7	4.9

Figure 5.14 Variation of 'Self-perceived age' by frequency of physical activities and gender.



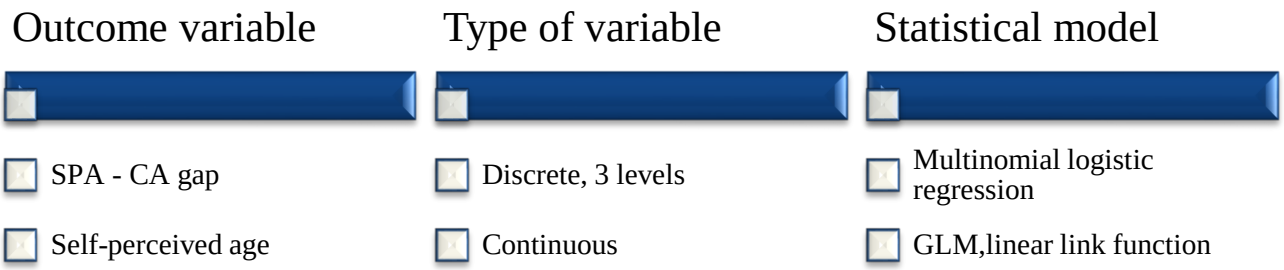
Control variables

The aim of control variables is to isolate the effects of explanatory variables of interest on the dependent variable and assist the research conclusions. The results of control variables are not discussed in detail in this study. However the main data preparation steps are outlined. The following variables will be included in the models as control variables:

- Chronological Age and Chronological Age squared /100
- The number of children of the respondent
- Country of residence for Europeans
- Body Mass Index (BMI)
- Prediction interval (in years)

5.3. Statistical modelling

In the analysis multinomial and generalised linear models (GLMs) are used, according to the type of outcome variable. SPSS 21 is used for the analyses.



Multinomial regression models with a tolerance level of 3, 5 and 7 years, are used to investigate the impact of explanatory variables on the age gap between 'Self-perceived' and Chronological Age. A different model has been estimated for every combination of region and gender (i.e. European males, European females, American males and American females). This would reveal differences as well as similarities across genders and regions. In total, twelve different models were run to cover all combinations and to assess the sensitivity of Relative Risk Ratios to different levels of tolerance. Generalised Linear Models with a linear link function as well as ordinary least squares regressions are used to investigate the impact of explanatory variables on 'Self-perceived age'. Different models are estimated for the total combined sample, Europe and USA in order to assess the robustness of our estimates.

Boundary values of SSPs

Respondents who report SSP = 100% or SSP = 0% are treated as special cases. In particular, additional assumptions are required for the calculation of 'Self-perceived age'. Two alternative approaches are investigated in this study.

— *Approach #1:* If SSP = 100% then 'Self-perceived age' is set to 0 years; If SSP = 0% then 'Self-perceived age' is set to 110 (the oldest age in the life tables) and the associated OSP is 0%. This approach is simple and easy to calculate but it does not take into account the socio-demographic characteristics of the respondents.

— *Approach #2:* The boundary values of SSPs are treated as missing values and they are estimated from a statistical model. This approach takes into account the socio-demographic characteristics of the respondents but, at the same time, it reduces the variation of SSPs.

5.4.Results

The results below (Tables 5.11 – 5.15) are estimated using Multinomial Logistic regression while the boundary values for SSPs are set according to Approach #1.

Table 5.11 Relative Risk Ratios (RRRs) based on Multinomial Logistic regression. Comparison of 'Self-perceived age' and Chronological Age. Tolerance level is set to 3 years.

'Self-perceived age' younger than 3 years from Chronological Age vs 'Self-perceived age' up to 3 years different from Chronological Age	**European males**	**European females**	**American males**	**American females**
Socio-demographic characteristics				
Number of parents alive	0.045	0.000	0.045	0.129*
Marital status (reference : Married)				
Never married	0.051	0.065	0.121	0.196
Widowed	0.096	-0.017	0.11	0.025
Divorced	0.150	0.114	-0.001	0.164
Education (reference = Tertiary)				
Less than upper secondary	0.164**	0.030	0.019	0.060
Upper secondary	0.134**	0.051	-0.017	-0.148*
Physical Health				
ADLs	0.992	1.003	1.037	1.074
Mobility Index	0.928*	0.929**	1.022	0.991
Self-rated health	0.755**	0.819**	0.746**	0.758**
Cognitive function				
Serial 7s test	0.943**	0.964*	0.920**	0.922**
Word Recall score	1.009	1.006	0.998	1.001
Lifestyle & Behavioural Risk Factors				
Frequency of Vigorous activities	0.956**	0.967*	0.966	0.956
Never smoked daily (reference: smoked daily)	1.012	1.002	0.970	1.103
'Self-perceived age' older than 3 years from Chronological Age vs 'Self-perceived age' up to 3 years different from Chronological Age	**European males**	**European females**	**American males**	**American females**
Socio-demographic characteristics				
Number of parents alive	-0.086*	-0.134**	-0.225**	-0.154**
Marital status (reference : Married)				
Never married	0.033	-0.108	-0.207	-0.018
Widowed	0.234*	-0.026	0.067	0.011
Divorced	0.098	-0.124	-0.127	-0.016
Education (reference = Tertiary)				
Less than upper secondary	0.085	-0.012	0.129	0.617**
Upper secondary	0.107*	0.053	0.059	0.126
Physical Health				
ADLs	1.074	1.087**	0.975	1.011
Mobility Index	1.087**	1.101**	1.111**	1.078**
Self-rated health	1.325**	1.333**	1.480**	1.385**
Cognitive function				
Serial 7s test	0.969	1.000	0.988	0.985
Word Recall score	0.975**	0.972**	0.986	0.994
Lifestyle & Behavioural Risk Factors				
Frequency of Vigorous activities	1.036*	1.041**	1.040	1.050
Never smoked daily (reference: smoked daily)	0.869**	0.949	0.908	1.066

The dependent variable is an unordered categorical variable reflecting the difference in years between 'self-perceived age' and chronological age. * p<5% . ** p<1%. Controlling for: country of residence; chronological age; chronological age-squared/100; BMI; prediction interval (in years); wealth; income and educational level.

Table 5.12 Relative Risk Ratios (RRRs) based on Multinomial Logistic regression. Comparison of 'Self-perceived age' and Chronological Age. Tolerance level is set to 5 years.

'Self-perceived age' younger than 5 years from Chronological Age vs 'Self-perceived age' up to 5 years different from Chronological Age	**European males**	**European females**	**American males**	**American females**
Socio-demographic characteristics				
Number of parents alive	0.047	-0.031	0.074	0.014**
Marital status (reference : Married)				
Never married	-0.111	0.081	0.18	0.204
Widowed	-0.065	0.024	0.098	0.012
Divorced	0.169*	0.152*	0.011	0.178*
Education (reference = Tertiary)				
Less than upper secondary	0.183**	0.074	0.013	0.027
Upper secondary	0.131**	0.066	0.007	-0.071
Physical Health				
ADLs	1.007	1.020	1.051	1.094*
Mobility Index	0.939*	0.931**	1.009	0.988
Self-rated health	0.746**	0.794**	0.685**	0.736**
Cognitive function				
Serial 7s test	0.962*	0.953**	0.916**	0.898**
Word Recall score	1.014**	1.008	0.994	0.994
Lifestyle & Behavioural Risk Factors				
Frequency of Vigorous activities	0.957**	0.944**	0.921**	0.958
Never smoked daily (reference: smoked daily)	1.030	0.984	1.110	1.057
'Self-perceived age' older than 5 years from Chronological Age vs 'Self-perceived age' up to 5 years different from Chronological Age	**European males**	**European females**	**American males**	**American females**
Socio-demographic characteristics				
Number of parents alive	-0.099**	-0.191**	-0.245**	-0.161**
Marital status (reference : Married)				
Never married	-0.064	-0.127	-0.075	-0.145
Widowed	0.117	0.081	0.031	-0.032
Divorced	0.112	-0.057	-0.093	-0.041
Education (reference = Tertiary)				
Less than upper secondary	0.053	0.038	0.192	0.637**
Upper secondary	0.052	0.066	0.107	0.204**
Physical Health				
ADLs	1.105**	1.104**	0.966	1.033
Mobility Index	1.107**	1.138**	1.117**	1.077**
Self-rated health	1.42**	1.377**	1.493**	1.402**
Cognitive function				
Serial 7s test	0.978	0.987	0.986	0.966
Word Recall score	0.975**	0.968**	0.979	0.983
Lifestyle & Behavioural Risk Factors				
Frequency of Vigorous activities	1.044**	1.024	1.007	1.073**
Never smoked daily (reference: smoked daily)	0.884**	0.895**	0.955	1.009

The dependent variable is an unordered categorical variable reflecting the difference in years between 'self-perceived age' and chronological age. * p<5% . ** p<1%. Controlling for: country of residence; chronological age; chronological age-squared/100; BMI; prediction interval (in years); wealth; income.

Table 5.13 Relative Risk Ratios (RRRs) based on Multinomial Logistic regression. Comparison of 'Self-perceived age' and Chronological Age. Tolerance level is set to 7 years.

'Self-perceived age' younger than 7 years from Chronological Age vs 'Self-perceived age' up to 7 years different from Chronological Age	**European males**	**European females**	**American males**	**American females**
Socio-demographic characteristics				
Number of parents alive	0.048	-0.008	0.038	0.128*
Marital status (reference : Married)				
Never married	-0.067	0.145	0.159	0.175
Widowed	-0.065	0.012	0.096	0.062
Divorced	0.179**	0.183**	0.048	0.164*
Education (reference = Tertiary)				
Less than upper secondary	0.187**	0.087	-0.011	-0.010
Upper secondary	0.139**	0.078	-0.029	-0.076
Physical Health				
ADLs	1.031	0.968	1.086	1.106
Mobility Index	0.929**	0.934**	0.991	0.984
Self-rated health	0.725**	0.787**	0.671**	0.731**
Cognitive function				
Serial 7s test	0.948**	0.957**	0.898**	0.87**
Word Recall score	1.014**	1.012*	0.996	0.994
Lifestyle & Behavioural Risk Factors				
Frequency of Vigorous activities	0.956**	0.955**	0.925**	0.962
Never smoked daily (reference: smoked daily)	1.013	0.971	1.141*	1.007
'Self-perceived age' older than 7 years from Chronological Age vs 'Self-perceived age' up to 7 years different from Chronological Age	**European males**	**European females**	**American males**	**American females**
Socio-demographic characteristics				
Number of parents alive	-0.124**	-0.209**	-0.246**	-0.167**
Marital status (reference : Married)				
Never married	-0.002	-0.088	-0.066	-0.263*
Widowed	0.109	0.088	0.109	-0.081
Divorced	0.137	-0.013	-0.074	-0.098
Education (reference = Tertiary)				
Less than upper secondary	-0.006	0.052	0.264*	0.633**
Upper secondary	0.036	0.070	0.116	0.203**
Physical Health				
ADLs	1.105**	1.050	1.013	1.045
Mobility Index	1.14**	1.164**	1.073*	1.079**
Self-rated health	1.461**	1.413**	1.59**	1.433**
Cognitive function				
Serial 7s test	0.952	0.983	0.971	0.947
Word Recall score	0.974**	0.970**	0.977*	0.981*
Lifestyle & Behavioural Risk Factors				
Frequency of Vigorous activities	1.041*	1.052**	1.027	1.080**
Never smoked daily (reference: smoked daily)	0.837**	0.859**	1.009	0.994

The dependent variable is an unordered categorical variable reflecting the difference in years between 'self-perceived age' and chronological age. * p<5% . ** p<1%. Controlling for: country of residence; chronological age; chronological age-squared/100; BMI; prediction interval (in years); wealth; income.

The following results are estimated using GLMs but the boundary values for SSPs are set according to Approach #2.

Table 5.14 Coefficients based on generalised linear models. The dependent variable is 'self-perceived age' in years and its estimation is based on 'fitted SSPs'.

Independent variables	European males	European females	American males	American females
Intercept	47.667**	54.737**	44.304**	45.002**
Socio-demographic characteristics				
Number of parents alive (reference: Both parents alive)				
No parent alive	1.890**	1.566**	3.283**	2.476**
One of the parents alive	0.825**	0.732**	2.016**	1.128*
Marital status (reference : Married)				
Never married	-0.250*	-0.464**	-0.686	-1.077*
Widowed	0.335**	0.329**	0.036	-0.263
Divorced	-0.686**	-0.620**	-0.496	-0.982**
Education (reference: Tertiary)				
Less than upper secondary	0.066	0.065	0.400	2.158**
Upper secondary	-0.076	0.015	-0.142	0.362
Income quartiles (reference : Q4)				
Income Q1	1.378**	1.428**	3.123**	1.866**
Income Q2	0.769**	0.766**	1.903**	1.417**
Income Q3	0.668**	0.563**	0.417	0.193
Physical Health				
ADLs	0.528**	0.490**	0.081	0.388**
Mobility Index	0.678**	0.688**	1.026**	1.025**
Self-rated health (reference : Poor)				
Excellent	-13.138**	-12.193**	-20.750**	-16.320**
Very good	-9.869**	-8.540**	-18.980**	-14.632**
Good	-7.262**	-6.373**	-16.565**	-12.298**
Fair	-4.746**	-4.365**	-10.653**	-7.250**
Cognitive function				
Serial 7s test	0.197**	0.216**	0.162	0.175*
Word Recall score	-0.161**	-0.150**	-0.238**	-0.188**
Lifestyle & Behavioural Risk Factors				
Frequency of Vigorous activities	0.446**	0.390**	0.299**	0.415**
Never smoked daily (reference: smoked daily)	-0.366**	-0.405**	-0.255*	-0.057

* p<5% . ** p<1%. Controlling for: country of residence; chronological age; BMI; prediction interval (in years). The signs of the coefficients indicate whether a variable has positive or negative contribution to 'Self-perceived age'

Almost all predictors are significant at 1%. This is partly a result of the approach adopted for the estimation of boundary values for SSPs. In other words, the use of statistical modelling to estimate SSPs has made the trends clear while removing the random variation of self-reported SSPs. Therefore, the calculated 'Self-perceived' age based on 'fitted SSPs' has clearly identified trends and small variation. For comparison purposes, the models are re-run using 'Self-perceived' age as a dependent variable. The estimation is now based on self-reported SSPs.

Table 5.15 Coefficients based on generalised linear models. The dependent variable is 'self-perceived age' in years and its estimation is based on 'self-reported SSPs'.

Independent variables	European males	European females	American males	American females
Intercept	56.246**	67.515**	52.551**	52.049**
Socio-demographic characteristics				
Number of parents alive (reference: Both parents alive)				
No parent alive	1.418**	1.493**	2.219**	2.260**
One of the parents alive	0.761*	0.551*	1.200	0.956
Marital status (reference : Married)				
Never married	0.564*	-0.010	-0.403	-0.243
Widowed	0.335	0.202	-1.118*	0.083
Divorced	0.021	-0.260	-0.544	-0.521
Education (reference: Tertiary)				
Less than upper secondary	-0.247	0.591**	0.584	2.150**
Upper secondary	0.164	0.640**	-0.203	1.073**
Income quartiles (reference : Q4)				
Income Q1	0.742**	1.070**	0.505	0.258
Income Q2	0.590**	0.658**	0.148	0.550
Income Q3	0.215	0.546**	0.083	0.681*
Physical Health				
ADLs	0.173	0.025	-0.302	-0.067
Mobility Index	0.575**	0.522**	0.286*	0.309**
Self-rated health (reference : Poor)				
Excellent	-8.996**	-6.359**	-8.926**	-6.518**
Very good	-7.692**	-5.501**	-5.680**	-4.678**
Good	-5.473**	-3.769**	-3.076**	-2.359**
Fair	-2.285**	-2.160**	-0.732	-0.849
Cognitive function				
Serial 7s test	0.128	0.007	0.348**	0.145
Word Recall score	-0.161**	-0.108**	-0.012	-0.017
Lifestyle & Behavioural Risk Factors				
Frequency of Vigorous activities	0.251**	0.187**	0.166	0.257**
Never smoked daily (reference: smoked daily)	-0.844**	-0.451**	-0.419	-0.303

The dependent variable is 'self-perceived age' in years and its estimation is based on 'self-reported SSPs'. Respondents who reported SSP = 0% or SSP = 100% are excluded from the analyses. * p<5% . ** p<1%. Controlling for: country of residence; chronological age; BMI; prediction interval (in years). The signs of the coefficients indicate whether a variable has positive or negative contribution to 'Self-perceived age'.

Synopsis of the results

I summarise the main outcomes from all models by group of variables.

Socio-demographic characteristics

Parental mortality is a key differentiation of 'Self-perceived age'. If parents are alive then 'Self-perceived age' is younger and vice versa. The results of the multinomial models suggest that the number of parents who are alive is negatively associated with older 'Self-perceived age' (RRRs vary from -0.086 to -0.124 for European males). American males have the strongest association (RRRs vary from -0.225 to -0.246) whereas European males the weakest. Moreover, the results of linear models

suggest that the association is stronger for Americans (coefficients vary from 0.956 to 3.283) than for Europeans (coefficients vary from 0.551 to 1.890).

Physical Health

European males and females who report more ADLs tend to have a 'Self-perceived' age older than their Chronological Age; Relative Risk Ratios (RRRs) indicate that an additional limitation increases chances by approximately 10%. The association of ADLs with 'Self-perceived' age is less clear for Americans and depends on the tolerance level. For example, for a tolerance level of 3 or 5 years, the RRRs indicate that for American females an additional limitation increases chances of having older 'Self-perceived' age by 1.1% to 3.3% while for American males the findings are non-significant but indicate higher chances of having a younger 'Self-perceived' age. On the other hand, for a tolerance level of 7 years, both Europeans and Americans tend to have a 'Self-perceived' age older than their Chronological Age and the findings are consistent for both genders.

As the number of mobility difficulties increases, individuals tend to have a 'Self-perceived' age older than their Chronological Age. The RRRs indicate that for European males an additional mobility difficulty increases chances of having a 'Self-perceived' age older than their Chronological Age by 8.7% - 14% while for American males relative chances increase by at least 7.3%. Mobility difficulties affect European females the most (RRR= 1.101 - 1.164) and American females the least (RRR= 1.078 - 1.079).

Poor self-rated health implies an older 'Self-perceived' age compared to Chronological Age. Associations are strong and significant in all instances. The RRRs indicate that for European males, a one-point deterioration in the 5-point scale of self-rated health increases chances of having older 'Self-perceived age' by 32.5% to 46.1% while the relative increase for American males is 48% to 59%. Similar results hold for females.

Cognitive function

Better cognitive skills, based on the subtraction test, imply for both Europeans and Americans a 'Self-perceived age' closer to Chronological Age. The RRRs indicate that, for European males, a one-point improvement in the serial 7s test score reduces the chances of having younger 'Self-perceived age' by approximately 5% and the chances of having older 'Self-perceived age' by approximately 3% to 5%. The impact of cognitive skills is more significant for Americans. The RRRs indicate that for American males, a one-point improvement in the serial 7s test score reduces the chances of having younger 'Self-perceived age' by approximately 8% while it reduces the chances of having older 'Self-perceived age'

by 1.2% - 2.9%. Better memory also implies that 'Self-perceived age' is closer to Chronological Age for Americans while for Europeans it indicates a younger 'Self-perceived age'.

Lifestyle & Behavioural Risk Factors

Individuals who do vigorous activities frequently have a 'Self-perceived age' younger than their Chronological Age; this holds for both genders and irrespectively of place of origin. For instance, the RRRs indicate that, for European males, as frequency of vigorous activities decreases, chances of having a 'Self-perceived age' older than their Chronological Age increase by 3.6% - 4.4%. Similarly, American females, who do vigorous activities less frequently, tend to have a 'Self-perceived age' older than their Chronological Age; relative chances increase by 5% - 8%, depending on the tolerance level.

Male nonsmokers have a 'Self-perceived age' younger than their Chronological Age, compared to male smokers. The RRRs indicate that American male non-smokers exhibit up to 14% higher chances of having younger 'Self-perceived age'. Similarly, chances for European males are up to 3.0% higher. It is worth noting that the results depend on the tolerance level and that the results for females are contradictory. On the one hand, the RRRs indicate that European female nonsmokers have a 'Self-perceived age' closer to their Chronological Age. On the other hand, American female nonsmokers tend to have a 'Self-perceived age' older than their Chronological Age.

In addition, the results based on generalised linear models (Table 4.15) show that non-smokers tend to have younger 'Self-perceived age' compared to smokers. This holds for both genders as well as for Americans and Europeans. More specifically, smoking status has a stronger negative association for European males (b= - 0.844) whereas for American females the association is weaker (b= - 0.303).

5.5. Comparison of 'Self-perceived age' and mortality patterns

In this final section the patterns of 'Self-perceived age' are compared to the patterns of actual mortality. Each factor is allocated in one of the following homogeneous groups (Table 5.16).

i. Factors linked to better actual mortality and younger 'Self-perceived age'
ii. Factors linked to worse actual mortality and older 'Self-perceived age'
iii. Factors exhibiting inconsistent associations with 'Self-perceived age' and actual mortality

Table 5.16 Grouping of factors based on 'Self-perceived age' consistency to actual mortality patterns.

Factors	Factors linked to better actual mortality and younger 'Self-perceived age'	Factors linked to worse actual mortality and older 'Self-perceived age'	Factors exhibiting inconsistent associations with Self-perceived age' and actual mortality
Demographic			
Both parents are alive	✓		
No parents are alive		✓	
Marital Status: Divorced			✘
Marital Status: Widowed		✓ *(for Europeans)*	✘ *(for Americans)*
Socio-economic Status			
Higher education	✓		
Higher Income	✓		
Physical Health			
Increase in ADLs		✓	
Mobility Difficulties		✓	
Deterioration of Self-rated health		✓	
Cognitive function			
Better Serial 7s test			✘
Better Word Recall score	✓ *(for Europeans*		✘ *(for Americans)*
Lifestyle & Behavioural Risk Factors			
Doing vigorous activities frequently	✓		
Never smoked	✓ *(for Europeans)*		✘ *(for Americans)*

The factors included in the first homogeneous group are associated with better actual mortality and younger ‘Self-perceived age’. These factors are:

- Better parental longevity
- Higher income and educational attainment
- Better memory and non-smoking habit for Europeans
- Doing vigorous physical exercise

The factors included in the second homogeneous group are associated with worse actual mortality and older ‘Self-perceived age’. These factors are:

- Increase in the number of chronic conditions and mobility difficulties
- Poor self-rated health
- Widowhood for Europeans

The factors included in the third homogeneous group exhibit inconsistent associations with ‘Self-perceived age’ and actual mortality patterns. These factors are:

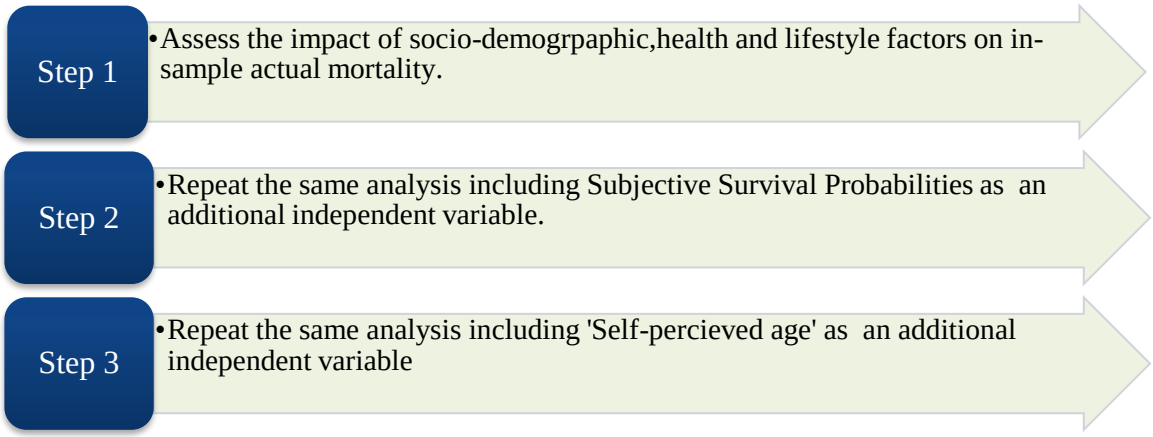
- Smoking status and better memory for Americans
- Better cognitive function
- Divorced
- Widowhood for Americans

The results are discussed in more detail in Section 7.

6. Chapter 6 – Subjective survival and in-sample actual survival

6.1. Introduction

One of the most interesting and challenging topics on subjective survival is the examination of the accuracy of subjective survival expectations. A closely related research question is whether subjective survival expectations include important information in addition to the factors known to affect survival. To achieve the aforementioned objective, the following process is adopted.



To achieve the objective, data from the SHARE wave 6 and wave 7 were used. In particular, the data for respondents who participated in SHARE Wave 6 and died afterwards were collected from Wave 7. The data and statistical modelling are described in the following sections. All data sources are also described in Section 2.5.

6.2.Methods and sample description

6.2.1. Force of subjective mortality

The main challenge that researchers face assessing the predictive power of subjective survival probabilities on mortality is the time interval of the prediction. For example, if the survival prediction refers to the next 10 years researchers have to wait 10 years to confirm the accuracy of the prediction. In the 'Expectations' module of the SHARE questionnaire respondents were asked to state their survival expectations on a scale from 0 to 100 as follows:

What are the chances that you will live to be age [T] or more?

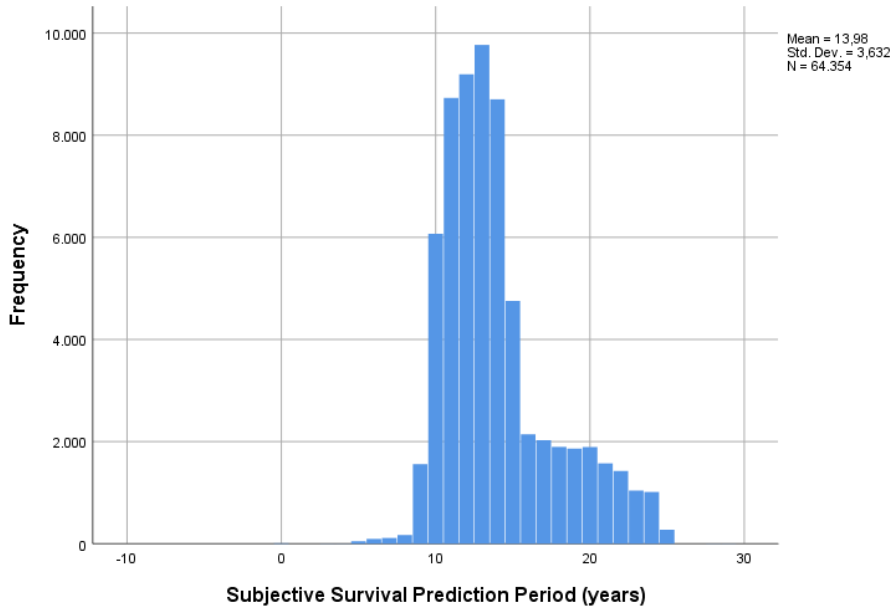
The target age T depends on the age of the respondent at the interview and is set as follows.

Table 6.1 Chronological age and Target age.

Age band	Target age
50 – 65	75
66 – 70	80
71 – 75	85
76 – 80	90
81 – 85	95
86 – 95	100
96 - 100	105
101 +	110

The Time horizon of Subjective Survival Probabilities is the difference of Target age and Chronological age at the time of the interview.

Figure 6.1 Time horizon of Subjective Survival Probabilities.



On average, SHARE Wave 6 respondents report subjective survival probabilities which refer to the next 14 years (Figure 6.1). On the other hand, SHARE Wave 7 data, including End of Life interviews which were collected in 2017, two years after SHARE Wave 6. So, there is a mismatch between the time interval of SHARE waves (about 2 years) and the time interval which corresponds to subjective survival probabilities (on average 14 years). The subjective survival probabilities which correspond to the next two years are estimated using the 'Force of subjective mortality'.

Force of mortality

In Actuarial Science, 'Force of mortality' represents the 'rate of mortality' among persons who have survived to that age. It describes the behaviour of a mortality rate over an infinitely small duration of time. In mathematical terms it is the conditional probability of person surviving up to age x, given that he has survived up to that age (Bowers et al. 1997).

The 'Force of mortality' is defined as follows.

$$\mu(x) = -\frac{S'(x)}{S(x)}$$

Where,

S(x) is a survival function.

$-S'(x)$ is a survival probability density function.

Furthermore, the probability of person surviving up to age x, to die between ages x and x+u can be calculated as follows.

$$S_x(u) = e^{-\int_x^{x+u} \mu(y)dy} \quad \text{(Equation 6.1)}$$

Where,

$S_x(.)$ is a survival function

$\mu(y)$ is the 'Force of mortality'

u is the time horizon of the survival probability

Assume a constant 'Force of mortality' $\mu(x) = \lambda$; then the survival probability becomes:

$$S_x(u) = e^{-\lambda u} \quad \text{(Equation 6.2)}$$

Force of subjective mortality

As explained earlier, the aim is to estimate the subjective survival probability which refers to the next two years for all SHARE Wave 6 respondents. To do this, I introduce the new term 'Force of subjective mortality'. The 'Force of subjective mortality' is assumed to be constant and can be calculated from the self-reported subjective survival probabilities as follows.

$$SSP_x^u = e^{-\lambda^s u} \quad \text{(Equation 6.3)}$$

Where,

SSP_x^u is a subjective survival probability of a respondent aged x, who reports the chance of own survival for the next 'u' years (see Table 6.1)

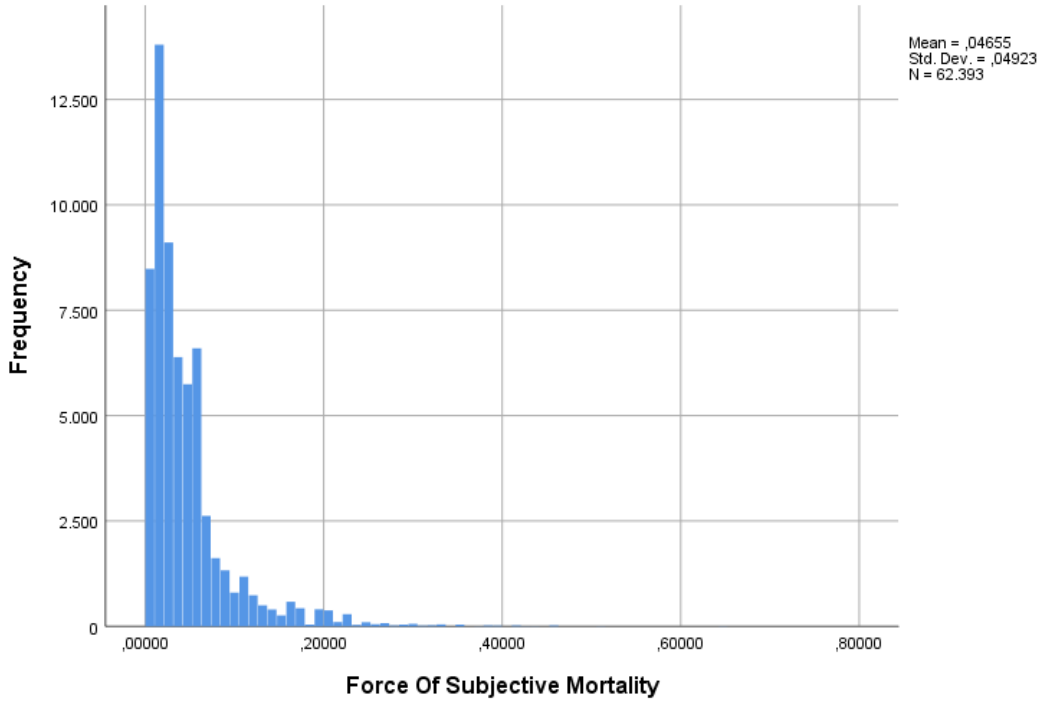
λ^s is the 'Force of subjective mortality' and it is assumed to be constant across all years of age.

Two special cases have to be dealt with, namely, SSP=0% and SSP=100%. To set these boundary values, I follow the 'Approach #2', which has been used for the calculation of 'Self-perceived age' (see Section 5.3). In other words, the boundary values of SSPs are treated as missing values and they are estimated based on a statistical model.

The 'Force of subjective mortality' is estimated by Equation 6.4 below, whereas the subjective survival probability which corresponds to the next two years is estimated using Equation 6.3.

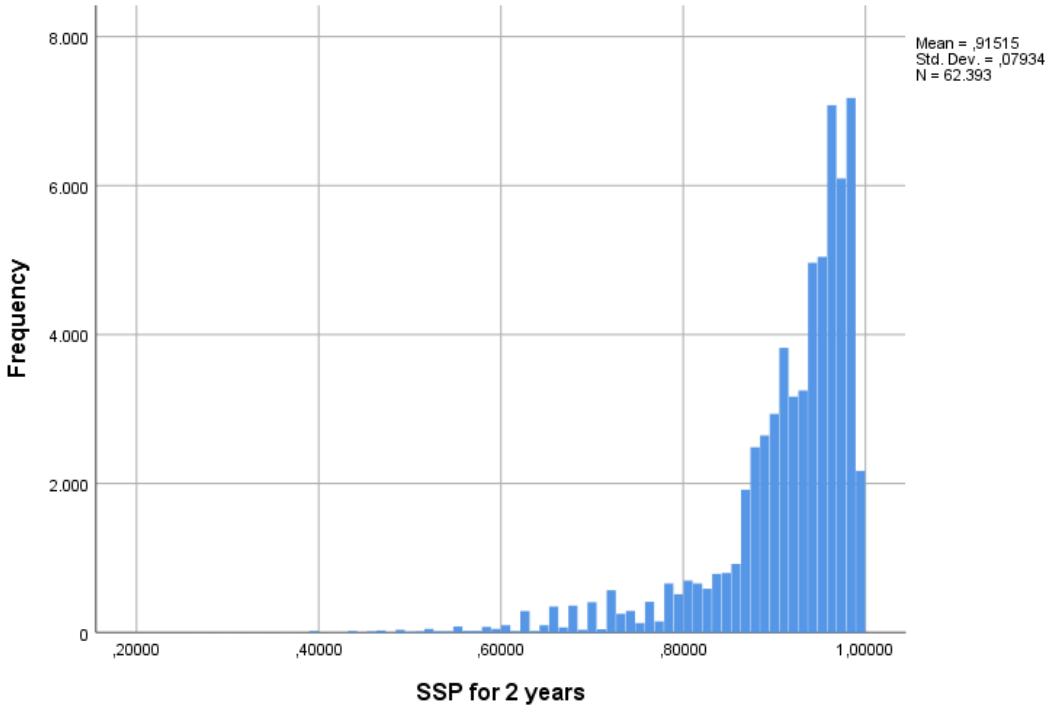
$$\lambda^s = -\frac{\ln(SSP_x^u)}{u} \quad \text{(Equation 6.4)}$$

Figure 6.2 Distribution of 'Force of subjective mortality'.



The distribution of 'Force of subjective mortality' is concentrated on 4.6% but there are large extreme values in the sample. Small subjective survival probabilities imply large values of 'Force of subjective mortality' and vice versa.

Figure 6.3 Distribution of Subjective survival probabilities for the next 2 years.



The distribution of SHARE Wave 6 subjective survival probabilities with a two-year time horizon is concentrated on 91% but there are extreme small values in the sample. Subjective survival probabilities decrease in value as time horizon increases.

The distribution of subjective survival probabilities as reported and the distribution of subjective survival probabilities with a two-year time horizon are compared in Table 6.2.

Table 6.2 Distributions of SHARE Wave 6 SSPs as reported and SSPs (2 years).

Percentile	SSPs as reported	SSP (2 years)
5%	15.0%	74.6%
10%	30.0%	81.8%
15%	40.0%	85.8%
20%	50.0%	88.2%
25%	50.0%	89.1%
30%	50.0%	90.0%
35%	50.0%	91.1%
40%	55.3%	92.2%
45%	60.0%	93.1%
50%	66.0%	94.0%
55%	70.0%	94.7%
60%	70.0%	95.4%
65%	75.0%	96.0%
70%	80.0%	96.3%
75%	80.0%	96.8%
80%	80.0%	97.2%
85%	83.9%	97.9%
90%	90.0%	98.3%
95%	90.0%	98.7%
Mean	61.6%	91.6%
St. deviation	22.4%	7.9%

As expected, SHARE Wave 6 subjective survival probabilities decrease in value as time horizon increases. The average SSP is reposted as 61.6% whereas the average SSP which corresponds to the next 2 years is 91.6%. The standard deviation of SSPs with a two-year time horizon is also much lower (7.9% vs 22.4%).

6.2.2. Dependent and explanatory variables

Dependent variable

A binary variable indicating the death of a respondent is defined as follows.

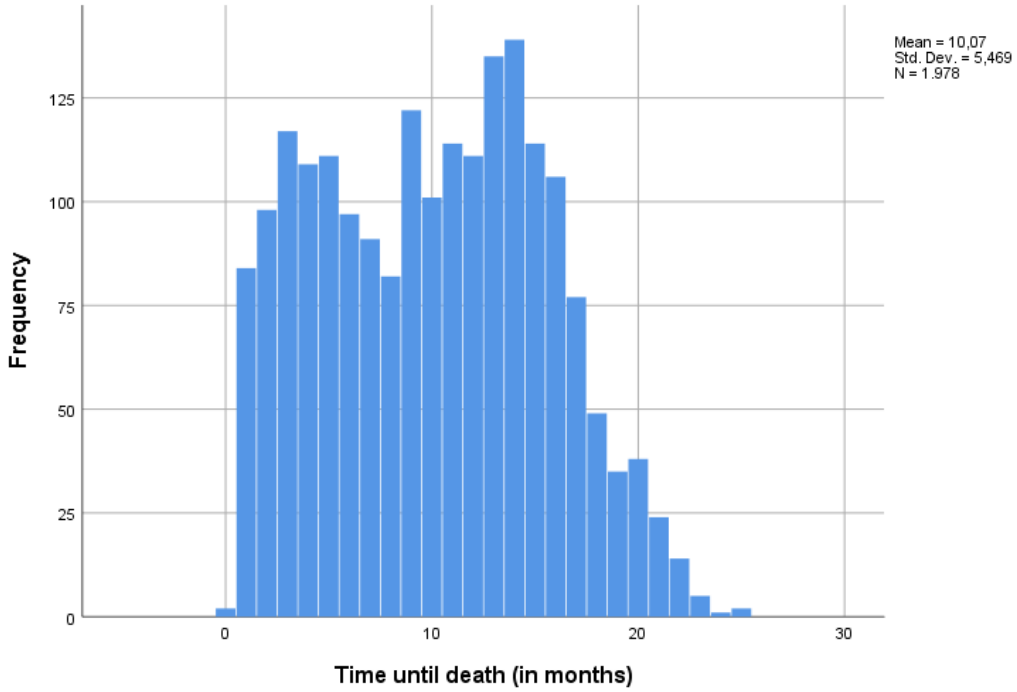
$$In-sample\ mortality\ indicator = \begin{cases} 1; if\ a\ respondent\ participated\ in\ Wave\ 6\ and\ then\ died \\ 0; otherwise \end{cases}$$

Table 6.3 Status of SHARE W6 respondents, aged 50 or older, at SHARE W7.

Status of respondent in W7	Count
Respondents who participated in W6, including W7 non-respondents	65,368
Respondents who participated in W6 but died before W7	1,978
Total respondents, aged 50 or older	67,346

Respondents who participated in W6, including W7 non-respondents, are treated as censored observations in the statistical models. The SHARE W6 in-sample crude mortality rate is 3%.

Figure 6.4 Distribution of time until death for SHARE W6 respondents who had died by Wave 7 (in months).



The average time until death for SHARE W6 respondents who had died by Wave 7 is about 10 months.

Explanatory variables

Demographic characteristics

This group includes chronological age (in years), gender, and the number of children of the respondent as well as country of residence as a control variable.

Table 6.4 Descriptive statistics of demographic variables.

Demographic characteristics	Respondents who participated in W6 , including W7 non-respondents	Respondents who participated in W6 but died before W7
Age (mean)	67.7	75.4
Males	96.7%	3.3%
Females	97.4%	2.6%
Number of children (mean)	2.12	2.10

Respondents who participated in W6 but died before W7 are on average older and have fewer children. On a relative basis, males have higher mortality than females.

Socioeconomic status

Socioeconomic status is represented by the "equivalised" individual income in quartiles and educational attainment, considered in 4 levels, based on the ISCED-97 classification, Primary (code 1), Lower Secondary (code 2), Upper Secondary (codes 3 & 4) and Tertiary (codes 5 & 6).

Table 6.5 Descriptive statistics of socioeconomic variables.

Socioeconomic variables		Respondents who participated in W6, including W7 non-respondents	Respondents who participated in W6 but died before W7
Equivalised Income	1st Quartile (mean in €)	€ 4,691	€ 4,906
	2nd Quartile (mean in €)	€ 12,937	€ 12,784
	3rd Quartile (mean in €)	€ 25,477	€ 25,013
	4th Quartile (mean in €)	€ 74,166	€ 65,648
Education level	ISCED-97 code 0 & 1 (Primary)	95.7%	4.3%
	ISCED-97 code 2 (Lower secondary)	96.5%	3.5%
	ISCED-97 codes 3 & 4 (Upper secondary)	97.6%	2.4%
	ISCED-97 codes 5 & 6 (Tertiary)	98.2%	1.8%

Respondents who participated in W6 but died before W7 have on average lower income and educational attainment.

Physical Health & Cognitive function

Physical health group includes the number of limitations in Activities of Daily Living (out of a list of 6 basic, everyday tasks) and self-rated health (ranging from 1=excellent to 5=poor). Furthermore, cognitive function includes the score of a numeracy test (1=bad to 5=good) and the score of orientation in time test (0=bad to 4=good).

Table 6.6 Descriptive statistics of Physical Health & Cognitive function variables.

Physical Health & Cognitive function		**Respondents who participated in W6, including W7 non-respondents**	**Respondents who participated in W6 but died before W7**
Number of ADLs (mean)		0.9	1.7
Self-rated health	Excellent	98.9%	1.1%
	Very good	98.4%	1.6%
	Good	97.9%	2.1%
	Fair	96.4%	3.6%
	Poor	93.3%	6.7%
Numeracy test (mean)		1.03	1.04
Orientation in time (mean)		3.8	3.6

Respondents who participated in W6 but died before W7 have on average more chronic conditions and ADLs, worse self-rated health and lower orientation in time score. In contrast, they have marginally better score on numeracy test.

Lifestyle & Behavioural risk factors

This set of variables includes the BMI in four categories (underweight, normal weight, overweight and obese), whether the respondent does vigorous or moderate physical activities, and whether the respondent ever smoked daily. Life satisfaction is also included in the analysis.

Table 6.7 Descriptive statistics of Lifestyle & Behavioural risk factors

Lifestyle & Behavioural risk factors		**Respondents who participated in W6 , including W7 non-respondents**	**Respondents who participated in W6 but died before W7**
Do vigorous or moderate physical activities		97.8%	2.2%
Physically inactive		92.4%	7.6%
Smoked daily		97.1%	2.9%
Non-smoker		97.2%	2.8%
BMI	Underwieght	93.5%	6.5%
	Overweight	97.4%	2.6%
	Obese	97.5%	2.5%
	Normal	96.5%	3.5%
Life satisfaction (mean)		7.6	7.0

Respondents who participated in W6 but died before W7 tend to be physically inactive as well as smokers. In addition, they tend to be less satisfied with their lives. They also include a higher proportion of underweight individuals; underweight respondents tend to have relatively higher mortality compared to normal, overweight and obese.

Subjective Survival Probabilities and 'Self-perceived age'

Subjective Survival Probabilities with a time horizon of two years and 'Self-perceived age' will be included in the statistical models as independent variables. Subjective Survival Probabilities are defined in Section 2.5.2. 'Self-perceived age' is defined, calculated, and explained in Chapter 5.

Table 6.8 Descriptive statistics of Subjective Survival variables.

Subjective Survival	Respondents who participated in W6, including W7 non-respondents	Respondents who participated in W6 but died before W7
SSP 2-years (mean)	92%	87%
'Self-perceived age' (mean)	56 years	63 years

Respondents who participated in W6 but died before W7 have lower Subjective Survival Probabilities and older 'Self-perceived age'.

6.3. Statistical modelling

In the analysis we use Binary Logistic regression and Cox regression models. A binary dependent variable indicates whether a respondent is alive or not at the time of Wave 7 data collection while time to death (in months) is used as an additional input in Cox regression models. In total three sets of models are estimated.

i. The first set of models is used to estimate the impact of the explanatory variables on the actual mortality experience of SHARE Wave 6 respondents.

ii. The second set of models is used to estimate the impact of the explanatory variables as well as subjective survival probabilities with a time horizon of two years on the actual mortality experience of SHARE Wave 6 respondents. The value and sign of the regression coefficient would reveal whether subjective survival probabilities contain 'supplementary survival information', in addition to variables known to predict mortality.

iii. The third set of models is used to estimate the impact of the explanatory variables as well as 'Self-perceived age' on the actual mortality experience of SHARE Wave 6 respondents. The value and sign of the regression coefficient would reveal whether 'Self-perceived age' contains 'supplementary survival information' in addition to variables known to predict mortality.

It is worth mentioning that one of the advantages of 'Self-perceived age' is that no adjustment about the time horizon in required.

6.4.Results

This section includes the results of the statistical models as well as their interpretation. Table 6.9 presents the results of the models including known mortality predictors (actual mortality model); Table 6.10 presents the results of the models that further include subjective survival probabilities (survival probabilities model); and Table 6.11 presents the results of the models including known mortality predictors as well as 'Self-perceived age' (self-perceived age model).

Table 6.9 Odds and hazard ratios of Binary logistic and Cox proportional hazard models. Actual mortality model (including known mortality predictors).

Predictor	Binary logistic	Cox proportional hazard
	OR (95% CI)	Hazard ratios (95% CI)
Demographic Characteristics		
Chronological Age	1.054** (1.049 – 1.060)	1.053** (1.047 – 1.058)
Male (reference: Female)	1.457** (1.318 – 1.611)	1.440** (1.307 – 1.587)
Number of children	0.981 (0.949 – 1.014)	0.983 (0.953 – 1.014)
Socio-economic status		
Education (reference: Tertiary)		
Primary	1.193 (0.923 – 1.294)	1.100 (0.934 – 1.296)
Lower Secondary	1.279** (1.083 – 1.510)	1.274** (1.084 – 1.497)
Upper Secondary	1.141 (0.983 – 1.326)	1.143 (0.988 – 1.322)
Equivalised Income (reference: Q4)		
Q1	2.128** (1.737 – 2.608)	2.069** (1.698 – 2.521)
Q2	1.819** (1.503 – 2.202)	1.779** (1.477 – 2.144)
Q3	1.479** (1.242 – 1.762)	1.468** (1.237 – 1.742)
Physical health & Cognitive function		
ADLs	1.122** (1.083 – 1.163)	1.110** (1.074 – 1.147)
Self-rated health (reference: Poor)		
Excellent	0.540** (0.396 – 0.737)	0.541** (0.398 – 0.734)
Very good	0.695** (0.569 – 0.850)	0.695** (0.572 – 0.845)
Good	0.757** (0.649 – 0.883)	0.760** (0.655 – 0.881)
Fair	0.952 (0.833 – 1.163)	0.954 (0.833 – 1.084)
Orientation in time	0.926* (0.869 – 0.987)	0.934* (0.880 – 0.992)
Numeracy	1.004 (0.954 – 1.056)	1.003 (0.955 – 1.014)
Lifestyle and behavioral risk factors		
Physical activity (reference: Physically Inactive)		
Physically Active	0.664** (0.585 – 0.754)	0.675** (0.597 – 0.763)
Ever smoked daily (reference: Yes)		
Never smoked daily	0.871** (0.788 – 0.964)	0.875** (0.794 – 0.965)
BMI (reference: Normal)		
Underweight	1.330 (0.976 – 1.813)	1.292 (0.970 – 1.722)
Overweight	0.709** (0.637 – 0.788)	0.717** (0.647 – 0.794)
Obese	0.687** (0.605 – 0.781)	0.696** (0.615 – 0.788)
Life satisfaction	0.897** (0.876 – 0.919)	0.903** (0.882 – 0.924)

***p<1%, *p<5%. Controlling for country of residence. The binary dependent variable indicates whether a respondent is alive or not at the time of Wave 7 data collection.*

Table 6.10 Odds and hazard ratios of Binary logistic and Cox proportional hazard models. Subjective survival probability model.

Predictor	Binary logistic	Cox proportional hazard
	OR (95% CI)	**Hazard ratios (95% CI)**
SSP (2-years)	0.191** (0.109 – 0.336)	0.215** (0.126 – 0.367)
Demographic Characteristics		
Chronological Age	1.054** (1.047 – 1.060)	1.052** (1.046 – 1.059)
Male (reference: Female)	1.504** (1.339 – 1.689)	1.485** (1.327 – 1.662)
Number of children	0.976 (0.939 – 1.015)	0.978 (0.941 – 1.016)
Socio-economic status		
Education (reference: Tertiary)		
Primary	1.105 (0.912 – 1.341)	1.112 (0.922 – 1.342)
Lower Secondary	1.216* (1.005 – 1.471)	1.214* (1.008 – 1.461)
Upper Secondary	1.168 (0.988 – 1.381)	1.172 (0.995 – 1.381)
Equivalised Income (reference: Q4)		
Q1	1.794** (1.419 – 2.268)	1.760** (1.400 – 2.212)
Q2	1.575** (1.265 – 1.960)	1.547** (1.248 – 1.917)
Q3	1.383** (1.137 – 1.682)	1.379** (1.138 – 1.671)
Physical health & Cognitive function		
ADLs	1.064* (1.010 – 1.120)	1.057* (1.007 – 1.111)
Self-rated health (reference: Poor)		
Excellent	0.487** (0.338 – 0.701)	0.488** (0.341 – 0.698)
Very good	0.628** (0.496 – 0.796)	0.629** (0.499 – 0.793)
Good	0.724** (0.605 – 0.867)	0.727** (0.611 – 0.866)
Fair	0.960 (0.820 – 1.124)	0.963 (0.828 – 1.121)
Orientation in time	0.903** (0.837 – 0.974)	0.912* (0.849 – 0.980)
Numeracy	0.920 (0.949 – 1.064)	1.003 (0.948 – 1.062)
Lifestyle and behavioral risk factors		
Physical activity (reference: Physically Inactive)		
Physically Active	0.638** (0.552 – 0.738)	0.653** (0.567 – 0.751)
Ever smoked daily (reference: Yes)		
Never smoked daily	0.805** (0.717 – 0.904)	0.807** (0.721 – 0.903)
BMI (reference: Normal)		
Underweight	1.104 (0.740 – 1.647)	1.077 (0.738 – 1.571)
Overweight	0.673** (0.594 – 0.761)	0.680** (0.603 – 0.767)
Obese	0.683** (0.590 – 0.711)	0.692** (0.600 – 0.797)
Life satisfaction	0.900** (0.875 – 0.926)	0.904** (0.880 – 0.929)

***p<1%, *p<5%. Controlling for country of residence. The binary dependent variable indicates whether a respondent is alive or not at the time of Wave 7 data collection. Subjective Survival Probabilities with a time horizon of two years are included as predictors of in-sample mortality.*

Table 6.11 Odds and hazard ratios of Binary logistic and Cox proportional hazard models. 'Self-perceived age' model.

Predictor	Binary logistic	Cox proportional hazard
	OR (95% CI)	**Hazard ratios (95% CI)**
'Self-perceived age'	1.020** (1.017 – 1.023)	1.020** (1.017 – 1.022)
Demographic Characteristics		
Chronological Age	1.044** (1.038 – 1.050)	1.043** (1.037 – 1.048)
Male (reference: Female)	1.532** (1.375 – 1.707)	1.508** (1.357 – 1.674)
Number of children	0.984 (0.949 – 1.020)	0.985 (0.951 – 1.020)
Socio-economic status		
Education (reference: Tertiary)		
Primary	1.083 (0.904 – 1.296)	1.088 (0.913 – 1.296)
Lower Secondary	1.232* (1.032 – 1.471)	1.230* (1.035 – 1.462)
Upper Secondary	1.181* (1.008 – 1.383)	1.183* (1.014 – 1.381)
Equivalised Income (reference: Q4)		
Q1	1.999** (1.603 – 2.492)	1.945** (1.569 – 2.412)
Q2	1.715** (1.393 – 2.110)	1.679** (1.370 – 2.058)
Q3	1.452** (1.202 – 1.755)	1.441** (1.197 – 1.735)
Physical health & Cognitive function		
ADLs	1.073** (1.021 – 1.127)	1.064** (1.016 – 1.115)
Self-rated health (reference: Poor)		
Excellent	0.664* (0.481 – 0.915)	0.667* (0.487 – 0.914)
Very good	0.790* (0.637 – 0.979)	0.792* (0.642 – 0.976)
Good	0.835* (0.705 – 0.990)	0.840* (0.713 – 0.989)
Fair	1.036 (0.890 – 1.206)	1.064 (0.897 – 1.200)
Orientation in time	0.914* (0.851 – 0.982)	0.924* (0.864 – 0.989)
Numeracy	0.998 (0.944 – 1.055)	1.000 (0.948 – 1.055)
Lifestyle and behavioral risk factors		
Physical activity (reference: Physically Inactive)		
Physically Active	0.660** (0.576 – 0.758)	0.674** (0.591 – 0.769)
Ever smoked daily (reference: Yes)		
Never smoked daily	0.829** (0.744 – 0.924)	0.833** (0.751 – 0.925)
BMI (reference: Normal)		
Underweight	1.186 (0.818 – 1.720)	1.148 (0.810 – 1.627)
Overweight	0.700** (0.624 – 0.786)	0.710** (0.635 – 0.794)
Obese	0.696** (0.606 – 0.798)	0.706** (0.618 – 0.807)
Life satisfaction	0.929** (0.905 – 0.954)	0.932** (0.909 – 0.956)

***p<1%, *p<5%. Controlling for country of residence. The binary dependent variable indicates whether a respondent is alive or not at the time of Wave 7 data collection. 'Self-perceived age' is included as predictors of in-sample mortality.*

Comparison of the statistical models

The log-likelihood and two pseudo R-square measures, Cox & Snell and Nagelkerke R^2, will be used to compare the models. Pseudo R-square measures assess the improvement of the fitted model compared to the null model, excluding all predictors.

Table 6.12 Pseudo R-square measures.

Statistical model	Measure	Binary logistic regression	Cox regression
Actual mortality model	-2 Log-likelihood	16,003	42,085
	Cox & Shell R^2	2.7%	
	Nagelkerke R^2	11.6%	
Subjective survival probability model	-2 Log-likelihood	14,166	35,772
	Cox & Shell R^2	2.2%	
	Nagelkerke R^2	10.0%	
'Self-perceived age' model	-2 Log-likelihood	13,911	35,516
	Cox & Shell R^2	2.5%	
	Nagelkerke R^2	11.8%	

Note: Lower Log-likelihood and higher pseudo - R^2 values indicate better fit.

Based on the Log-likelihood and the pseudo – R^2 measures the 'Self-perceived age' model has the best fit. The Subjective survival probability model lies in the middle, but the pseudo – R^2 measures indicate that it has poorer fit that the Actual mortality model.

The results of the statistical models are summarized below. Overall, the values and the patterns of odds ratios and hazard ratios are similar.

Demographic Characteristics

Older chronological age and male gender are predictors of mortality (OR = 1.054 and OR = 1.457). The number of children is a weak predictor of mortality (OR=0.981). Based on the changes of regression coefficients, the impact of these demographic factors is not materially affected by the incorporation of subjective survival probabilities and 'Self-perceived age' in the statistical models.

Socio-economic status

Low income and low educational attainment are strong predictors of mortality. Respondents whose income falls in the 1st quartile are 2.128 times more likely to die compared to those whose income falls in the 4th quartile. Respondents who attained Lower secondary education are 1.279 times more likely to die compared to those who completed tertiary education.

Based on the changes of regression coefficients, the impact of income on mortality is reduced by the incorporation of subjective survival probabilities and 'Self-perceived age' in the statistical models. In particular, respondents whose income falls in the 1st quartile are 1.794 – 1.999 times more likely to die compared to those whose income falls in the 4th quartile. The values of hazard ratios are reduced as well.

Physical health & Cognitive function

More ADLs, poor self-rated health and worse orientation in time are strong predictors of mortality. In contrast, numeracy is not significant. An additional limitation in daily activities increases the chances of death by 12.2%; respondents who have excellent self-rated health are 1.85 time less likely to die and better orientation in time by one scale decreases the chances of death by 8%.

Based on the changes of regression coefficients, the impact of ADLs and self-rated health on mortality is reduced by the incorporation of subjective survival probabilities and 'Self-perceived age' in the statistical models. In particular, an additional limitation in daily activities increases the chances of death by 6.4% - 7.3%; respondents who have excellent self-rated health are 1.5 – 2.05 times less likely to die. On the other hand, the impact of orientation in time is not materially affected by the incorporation of subjective survival probabilities and 'Self-perceived age' in the statistical models.

Lifestyle and behavioral risk factors

Respondents who are physically active (OR=0.664) as well as nonsmokers (OR=0.871) have better chances of survival. Furthermore, respondents who are satisfied with their lives have also better chances of survival (OR=0.897). Based on the changes of regression coefficients, the impact of these factors is not materially affected by the incorporation of subjective survival probabilities and 'Self-perceived age' in the statistical models.

Underweight are 1.33 times more likely to die compared to persons of normal weight. In contrast, overweight and obese are 1.41 and 1.45 times less likely to die compared to persons having normal weight. In other words, in the SHARE W6 sample only underweight respondents have lower chances of survival. This finding, which relates to the pattern of in-sample actual mortality, is important for understanding the variation of subjective survival probabilities based on BMI.

Based on the changes of regression coefficients, the impact of underweight on mortality is reduced after the incorporation of subjective survival probabilities and 'Self-perceived age' in the statistical models. In particular, underweight are 1.104 and 1.186 times more likely to die compared to normal.

Subjective Survival Probabilities and 'Self-perceived age'

The results indicate clearly that both Subjective Survival Probabilities which correspond to the next two years and 'Self-perceived age' are both strong and independent predictors of mortality. Based on the changes of regression coefficients, the impact of income, ADLs, self-rated health and underweight on mortality, is reduced by the incorporation of subjective survival probabilities and 'Self-perceived age' in the statistical models. The results imply that both Subjective Survival Probabilities and 'Self-perceived age' include information important for predicting actual survival.

6.5. Comparison of in-sample actual mortality patterns with literature

In this final section the patterns of in-sample actual mortality are compared to the patterns of actual mortality based on relevant literature. Each factor is allocated in one of the following homogeneous groups (Table 6.13).

i. Factors linked to better actual mortality in both SHARE W6 and literature
ii. Factors linked to worse actual mortality in both SHARE W6 and literature

Table 6.13 Grouping of factors based on the consistency of SHARE W6 actual mortality patterns with the relevant literature.

Factors	Factors linked to better actual mortality in both SHARE W6 and literature	Factors linked to worse actual mortality in both SHARE W6 and literature
Demographic		
Age	✓	
Males vs Female		✓
Number of children	✗	
Socio-economic Status		
Higher education	✓	
Higher Income	✓	
Physical Health		
Increase in ADLs		✓
Deterioration of Self-rated health		✓
Cognitive function		
Better Orientation in time	✓	
Better Numeracy	✗	
Lifestyle & Behavioural Risk Factors		
Doing vigorous activities frequently	✓	
Never smoked	✓	
Underweight vs Normal BMI	✓	
Overweight vs Normal BMI		✗
Obese vs Normal BMI		✗
Better life satisfaction	✓	

The factors included in the first homogeneous group are associated with better actual mortality in both SHARE W6 and literature. These factors are:

- Younger chronological age
- Higher income and educational attaintment
- Physical exercise and non-smoking
- Greater life satisfaction

Better numeracy and more children are linked to lighter mortality in the literature but this is not evident in SHARE W6.

The factors included in the second homogeneous group are associated with worse actual mortality in both SHARE W6 and literature. These factors are:

- Male gender
- A higher number of ADLs
- Poor self-rated health
- Underweight vs Normal BMI

Overweight and obese are linked to worse mortality in the literature but this is not evident in SHARE W6.

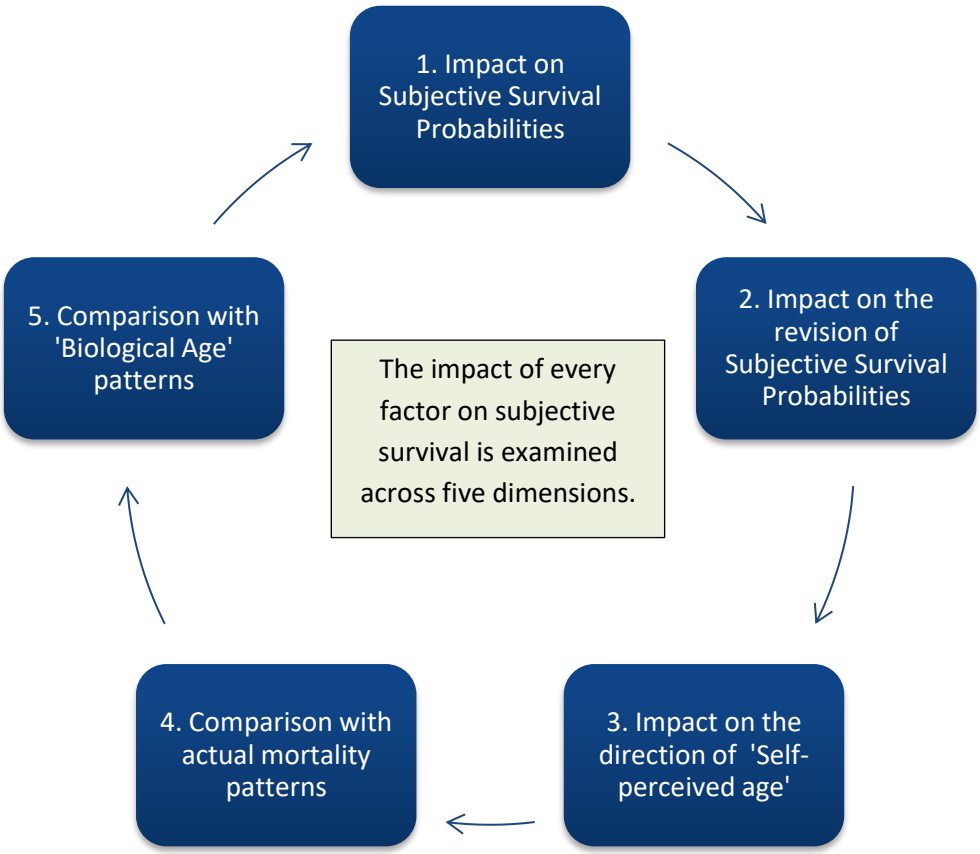
The results and the relevant literature are discussed in more detail in Section 7.

7. Chapter 7 – Discussion of the results

7.1. Introduction

The overall aim of this study is to investigate the impact of several socio-demographic, health and behavioral factors on self-reported subjective survival probabilities. This investigation included a comparison of the subjective survival patterns with the actual mortality patterns. If the patterns are similar, then that would imply that individuals take into account these factors and their survival prediction is in the right direction. In other words, certain survival prediction, but not all, can be considered as 'accurate'. Each factor is examined across the five dimensions, presented below.

Figure 7.1 Overview of the investigation process followed for every factor included in this study.



Factors are allocated in the following three homogeneous groups based on the direction of their associations.

i. Factors associated with higher subjective survival, lower aging and longevity
ii. Factors associated with lower subjective survival, deteriorating aging and higher mortality
iii. Factors which exhibit inconsistent associations with subjective survival, biological age and actual mortality

7.2. Factors associated with higher subjective survival, younger biological age and longevity

This group includes factors which exhibit a clear positive association with:

- Higher subjective survival probabilities
- Upward revision of subjective survival probabilities
- Younger 'Self-perceived age'
- Lower actual mortality, and
- Younger 'Biological Age'

According to our results, more children, higher socio-economic status, physical activities, frequent consumption of fruits or vegetables and eggs or legumes as well as better quality of life are associated with higher subjective survival probabilities. In addition, persons who do vigorous activities frequently tend to have a 'Self-perceived age' younger than their Chronological Age. Male non-smokers tend to have 'Self-perceived age' younger than their Chronological Age. Furthermore, increases in income, improvements in life satisfaction as well as in self-rated health are associated with upward revisions of subjective survival probabilities. In-sample actual mortality is also better for physically active respondents with higher socio-economic status. Moreover, better orientation in time and life satisfaction is liked to lower in-sample actual mortality.

In contrast, a drop in income results in negative revisions of subjective survival probabilities for males but positive revisions of subjective survival probabilities for females. In other words, the effect of income changes on the revision of subjective survival probabilities depends on gender.

The results are broadly consistent with the literature. Higher socio-economic status, better life satisfaction and better self-rated health are associated with lower actual mortality (Kaplan et al. 1996, St. John et al. 2015, Idler and Benyamini 1997) as well as with higher subjective life expectancy (Rarrange et al. 2016; Hurd and McGarry 1995; van Solinge and Henkens 2018). Moreover, reductions in income are important for the mortality of individuals in the lower end of the socio-economic spectrum (Backlund et al. 1999). Smoking reduces life expectancy (Doll et al. 1994) and is associated with accelerated biological aging (Beach et al. 2015).

Past studies noted that individuals who do physical activity regularly, have a 'Biological Age' younger than their Chronological Age. (Nakamura et al. 1989). In addition, physically active individuals over-estimate subjective life expectancy (Griffin et al. 2013; Rarrange et al. 2016; Liu et al. 2007) and they

tend to have to lower mortality and younger Biological Age (Gregg et al. 2003 , Nakamura et al. 1989). Better physical functioning of older adults is also associated with younger 'Subjective Age' (Stephan et al. 2013).

Respondents who live with young or adult children report higher survival expectations (Liu et al. 2007; Mirowsky 1999; Ross and Mirowsky 2002) and the lifespan of parents increases in line with the number of children (McArdle et al. 2006).

The consumption of fruits or vegetables is positively associated with subjective life expectancy (Griffin et al. 2013) and Knoops et al. (2004) noted that elderly people, who adopted Mediterranean diet and healthy lifestyle, reduced their actual mortality rates. A positive relationship between life satisfaction, emotional support and subjective life expectations has been identified in past research (Van Solinge and Henkens 2018, Ross and Mirowsky 2002). Furthermore, better quality of life, greater life satisfaction and higher social connectedness are associated with lower mortality (Netuveli et al. 2012, Buono et al. 1998). Finally, 'receiving help from others' frequently reflects a higher degree of perceived social support which is associated with lower mortality (Lyyra and Heikkinen 2006, Kaplan et al. 1988).

Overall, the factors presented in this section (Figure 6.2) imply that respondents report higher subjective survival probabilities and have younger Self-perceived age' than Chronological Age. These subjective survival patterns are validated by actual mortality and Biological Age patterns.

Figure 7.2 Factors associated with higher subjective survival, younger biological age and longevity.

The conclusion is that for this group of factors, self-reported subjective survival probabilities tend to be ‘accurate’ in predicting future survival. In this context, ‘accuracy’ is interpreted as the tendency of a subjective survival prediction to be correct compared to objective survival (based on life tables) and actual mortality patterns.

7.3. Factors associated with lower subjective survival, older biological age and higher mortality

This group includes factors which exhibit a clear association with:

- Lower subjective survival probabilities
- Downward revision of subjective survival probabilities
- Older 'Self-perceived age'
- Higher actual mortality, and
- Older 'Biological Age'

According to our results, poor self-rated health, more limitations in Activities of Daily Living, a larger number of chronic diseases, poor memory, poor writing skills and depression are associated with lower subjective survival probabilities. Our findings also show that poor self-rated health, more ADLs, poorer memory and more mobility difficulties imply a 'Self-perceived age' older than Chronological Age. In-sample actual mortality is also worse for smokers with poor self-rated health and more ADLs.

It is worth noting that smokers report lower subjective survival probabilities and they tend to have older 'Self-perceived age' compared to never smokers. Furthermore, increase in the number of chronic conditions and ADLs and deterioration in self-rated health lead to negative revisions of subjective survival probabilities.

Our findings are consistent with previous evidence on subjective survival probabilities as well as on actual mortality. Poor health is associated with lower survival expectations (Van Solinge and Henkens 2018; Rarrange et al. 2016; Liu et al. 2007; Hurd and McGarry 1995). Self-rated health is a strong predictor of mortality, (Idler and Benyamini 1997) and this also holds for disability and chronic diseases (Verropoulou 2014). The deterioration of functional health and the increase in the number of comorbidities are also linked to higher actual mortality (Scott et al. 1997, Lee et al. 2008).

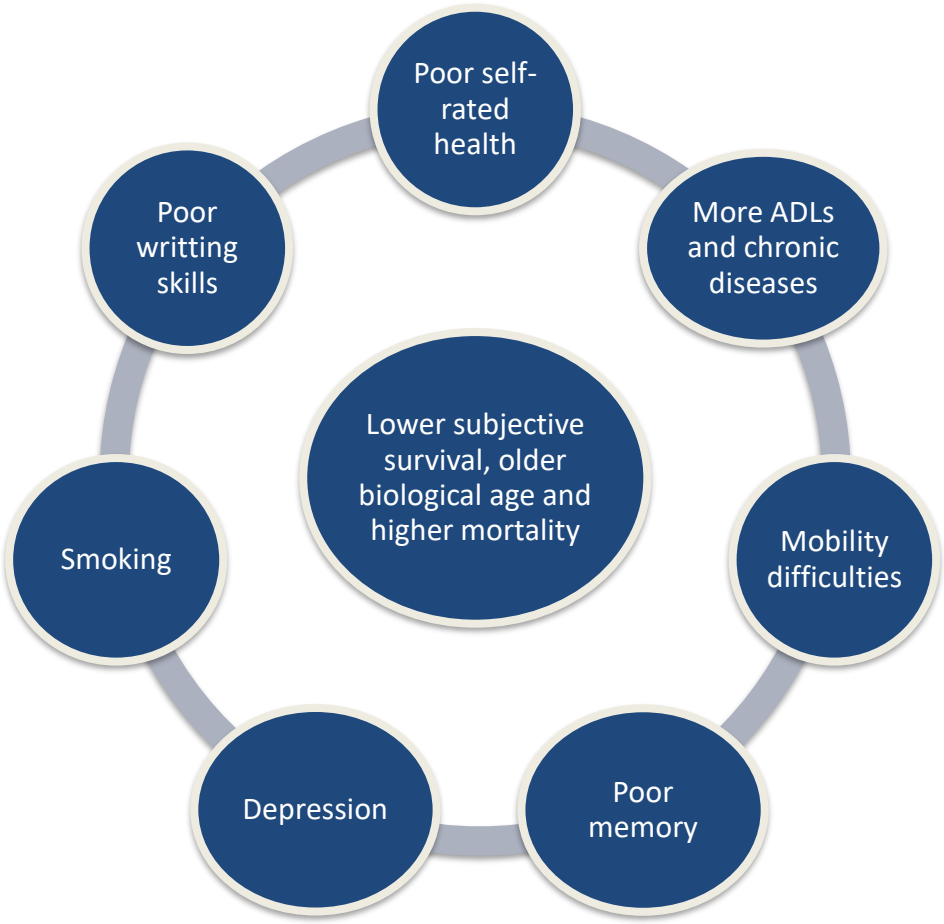
Regarding the aging process, Belsky et al. (2015) and Jylhävä et al. (2017) suggested that persons with older 'Biological Age' tend to have poor health. In addition, individuals with older 'Biological Age' have poorer physical functioning; worse balance, less strength and more mobility difficulties. Poorer health is also associated with older 'Subjective Age' (Stephan et al. 2012) and is interpreted as a counter-indicator of youthfulness.

Prior research noted that psychological distress and depression are negatively associated with survival expectations (Griffin et al. 2013; Rarrange et al. 2016). Further, depression is associated with higher mortality (Wulsin et al. 1999) and this also holds regarding poor cognitive function (Smits et al. 1999).

Smoking increases the chances of mortality (Ezzati & Lopez 2003) but the evidence is unclear for its impact on subjective survival. On the one hand, past analyses noted that current smokers report lower survival expectations (Hurd & McGarry 1995; Rarrange et al. 2016) whereas those who stopped smoking report higher survival expectations (Rarrange et al. 2016). On the other hand, several analyses show that current smokers over-estimate survival (Liu et al. 2007; Balia 2011).

Overall, the factors presented in this section (Figure 6.3) imply that respondents report lower subjective survival probabilities and have older Self-perceived age' than Chronological Age. These subjective survival patterns are validated by actual mortality and Biological Age patterns.

Figure 7.3 Factors associated with lower subjective survival, older biological age and higher mortality.



The conclusion is that for this group of factors, self-reported subjective survival probabilities tend to be 'accurate' in predicting future survival. In this context, 'accuracy' is interpreted as the tendency of

a subjective survival prediction to be correct compared to objective survival (based on life tables) and actual mortality patterns.

7.4. Factors which exhibit inconsistent associations with subjective survival, biological age and actual mortality

This group includes factors which do not exhibit a stable and consistent positive or negative association among subjective survival, biological aging and actual mortality. The most common case is such a factor to exhibit a positive association with subjective survival probabilities and a negative association with actual mortality.

Gender

Gender differentiates actual and subjective survival. According to our findings, males estimate that their future survival will be higher than that of the general population, compared to females and this finding is consistent with past research (Arpino et al. 2017; Liu et al. 2007; Mirowsky 1999). By contrast, it is well known that women experience lower mortality. In addition, the results of the Thesis show that in-sample actual mortality is better for females than males.

Two alternative explanations are provided. First, males have on average higher socioeconomic status, level of education and income while they also exhibit lower levels of morbidity. Higher socioeconomic status and lower morbidity are both associated with higher subjective survival probabilities. Second, according to the well documented 'gender paradox'; women live longer but report poorer health than men (Mathers et al. 2001, Austad 2006).

Chronological age

Chronological age also differentiates actual and subjective survival. The results of this study show that older people tend to report higher subjective survival probabilities. In other words, the gap between subjective and objective survival probabilities increases with age. In addition, the results of the Thesis show that in-sample actual mortality deteriorates in line with chronological age.

Several researchers have also estimated a positive association between chronological age and subjective life expectancy (Griffin et al. 2013; Mirowsky 1999; Ross and Mirowsky 2002) and concluded that people become more optimistic with age. It is worth noting that for respondents who provided proxy interviews, no information on subjective survival expectations was collected (see Table A1 in the Appendix). The persons excluded were, on average, older and had worse health compared to the other participants. This fact might have affected the estimated effect of age on the difference between subjective and objective survival probabilities.

Older people may report higher subjective survival probabilities because they have better psychology. Freund and Baltes (2002) found a positive association between life-management behaviors, such as the Selection-Optimization-Compensation model, and levels of subjective well-being. This theory could provide an alternative explanation about the optimism of subjective survival observed among older respondents. 'Denial of aging' is another theory that could explain why older individuals feel younger (Barak et al. 2001). According to this theory older adults tend to maintain young subjective age identities.

Marital status

The results of this study suggest that widowed, divorced and never married tend to estimate that their future survival will exceed that of the general population compared to married persons. Furthermore, widowhood is associated with negative revisions of subjective survival probabilities for both males and females. In contrast, a divorce implies positive revisions for males and negative revisions for females.

The impact of marital status on subjective survival probabilities is not fully understood. On the one hand, it has been suggested that widows report higher SSPs (Balia 2011) while living alone is negatively associated with subjective survival (Liu et al. 2007; Rarrange et al. 2016). On the other hand, past studies noted that partnership status is not associated with subjective life expectancy (Van Solinge and Henkens 2018).

Actual mortality patterns differentiate based on marital status. First, married persons exhibit lower mortality compared to all other groups (Manzoli et al. 2007; Kaplan and Kronick 2006). Further, excess mortality for widows is observed during the first year following bereavement (Kaprio et al. 1987) but that risk is reduced over the following years. This is in line with the negative revisions of subjective survival probabilities following widowhood. On the other hand, divorce is associated with higher mortality for both genders (Dupre et al. 2009). This is in line with the negative revisions of subjective survival probabilities observed for females, after a divorce. It is worth noting that divorced and never married persons in the present sample are, on average, younger (by about 2 years) compared to married, while they also have slightly better educational attainment. This could affect their optimism regarding future survival.

Numeracy and orientation in time

According to the results of this study, better numeracy score is associated with a decrease in the difference of subjective and objective survival probabilities. In other words, respondents with better numeracy score report subjective survival probabilities close to the general population. On the other hand, better orientation in time is associated with reporting lower subjective survival probabilities. In addition, the results of the Thesis show that in-sample actual mortality improves as orientation in time score increases. In contrast, the score of numeracy test is not a predictor of in-sample actual mortality.

Better subtraction score implies a 'Self-perceived age' closer to Chronological Age. It is worth highlighting that in the present study, the association of subjective survival probabilities with numeracy and orientation in time differentiates compared to that with memory and writing skills. Numeracy score has also a different association with subjective survival probabilities compared to the score of orientation in time.

Past research noted those individuals with better numeracy skills as well as better recall score may form more accurate probability assessments about own survival (d'Uva et al. 2017; Elder 2007). Better cognitive function is related to younger 'Subjective Age' (Stephan et al. 2016). Furthermore, Belsky et al. (2015) concluded that individuals with older 'Biological Age' had poorer cognitive function and Shipley et al. (2006) showed that British adults with better memory skills face lower mortality risks. Moreover, poor cognitive function is linked to worse actual mortality (Kelman et al. 1994).

Body Mass Index (BMI)

Our findings suggest that obese people tend to report higher subjective survival probabilities compared to those who maintain a normal weight. Regarding the revision of subjective survival probabilities, our results show that gender plays an important role. In particular:

- Males who change their BMI from underweight to normal or from obese to normal tend to revise upwards their subjective survival probabilities. The opposite trend is observed for females.
- Transitions from the BMI categories normal or underweight to overweight or obese are associated with positive revisions of subjective survival probabilities. This holds for both males and females.

In addition, the results of the Thesis show that in-sample actual mortality is worse for underweight compared to normal; while it is better for overweight and obese compared to normal.

Past analyses have noted that obesity is negatively associated with survival expectations (Rarrange et al. 2016; Liu et al. 2007). Furthermore, Ross & Mirowsky (2002) argue that obese tend to misestimate their survival expectations compared to persons of normal weight. Actual mortality is lower for individuals whose BMI is normal and higher for the other categories (Berrington et al. 2010). Moreover, underweight is a strong predictor of mortality (Takala et al. 1994) and obesity is a factor related to higher mortality (Solomon and Manson 1997). It is worth noting that in our data obese persons are more frequently younger males and these two features are associated with higher subjective survival probabilities.

Overall, the findings suggest that BMI leads to less 'accurate' subjective survival probabilities. In this context, 'accuracy' is interpreted as the tendency of a subjective survival prediction to be correct. Furthermore, as all BMI categories, apart from the normal, are associated higher mortality (Berrington et al. 2010) we conclude that:

- The in-sample actual mortality patterns are consistent with the literature only for underweight
- Therefore, the transition out of the normal BMI category to underweight results in a downward revision of subjective survival probabilities.

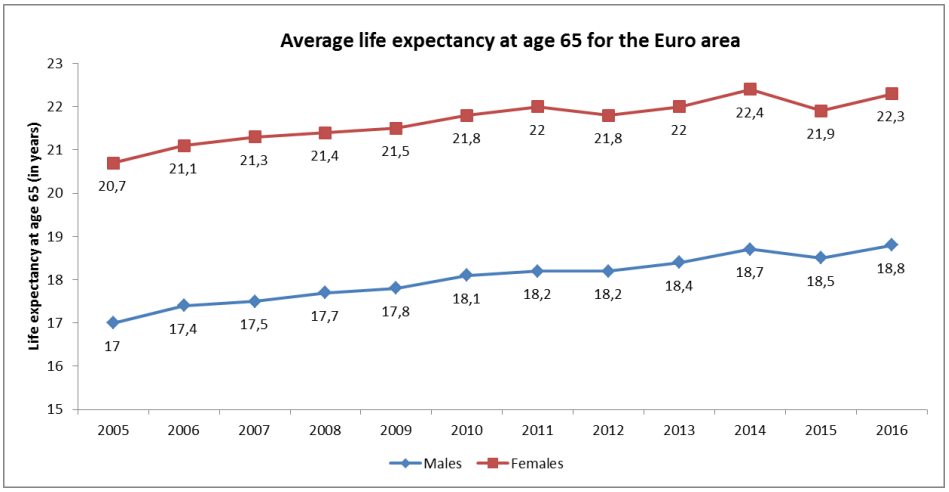
7.5. Study limitations

This study has several limitations which should be taken into account when considering the findings. First, SHARE includes only one question on subjective survival probabilities. This does not allow estimation of the whole distribution of subjective survival probabilities across all ages, for each individual. However, the impact of this limitation is not expected to be material, because of the large number of respondents who report subjective survival probabilities in Wave 6.

Second, in this study the HMD period life tables were used in order to estimate objective survival probabilities. The HMD period life tables are based on the whole population by country and sex. On the one hand, this choice reduces parameter error, because the estimation of the mortality rates is robust. On the other hand, these life tables reflect general population mortality and do not vary by other characteristics known to have an impact on life expectancy, such as socio-economic status.

Period life tables are based on several cohorts and thus they reflect the average mortality across these cohorts without allowing for cohort-specific mortality improvements. On this topic, Goldstein and Wachter (2006) noted that for populations whose mortality patterns evolve, period life expectancy is a lagged measure of cohort life expectancy. The size of the lag depends on the pace of mortality improvement. According to Eurostat, the average life expectancy at age 65 for individuals in Europe has increased annually by 0.92% for males and by 0.69% for females (Figure 6.1) over the past ten years. Since the average mortality improvement is lower for females, it is expected that the comparison of subjective and objective survival probabilities for females, is less affected by the decision not to incorporate cohort-specific mortality improvements in the analysis.

Figure 7.4 Average life expectancy at age 65 for the Euro-area (18 countries).



Source: Eurostat, Life expectancy at age 65, by sex.

To mitigate further this risk, sensitivity analysis was undertaken. Moe specifically, the results seem consistent across different tolerance levels and different models that were incorporated in the statistical analysis. The interpretation is based on the direction of the subjective survival probabilities in conjunction with the direction of actual mortality and biological aging. Ideally, objective survival probabilities should have been based on the actual mortality of the panel respondents of SHARE; however, due to high attrition rates between waves that was not possible (Peracchi and Perotti, 2014; Bergmann et al. 2017).

The mortality improvement affects the calculation of 'Self-perceived age'. The algorithm for the estimation of 'Self-perceived age' requires as inputs the objective survival probabilities. In this study, the HMD period life tables by country and sex, which are estimated for the general population were used. Cohort-specific mortality improvements were not included. The incorporation of mortality improvements is expected to marginally reduce 'Self-perceived age' compared to the current period-based 'Self-perceived age'.

Third, 'Biological Age' was not calculated for respondents in this sample, due to the lack of relevant data. Therefore, all comparisons are based only on the findings of the relevant literature

Fourth, the longitudinal sample of Wave 6 and Wave 7 is reduced to 49505 individuals (-27%). Sample attrition is a phenomenon that often affects longitudinal studies. Prior research noted that the youngest and oldest respondents show the highest attrition rates (Börsch-Supan et al. 2008). Furthermore, there is a tendency of good or better health respondents to participate in the next SHARE waves. The shrinkage of the longitudinal sample size in this instance could be partially due to country-specific retention rate targets, ranging from 75% to 85%, set by SHARE (Malter et al. 2018). As a result, the longitudinal sample has a lower representation of older, less educated, less satisfied, and less healthy individuals. However, even though the actual impact of these factors may have been underestimated in the present study due to selectivity, according to our findings the effect of life satisfaction and health status is consistent with actual mortality patterns for both males and females.

Fifth, in the analysis of the subjective survival probabilities revisions, the number of respondents is low for transitions between certain BMI categories. This may affect the validity of our results related to the effect of BMI on subjective survival probabilities revisions. Further analyses using larger datasets are required to confirm our findings regarding to the influence of BMI transitions on SSPs revisions.

7.6. Areas of further research

This section includes a detailed discussion on the next steps of research on subjective survival.

7.6.1. Further research on subjective survival probabilities

Patterns of subjective survival probabilities

Future research should involve a more detailed examination of the relative effects of various cognitive function indicators, of differences by country of residence as well as the exploration of the effect of marital status on subjective survival probabilities. Another interesting topic is the investigation of the interactions between the predictors. As an example, I mention the impact of the interaction between gender and marital status on the direction of subjective survival probabilities. The impact of psychological factors should be further explored.

Accuracy of subjective survival probabilities

A very interesting topic is the examination of the 'accuracy' of subjective survival probabilities. The results of this Thesis suggest that there are groups of factors which are taken into account when forming subjective survival probabilities in a way consistent with actual mortality patterns (see Figures 5.2 and 5.3). In this context, 'accuracy' is interpreted as the tendency of a subjective survival prediction to be confirmed by the patterns of actual mortality and biological aging. A further step would be to evaluate the accuracy of subjective survival probabilities using a panel dataset, which would include information from previous SHARE, HRS and ELSA waves and data about the actual survival of the respondents.

Revision of subjective survival probabilities

The revision of subjective survival probabilities after the occurrence of certain events is also a topic for further research. The next steps should involve a more detailed examination of the influences of a wide range of events on the revision of subjective survival probabilities, using large longitudinal datasets covering many aspects of respondents' lives.

7.6.2. Further research on 'Self-perceived age'

Calculation of 'Self-perceived age'

'Self-perceived age' is a quantity, expressed in years, which incorporates survival-related information from two sources, namely, population life tables and subjective survival probabilities. The algorithm for the calculation of 'Self-perceived age' requires mainly two sources of information, namely, subjective survival probabilities and population life tables. Further enhancement of the calculation algorithm include:

- The use of cohort-life tables instead of period life tables,
- The incorporation of mortality patters, estimated by the actual mortality of the respondents
- The investigation of a range of plausible values of 'Self-perceived age' in cases where subjective survival probabilities takes values of 0% or 100%.

Patterns of 'Self-perceived age'

In terms of the variation of 'Self-perceived age', the next steps include a detailed examination of the impact of all main sociodemographic, health and lifestyle factors on the size of the difference of 'Self-perceived age' from chronological age.

7.6.3. Bayesian Demography

Applications of Bayesian Demography

Bayesian Demography is a research field in Demography which incorporates the use of Bayesian techniques for data analysis and statistical inference. Bijak and Bryant (2006) note that two key applications of Bayesian Demography are population forecasts and statistical inference in the presence of limited data. For example the estimation of fertility rates and death rates for small populations or sub-populations.

Dellaportas, Smith and Stavropoulos (2001) used a Bayesian extension of Heligman and Pollard model (1980) to fit this model in mortality data from England and Wales. During their modelling process they treat all model parameters as unknowns and prior information via probability density functions is assigned. The means and standard deviations of the priors aim to capture the heterogeneity observed in death rates. Alexander, Zagheni and Barbieri (2017) used a Bayesian hierarchical model to estimate mortality at a subnational level for the USA and France. Mortality rates are assumed to follow a Poisson process which has a hierarchical structure. The first level models the log of the mortality rate as a linear equation of age, gender and geographical area. The second level specifies distributions for

the parameters of the first level (i.e. the distribution coefficients of the log-linear equation, given the mean and variance) whereas the third level specifies the distribution for the parameters at the second level (i.e. the distribution of mean and variance of the distribution of the coefficients).

Bayesian Demography and subjective survival probabilities

Bayesian techniques could be used to incorporate subjective survival probabilities as well as population life tables in order to explain the heterogeneity in mortality patterns. The findings of this Thesis conclude that certain factors affect the patterns of subjective survival probabilities in the direction of actual mortality. As part of the Bayesian analysis, prior distributions need to be assigned to the parameters of mortality models. The parameters of these priors could be based on the values of subjective survival probabilities. Bayesian mortality models in combination with subjective survival probabilities could be used to estimate population life tables which vary based on certain factors which enhance the 'accuracy' of survival predictions (e.g. self-rated health or socio-economic status)

7.7. Conclusions

The results of this Thesis show that there are common factors affecting self-reported subjective survival probabilities. On the one hand, factors such as a large number of children, higher socio-economic status, frequent physical activity, frequent consumption of fruits or vegetables and eggs or legumes as well as better quality of life are associated with higher subjective survival, younger biological age and longevity. On the other hand, factors such as poor self-rated health, more limitations in Activities of Daily Living, a larger number of chronic diseases, poor memory, poor writing skills and depression are associated with lower subjective survival, older biological age and higher mortality. For these groups of factors, self-reported subjective survival probabilities tend to be 'accurate' in predicting future survival. In this context, 'accuracy' is interpreted as the tendency of a subjective survival prediction to be correct.

The categorization of the findings of this Thesis, indicate a third group of factors which includes gender, chronological age, marital status, certain aspects of cognitive function and BMI. These factors do not exhibit a clear and consistent association between subjective survival probabilities, biological aging and actual mortality. This would imply that these factors reduce the predictive power of subjective survival probabilities.

The revision of subjective survival probabilities is also investigated in this Thesis. According to the findings, individuals update their subjective survival probabilities after the occurrence of specific events, consistently with actual mortality patterns. In particular, males and females tend to revise their subjective survival probabilities following self-rated and functional health deterioration or improvement, an increase in chronic conditions, widowhood, and increases in income as well as in life satisfaction. There is, however, differentiation by gender. Only males tend to revise their subjective survival probabilities following a drop in income consistently with actual mortality patterns. In contrast, only females tend to revise their subjective survival probabilities after a getting divorced consistently with actual mortality patterns.

The third part of this Thesis presents a methodology for estimating individual 'Self-perceived age', which incorporates information from subjective survival probabilities and population life tables. The findings of the analysis show that 'Self-perceived age' patters are similar to the patterns of subjective survival probabilities. The main advantage is that 'Self-perceived age' is expressed in years instead of percentage values. Therefore 'Self-perceived age' is directly comparable to chronological age.

Furthermore, the patterns of ‘Self-perceived age’ are also comparable to the patterns of Biological Age and Subjective Age.

The results of the last part of this Thesis indicate clearly that Subjective Survival Probabilities and ‘Self-perceived age’ are both strong and independent predictors of mortality. This implies that both Subjective Survival Probabilities and ‘Self-perceived age’ include ‘survival information’ important for predicting actual survival.

References

Alexander, M., Zagheni, E., & Barbieri, M. (2017). A flexible Bayesian model for estimating subnational mortality. *Demography*, 54(6), 2025-2041.

Antonovsky, A. (1967). Social class, life expectancy and overall mortality.

Arpino, B., Bordone, V., & Scherbov, S. (2018). Smoking, education and the ability to predict own survival probabilities. *Advances in Life Course Research*, *37*, 23-30.

Austad, S. N. (2006). Why women live longer than men: sex differences in longevity. *Gender medicine* 3(2), 79-92.

Balia, S. (2014). Survival expectations, subjective health and smoking: Evidence from SHARE. *Empirical Economics*, *47*(2), 753-780.

Baltes, M. M., & Carstensen, L. L. (1996). The process of successful ageing. *Ageing & Society*, 16(4), 397-422.

Barrett, A. E. (2003). Socioeconomic status and age identity: The role of dimensions of health in the subjective construction of age. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 58*(2), S101-S109.

Barrett, A. E. (2005). Gendered experiences in midlife: Implications for age identity. *Journal of Aging Studies*, *19*(2), 163-183.

Barak, B., Mathur, A., Lee, K., & Zhang, Y. (2001). Perceptions of age–identity: A cross-cultural inner-age exploration. *Psychology & Marketing, 18*(10), 1003-1029.

Bergmann, M., Kneip, T., De Luca, G., & Scherpenzeel, A. (2017). Survey participation in the survey of health, ageing and retirement in Europe (SHARE), Wave 1-6. *Munich: Munich Center for the Economics of Aging*.

Belsky, D.W., Caspi, A., Houts, R., Cohen, H.J., Corcoran, D.L., Danese, A., Harrington, H., Israel, S., Levine, M.E., Schaefer, J.D., Sugden, K., Williams, B., Yashin, A.I., Poulton, R., Moffitt, T.E., (2015). Quantification of biological aging in young adults. *Proceedings of the National Academy of Sciences*, 112(30), E4104-E4110.

Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., & Weber, G. (2008). First results from the Survey of Health, Ageing and Retirement in Europe (2004-2007). *Starting the longitudinal dimension. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).*

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S. and Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). International Journal of Epidemiology DOI: 10.1093/ije/dyt088

Börsch-Supan A. (2017) Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 6.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w6.600

Börsch-Supan, A. (2019). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 7. Release version: 7.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w7.700

Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A., & Nesbitt, C. J. (1997). *Actuarial mathematics. Second Edition. Chapter 3.*

Bijak, J., & Bryant, J. (2016). Bayesian demography 250 years after Bayes. *Population studies*, 70(1), 1-19.

Buono, M. D., Urciuoli, O., & LEO, D. D. (1998). Quality of life and longevity: a study of centenarians. *Age and ageing* 27(2), 207-216.

Dellaportas, P., Smith, A. F., & Stavropoulos, P. (2001). Bayesian analysis of mortality data. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* 164(2), 275-291.

Doll, R., Peto, R., Wheatley, K., Gray, R., & Sutherland, I. (1994). Mortality in relation to smoking: 40 years' observations on male British doctors. *Bmj*, *309*(6959), 901-911.

d'Uva, T. B., O'Donnell, O., & van Doorslaer, E. (2017). Who can predict their own demise? Heterogeneity in the accuracy and value of longevity expectations. *The Journal of the Economics of Ageing*.

Ebrahim, S., Wannamethee, G., McCallum, A., Walker, M., & Shaper, A. G. (1995). Marital status, change in marital status, and mortality in middle-aged British men. *American Journal of Epidemiology*, *142*(8), 834-842.

Eikemo, T. A., Bambra, C., Joyce, K., & Dahl, E. (2008). Welfare state regimes and income-related health inequalities: a comparison of 23 European countries. *The European Journal of Public Health*, 18(6), 593-599.

Elder T. E. (2007). Subjective Survival Probabilities in the Health and Retirement Study: Systematic Biases and Predictive Validity, Michigan Retirement Research Center Research Paper No. WP 2007-159. Available at SSRN: https://ssrn.com/abstract=1083823 or http://dx.doi.org/10.2139/ssrn.1083823

Ezzati, M., & Lopez, A. D. (2003). Estimates of global mortality attributable to smoking in 2000. *The Lancet* 362(9387), 847-852.

Flisi S., Goglio V., Meroni E., Vera-Toscano E., & Dragomirescu-Gaina C. (2015). Skills beyond education. An analysis of cognitive skill evolution and its implications for employment chances. *Publications Office of the European Union*, Luxembourg, doi, 10, p. 129633

Freund, A. M., & Baltes, P. B. (2002). Life-management strategies of selection, optimization and compensation: Measurement by self-report and construct validity. *Journal of personality and social psychology*, 82(4), 642.

Gateway to Global Aging Data, Produced by the USC Program on Global Aging, Health & Policy, with funding from the National Institute on Aging (R01 AG030153)

Goldstein, J. R., & Wachter, K. W. (2006). Relationships between period and cohort life expectancy: Gaps and lags. *Population Studies*, 60(3), 257-269

Gregg, Edward W., Jane A. Cauley, Katie Stone, Theodore J. Thompson, Douglas C. Bauer, Steven R. Cummings, Kristine E. Ensrud, and Study of Osteoporotic Fractures Research Group. (2003). Relationship of changes in physical activity and mortality among older women. *Jama* 289 (18) : 2379-2386.

Griffin, Barbara, Vanessa Loh, and Beryl Hesketh. "A mental model of factors associated with subjective life expectancy." *Social Science & Medicine* 82 (2013): 79-86.

Hagenaars, A. J., De Vos, K., & Asghar Zaidi, M. (1994). Poverty statistics in the late 1980s: Research based on micro-data.

Hamermesh, Daniel S. "Expectations, life expectancy, and economic behavior." *The Quarterly Journal of Economics* 100, no. 2 (1985): 389-408.

Health and Retirement Study, RAND HRS Longitudinal File 2014 v2 public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740).

Heligman, L., & Pollard, J. H. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries*, 107(1), 49-80.

Hurd, M. D., & McGarry, K. (1995). Evaluation of the subjective probabilities of survival in the health and retirement study. *Journal of Human resources*, S268-S292.

Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org (data downloaded on 26/09/2017).

Hurd M. D. and McGarry K. (2002). The predictive validity of subjective probabilities of survival. *The Economic Journal* 112(482): 966-985.

Hurd M.D. and McGarry K. (1995). Evaluation of the subjective probabilities of survival in the health and retirement study. *The Journal of Human Resources* pp. S268-S292.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior*, 21-37.

Ikeda, A., Iso, H., Toyoshima, H., Kondo, T., Mizoue, T., Koizumi, A., Inaba, I., Tamakoshi, A., & JACC Study Group. (2006). Parental longevity and mortality amongst Japanese men and women: the JACC Study. *Journal of internal medicine, 259(3)*, 285-295.

Johnson, T. E. (2006). Recent results: biomarkers of aging. *Experimental gerontology*, *41*(12), 1243-1246.

Jylhävä, J., Pedersen, N. L., & Hägg, S. (2017). Biological age predictors. *EBioMedicine*, *21*, 29-36.

Kaprio, J., Koskenvuo, M., & Rita, H. (1987). Mortality after bereavement: a prospective study of 95,647 widowed persons. *American Journal of Public Health, 77(3*), 283-287.

Kaplan, G. A., Pamuk, E. R., Lynch, J. W., Cohen, R. D., & Balfour, J. L. (1996). Inequality in income and mortality in the United States: analysis of mortality and potential pathways. Bmj, 312(7037), 999-1003.

Kaplan, G. A., Salonen, J. T., Cohen, R. D., Brand, R. J., Leonard Syme, S., & Puska, P. (1988). Social connections and mortality from all causes and from cardiovascular disease: prospective evidence from eastern Finland. American journal of epidemiology, 128(2), 370-380.

Kelman, H. R., Thomas, C., Kennedy, G. J., & Cheng, J. (1994). Cognitive impairment and mortality in older community residents. *American journal of public health* 84(8), 1255-1260.

Kleinspehn-Ammerlahn, A., Kotter-Grühn, D., & Smith, J. (2008). Self-perceptions of aging: Do subjective age and satisfaction with aging change during old age? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 63*(6), P377-P38

Klemera, P., & Doubal, S. (2006). A new approach to the concept and computation of biological age. *Mechanisms of ageing and development, 127(3),* 240-248.

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

Knoops K. T., de Groot L. C., Kromhout D., Perrin A. E., Moreiras-Varela O., Menotti A., and Van Staveren W. A. (2004). Mediterranean diet, lifestyle factors, and 10-year mortality in elderly European men and women: the HALE project. *Jama* 292(12): 1433-1439.

Kutlu-Koc V. and Kalwij A. (2017). Individual survival expectations and actual mortality: Evidence from Dutch survey and administrative data. *European Journal of Population* 33(4): 509-532.

Kutlu-Koc, V. and Kalwij, A. (2013). Individuals' survival expectations and actual mortality. Netspar Discussion Papers 13.

Liu J. T., Tsou M. W., and Hammitt J. K. (2007). Health information and subjective survival probability: Evidence from Taiwan. *Journal of Risk Research* 10(2): 149-175.

Lyyra, T. M., & Heikkinen, R. L. (2006). Perceived social support and mortality in older people. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61(3), S147-S152.

Levine, M. E. (2013). Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age?. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences, 68*(6), 667-674.

Levy, B. R., Slade, M. D., Kunkel, S. R., & Kasl, S. V. (2002). Longevity increased by positive self-perceptions of aging. *Journal of personality and social psychology*, *83*(2), 261.

Liu J. T., Tsou M. W., and Hammitt J. K. (2007). Health information and subjective survival probability: Evidence from Taiwan. *Journal of Risk Research* 10(2): 149-175.

Lyubomirsky S. (2001). Why are some people happier than others? The role of cognitive and motivational processes in well-being. *American psychologist* 56(3): 239.

Manzoli, L., Villari, P., Pirone, G. M., and Boccia, A. (2007). Marital status and mortality in the elderly: a systematic review and meta-analysis. *Social science & medicine* 64(1): 77-94.

Mathers, C. D., Sadana, R., Salomon, J. A., Murray, C. J., & Lopez, A. D. (2001). Healthy life expectancy in 191 countries, 1999. *The Lancet*, 357(9269), 1685-1691.

McArdle, P.F., Pollin, T.I., O'Connell, J.R., Sorkin, J.D., Agarwala, R., Schäffer, A.A., Streeten, E.A., King, T.M., Shuldiner, A.R. and Mitchell, B.D.(2006). Does having children extend life span? A genealogical study of parity and longevity in the Amish. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 61(2): 190-195.

Mirowsky J. (1999). Subjective life expectancy in the US: correspondence to actuarial estimates by age, sex and race. *Social science & medicine* 49(7): 967-979.

Nakamura, E., Moritani, T., & Kanetaka, A. (1989). Biological age versus physical fitness age. *European journal of applied physiology and occupational physiology*, *58*(7), 778-785.

Netuveli, G., Pikhart, H., Bobak, M., & Blane, D. (2012). Generic quality of life predicts all-cause mortality in the short term: evidence from British Household Panel Survey. *J Epidemiol Community Health* 66(10), 962-966.

Olaya, B., Moneta, M. V., Doménech-Abella, J., Miret, M., Bayes, I., Ayuso-Mateos, J. L., & Haro, J. M. (2017). Mobility difficulties, physical activity, and all-cause mortality risk in a nationally representative sample of older adults. *The Journals of Gerontology: Series A, 73(9)*, 1272-1279.

Park, M. H., Kwon, D. Y., Jung, J. M., Han, C., Jo, I., & Jo, S. A. (2013). Mini-Mental Status Examination as predictors of mortality in the elderly. *Acta Psychiatrica Scandinavica*, *127*(4), 298-304.

Peracchi, F., Perotti, V. (2010). Subjective survival probabilities and life tables: evidence from Europe. Einaudi Institute for Economics and Finance (EIEF) Working Paper, 10, 16.

Perozek, M. (2008). Using subjective expectations to forecast longevity: Do survey respondents know something we don't know? *Demography*, 45, 95–113. (EIEF) Working Paper 10, p. 16.

Post, T. and Hanewald, K. 2010. Stochastic Mortality, Subjective Survival Expectations, and Individual Saving Behavior. No. 2010-040. SFB 649 discussion paper.

Rappange, David R., Werner BF Brouwer, and Job van Exel. "Rational expectations? An explorative study of subjective survival probabilities and lifestyle across Europe." *Health Expectations* 19, no. 1 (2016): 121-137.

Ross C. E. and Mirowsky J. (2002). Family relationships, social support and subjective life expectancy. *Journal of health and social behavior* pp. 469-489.

Rubin, D. C., & Berntsen, D. (2006). People over forty feel 20% younger than their age: Subjective age across the lifespan. *Psychonomic bulletin & review*, *13*(5), 776-780.5.

Solomon, C. G., & Manson, J. E. (1997). Obesity and mortality: a review of the epidemiologic data. *The American journal of clinical nutrition* 66(4), 1044S-1050S.

Sebastiani, P., Thyagarajan, B., Sun, F., Schupf, N., Newman, A. B., Montano, M., & Perls, T. T. (2017). Biomarker signatures of aging. *Aging Cell*, *16*(2), 329-338.

Smits, C. H., Deeg, D. J., Kriegsman, D. M. and Schmand, B. (1999). Cognitive functioning and health as determinants of mortality in an older population. *American journal of epidemiology* 150(9): 978-986

Stephan, Y., Demulier, V., & Terracciano, A. (2012). Personality, self-rated health, and subjective age in a life-span sample: The moderating role of chronological age. *Psychology and aging*, *27*(4), 875.

Stephan, Y., Sutin, A. R., Caudroit, J., & Terracciano, A. (2016). Subjective age and changes in memory in older adults. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 71*(4), 675-683.

Stephan, Y., Chalabaev, A., Kotter-Grühn, D., & Jaconelli, A. (2013). "Feeling younger, being stronger": An experimental study of subjective age and physical functioning among older adults. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 68(*1), 1-7.

Teppa, Federica, Can the Longevity Risk Alleviate the Annuitization Puzzle? Empirical Evidence from Dutch Data (July 1, 2011). *De Nederlandsche Bank Working Paper* No. 302. Available at SSRN: https://ssrn.com/abstract=1951891 or http://dx.doi.org/10.2139/ssrn.1951891

Thyagarajan, B., Shippee, N., Parsons, H., Vivek, S., Crimmins, E., Faul, J., & Shippee, T. (2019). How Does Subjective Age Get "Under the Skin"? The Association Between Biomarkers and Feeling Older or Younger Than One's Age: The Health and Retirement Study. *Innovation in aging, 3*(4), igz035.

Van Solinge H. and Henkens K. (2009). Living longer, working longer? The impact of subjective life expectancy on retirement intentions and behaviour. *European journal of public health* 20(1): 47-51.

Van Solinge, H. and Henkens, K. (2018). Subjective life expectancy and actual mortality: results of a 10-year panel study among older workers. *European journal of ageing* 15(2), pp.155-164

Verropoulou, G. (2014). Specific versus general self-reported health indicators predicting mortality among older adults in Europe: disparities by gender employing SHARE longitudinal data. *International journal of public health* 59(4): 665-678.

White, A., & Holmes, M. (2006). Patterns of mortality across 44 countries among men and women aged 15–44 years. *Journal of Men's Health and Gender*, *3*(2), 139-151.

Wulsin, L. R., Vaillant, G. E. and Wells, V. E. (1999). A systematic review of the mortality of depression. *Psychosomatic medicine* 61(1): 6-17.

Appendix

Table A1 Characteristics of respondents (SHARE wave 6) for whom proxy interviews were contacted and have been excluded from the analysis (N=2906).

Self-rated health	Excellent	Very good	Good	Fair	Poor
Proportion of respondents	2%	5%	18%	29%	46%
Average number of chronic diseases	0.6	0.9	1.5	2.6	3.5
Average number of ADLs	0.2	0.1	0.5	1.2	3.2
Average age (in years)	65.6	65.4	69.7	76.5	78.6
Average Depression score	1.4	1.4	2.7	3.2	3.5
Average Life satisfaction score	8.3	8.3	7.4	7.2	7.1
Average Quality of life score	5	3	3.4	2.4	1.6

Table A2 Proportion of older individuals by country and gender.

Country	Proportion aged above 80 years		Proportion aged above 85 years		Proportion aged above 90 years	
Males	**SHARE W6***	**Eurostat** ***	**SHARE W6***	**Eurostat** ***	**SHARE W6***	**Eurostat** ***
Austria	11.32%	8.38%	4.85%	3.29%	1.83%	0.77%
Belgium	12.06%	9.45%	5.24%	3.61%	1.71%	0.88%
Czech Republic	11.27%	6.52%	3.43%	2.27%	0.67%	0.47%
Switzerland	13.66%	8.70%	5.00%	3.54%	0.71%	0.94%
Germany	8.81%	8.14%	3.42%	3.03%	0.53%	0.66%
Denmark	8.68%	7.53%	3.44%	3.05%	1.05%	0.80%
Estonia	11.67%	6.88%	3.98%	2.31%	0.54%	0.43%
Spain	18.38%	10.74%	7.80%	4.27%	2.51%	1.13%
France	13.79%	10.30%	5.47%	4.34%	1.55%	1.18%
Greece	11.95%	11.96%	4.71%	4.64%	0.56%	1.20%
Croatia	5.57%	6.83%	1.64%	2.03%	0.36%	0.38%
Italy	11.49%	10.47%	3.34%	4.14%	0.75%	1.06%
Israel	17.85%	N/A	8.17%	N/A	3.38%	N/A
Luxembourg	8.05%	7.69%	2.54%	2.64%	0.42%	0.56%
Poland	9.97%	6.50%	3.41%	2.26%	1.39%	0.50%
Portugal	8.91%	9.62%	3.06%	3.50%	0.80%	0.86%
Sweden	15.67%	9.44%	6.88%	4.10%	1.96%	1.16%
Slovenia	11.26%	6.81%	3.64%	2.27%	0.83%	0.47%
Females						
Austria	13.48%	14.15%	7.07%	7.13%	2.12%	2.30%
Belgium	13.96%	15.42%	6.53%	7.27%	1.94%	2.27%
Czech Republic	10.32%	11.50%	3.44%	4.89%	0.89%	1.25%
Switzerland	13.57%	14.10%	6.01%	6.93%	2.00%	2.31%
Germany	8.87%	13.80%	3.81%	6.75%	0.69%	2.12%
Denmark	10.91%	11.85%	4.41%	5.99%	1.74%	2.14%
Estonia	15.23%	14.23%	5.60%	6.09%	1.20%	1.51%
Spain	21.04%	16.33%	9.99%	7.76%	2.82%	2.54%
France	18.01%	16.47%	9.22%	8.38%	3.27%	2.90%
Greece	11.66%	14.83%	4.09%	6.27%	0.85%	1.68%
Croatia	7.80%	12.23%	3.22%	4.50%	0.64%	1.07%
Italy	10.31%	16.43%	4.42%	8.03%	1.37%	2.62%
Israel	16.47%	N/A	7.39%	N/A	1.95%	N/A
Luxembourg	9.81%	13.24%	3.39%	6.12%	1.17%	1.79%
Poland	11.80%	11.97%	5.03%	5.05%	1.35%	1.31%
Portugal	11.04%	14.41%	4.22%	6.28%	1.08%	1.81%
Sweden	14.44%	14.50%	6.65%	7.50%	1.75%	2.67%
Slovenia	13.18%	13.87%	4.89%	6.06%	1.04%	1.62%

**The proportion is calculated as the number of respondents (older than 80, 85 or 90 years) over the total number of respondents aged 50 or above. **The proportion of individuals (older than 80, 85 or 90 years) over the total population aged 50 or above.*

In the following section the VBA code for the calculation of Self-perceived age' is presented.

```
Public Sub SPA ()

Dim thisRecord , FirstRespondent , LastRespondent As Long

On Error Resume Next

For thisRecord = FirstRespondent To LastRespondent

    If subjective.Cells(thisRecord, 9) <> "" Then

         If subjective.Cells(thisRecord, 12) <> 0 Then

             If subjective.Cells(thisRecord, 4) = 1 Then ' if SSP=100%

                  subjective.Cells(thisRecord, 8) = 0 ' set SPA = 1

             ElseIf subjective.Cells(thisRecord, 4) <= 0.02 Then ' if SSP <= 2%

                  subjective.Cells(thisRecord, 8) = 110 ' set SPA = 110

                    subjective.Range("L" & thisRecord).GoalSeek Goal:=0,
                    ChangingCell:=subjective.Range("H" & thisRecord)

              Else

                 subjective.Cells(thisRecord, 8) = 40 'set initial value for SPA

                    subjective.Range("L" & thisRecord).GoalSeek Goal:=0,
                    ChangingCell:=subjective.Range("H" & thisRecord)

             End If

        End If

     End If

   Application.StatusBar = "Doing record  " & thisRecord

Next

Application.StatusBar = False

End Sub
```