SUBJECTIVE LIFE EXPECTANCY: DIFFERENCES BY SMOKING, EDUCATION AND GENDER

Sergei Scherbov, Bruno Arpino and Valeria Bordone

Long abstract prepared for the European Population Conference 2016

Abstract

Despite the well-known higher mortality rates among smokers than non-smokers, little investigation has focused on subjective survival probabilities (SSP) by smoking behaviour. We give attention to sub-group differences in subjective survival probabilities, comparing subjective predictions to objective ones (SP) and accounting for the role of education.

We use biannual data from the Health and Retirement Study (HRS) from 2000 to 2012 carried out in the USA. Based on a sample of 23,895 respondents aged 50-89, we calculate, for each respondent, the "gap" between SSP and the estimated survival probability (SP) from the HRS data.

We find that people currently smoking report lower survival probabilities especially if they are low educated. This is consistent with real mortality data that show higher mortality among these groups. When comparing subjective and objective survival probabilities we find that irrespectively of the smoking status, high educated people are more likely to correctly predict their survival probabilities than their low educated counterparts. Within education groups, people who smoked in the past are the best at predicting their mortality. Interestingly, those who currently smoke show the highest probability to incorrectly overestimate their survival probability (i.e., to underestimate the negative effect of smoking on mortality).

Introduction

Within the last two decades, the literature has found and confirmed that the subjective survival probability (SSP) survey question is a good predictor of mortality, even after controlling for mortality-related characteristics such as demographic and socioeconomic characteristics and objective health measures (1–8). SSP questions asks respondents to estimate the probability that they will live up to a certain age and it is a measure that incorporates private and subtle information germane to mortality yet unmeasurable through objective questions (2). Survey respondents seem to have a good knowledge of their life expectancy. Hence, SSP is used to predict and understand individuals' economic behaviors related to later stages of life (9,10) and health behaviors (e.g., 11,12).

Understanding the variability of subjective survival probabilities is important because they may affect life-cycle decisions such as labour market participation, consumption, and savings.

Yet, sub-groups within the population may be more or less able to predict the own survival curve. Several studies have assessed how subjective survival probabilities change with observed individual characteristics such as education, income, employment or health status. Building on this strand of the literature, we postulated that there would be reporting heterogeneity in SSP, focusing on the differences between smokers and non-smokers, with a further distinction between more and less educated. In this study, we first aim to compare subjective survival probabilities obtained from a population survey by education and smoking status. In a second step, we use a Gompertz survival model to study how educational attainment and smoking behaviour affect the objective survival probability obtained from a properly constructed life table on the same population. Third, we will compare subjective and objective life tables in two ways, i.e., looking at survival probabilities and at life expectancy, paying attention to sub-group differences (i.e., by education and smoking behaviour).

Methods

Data

We use the data from the Health and Retirement Study (HRS), an age-cohort–based longitudinal panel survey of persons aged 50 years and older in the United States. In particular, we consider respondents interviewed for the first time in 2000, 2002, 2004, 2006, 2008, 2010, and 2012 waves.

Our analysis applies to older adults aged 50–89 years and excludes respondent who are living in a nursing home. We also exclude respondents interviewed or who died in 2013 because most of the interviews of the last wave of SHARE were held during the year 2012 and for those interviewed in 2013 the exposure time would have been too short. The remaining working sample included 23,895 respondents.

Outcome variables

Subjective survival probability: At the SSP question level, there had been changes to the wording and skip patterns prior to 2000. Since then, SSP has been asked among self-respondents aged 50–89 years in a consistent manner that reads as follows: "I would like for you to give me a number from 0 to 100, where 0 means that you think there is absolutely no chance, and 100 means that you think the event is absolutely sure to happen. What is the percent chance that you will live to be [75 (if age is less 65 or less)/80 (if age is 66–69)/85 (if age is 70–74)/90 (if age is 75–79)/95 (if age is 80–84)/100 (if age is 85–89)] or more?" The target age in expectation is a function of respondents' current age (for details see the Expectations section content area on http://hrsonline.isr.umich.edu/index.php?p=qnaires). For those aged 50–65 years, SSP was first asked with the target age of 75 years; unless they reported 0, SSP was then asked with 80 years as the target age. We used only the first target in the analysis in order to have one target per person.

Survey survival probability: Benefiting from the longitudinal nature of HRS, we know whether a person died after first interview until the year 2013. In addition to the information on vital status obtained by HRS through tracking of respondents, the HRS seeks matches to the National Death Index for persons who are reported as deceased or who are not known to be alive through contact during tracking. For all submitted cases that were flagged as valid by the National Center for Health Statistics (NCHS) and verified by HRS staff, the Tracker file contains year and month of death, match score, and an alive/deceased flag.

Explanatory variables

Our explanatory variables are education and smoking behavior. Education distinguishes between higher (master degree, professional degree) and lower (no degree, GED, two year college degree, four year college degree) attainment. Smoking behavior considers whether the respondent has never smoked; whether the respondent has been a smoker in the past (at least 100 cigarettes), but currently does not smoke; or currently smokes cigarettes.

Control variables

We control for ethnicity (White/Caucasian; Black/African American; other) and health (by considering whether the respondent was diagnosed with cancer, stroke, lung problems, heart disease, or if the person is in a nursing home). We also include dummy variables to control for the wave at which the interview was carried out.

Statistical analysis

In a first step, we used linear regression models to analyze the association between smoking behaviors and education on the one hand and subjective survival probabilities on the other. In this case the outcome is the SSP, i.e. a variable that is bounded at 0 and 100. Therefore, we also carried out the models using a Tobit regression model. Since the linear and the Tobit models gave very similar results, we preferred the linear approach for simplicity of interpretation. From this model we obtained predicted survival probabilities for different sub-groups (i.e. considering smoking behavior and educational attainment) and plotted them to ease interpretation of results.

In a second step, we used a Gompertz survival model applied to real mortality data (i.e., survey data on whether each individual survived or not at each time point) to assess the association between our explanatory variables and objective mortality. From this model, we obtained estimates of objective survival probabilities (OSPs) by smoking behaviour and education. This allowed us, in a third step, to compare respondents´ SSPs and their predicted OSPs according to the Gompertz model.

Results

Descriptives

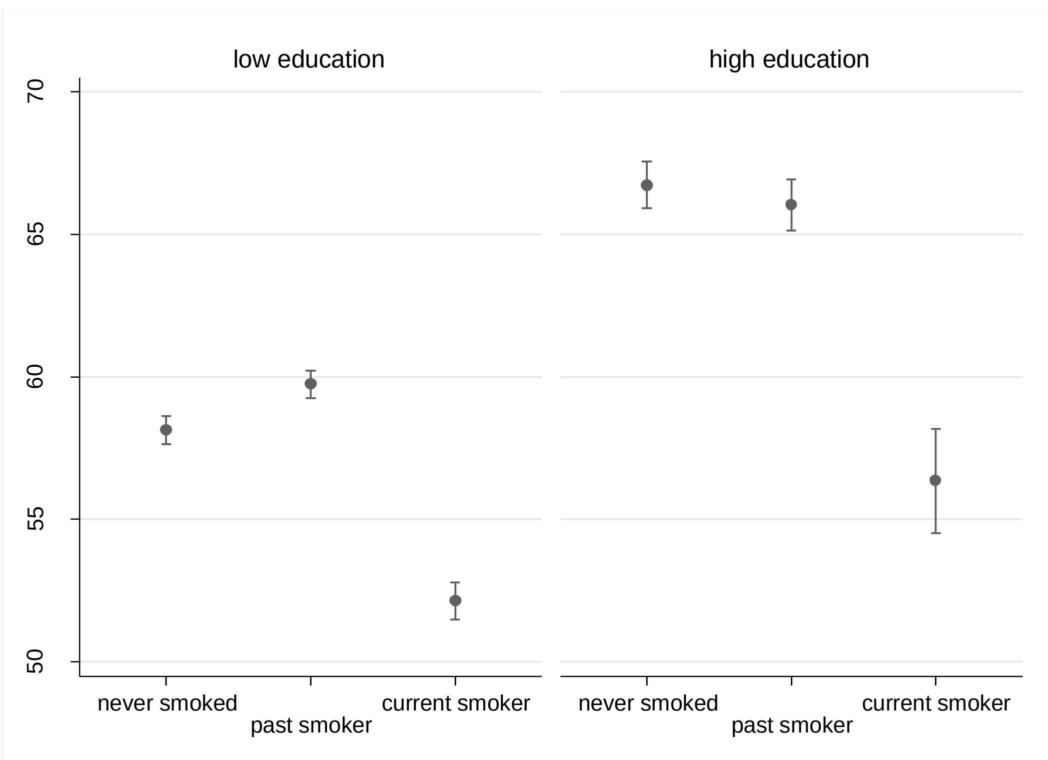
We examine SSP response patterns across three groups: those who have never smoked, those who had smoked sometime in the past (at least 100 cigarettes), but do not currently smoke, and those who currently smoke. 41.2% of respondents in our working sample have never smoked, 40.1% has smoked in the past, but is currently not smoking, and 18.7% is currently smoking. These percentages are significantly different between low and high educated. 21% of the sample has high education, of these only 9.7% are currently smoking, compared to 21.1% among the low educated.

Our working sample is on average aged 62.9 years, 33% of the respondents considered have poor health and 57% are females. The majority of the working sample is white (76.3%), followed by black (17.2%) and 6.4% identify themselves with other ethnicities.

In the first step of analyses we explore the SSP by sub-groups, i.e., between low and high educated by smoking status (never smoked, smoked in the past, currently smoking). Figure 1 shows the predictive margins for these sub-groups of the working sample, with the respective confidence intervals. These predicted SSPs are derived from linear regression models where the outcome variable is the individual SSP, the explanatory variables are smoking behaviour and education, and ethnicity, health and wave dummies are included as controls.

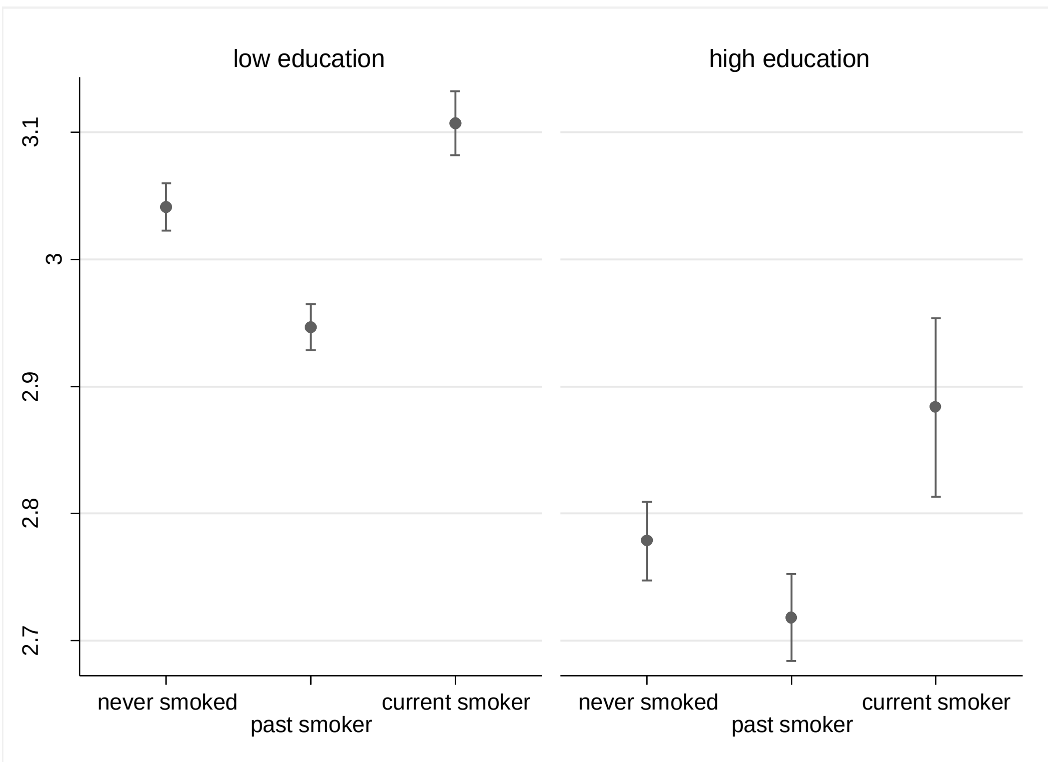
For each smoking status, high educated always report a significantly higher SSP that their low educated counterparts. Within the education sub-groups, we notice that respondents who never smoked report a significantly higher SSP than their counterparts who are currently smoking. However, among the low educated, the highest SSP is reported by past smoker, who among the high educated do not significantly differ from those who have never smoked in terms of SSP.

Figure 1. Predicted subjective survival probability by education and smoking behaviour



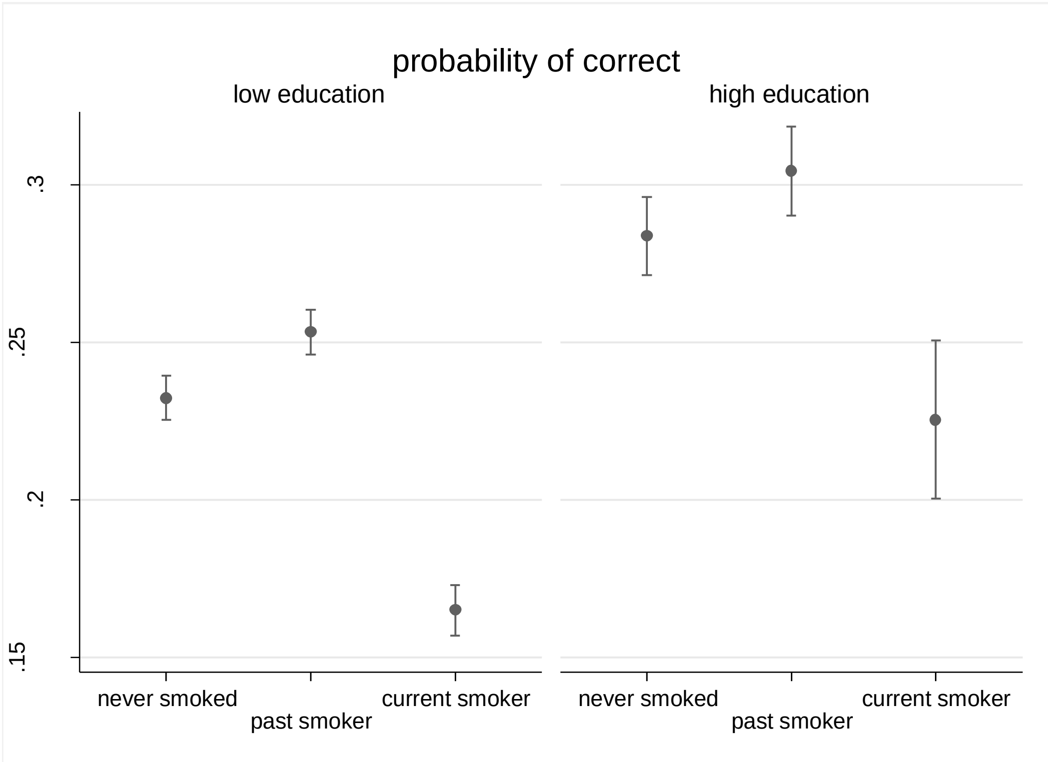
We then calculate, for each respondent, the "gap" between SSP and the estimated survival probability (SP) from the HRS data. Figure 2 shows the predictive margins obtained from this analysis. Note that here the logarithm of the absolute gap is reported, therefore the interpretation does not imply the direction of the gap (i.e., over or underestimation of OSP). For each smoking status, high educated always have a smaller gap between SSP and OSP than their low educated counterparts, meaning that the low educated are more likely to make mistakes in predicting their changes to survive. The patterns within the education sub-groups are however quite similar: those who are currently smoking are more likely to mistake their expectations of life. For them the gap between SSP and OSP is significantly bigger than for those who never smoked. Respondents who had smoked in the past are the best in predicting their survival probabilities However, their SSP is significantly closer to their OSP than it is their counterparts who never smoked only among the low educated.

Figure 2. Predicted logarithm of the gap between subjective and objective survival probabilities.



In order to be able to understand whether those who make more mistakes (i.e., they have a larger gap between SSP and SP) are under- or overestimating their survival probabilities, we carried out additional multinomial analyses (see Figures 3, 4, and 5). In these analyses the reference category of the outcome is the represented by a corrected subjective estimation of survival probability defined as a difference between SSP and OSP of maximum two percentage points in absolute terms. For each smoking status, high educated are more likely to correctly predict their survival probabilities than their low educated counterparts.

Figure 3. Probabilities of being correct in estimating survival probabilities



Within education groups it is evident a clear better prediction of those who have never smoked as compared to the currently smokers. Especially among the low educated, people who had smoked in the past but do not currently smoke are better able to predict their survival probability (Figure 3). But once we look at respondents who “make mistakes” in their SSP, it is very interesting to see that respondents who never smoked are the most likely to underestimate their survival probabilities (Figure 4); while currently smoking respondents are the most likely to overestimate their survival probabilities (Figure 5). Respondents who smoked in the past, but not currently smoking, hold an in-between position in both cases. Both among the low educated and among the high educated, there are statistically significant differences in the probabilities of under- and overestimating survival.

Figure 4. Probabilities of underestimating survival probabilities

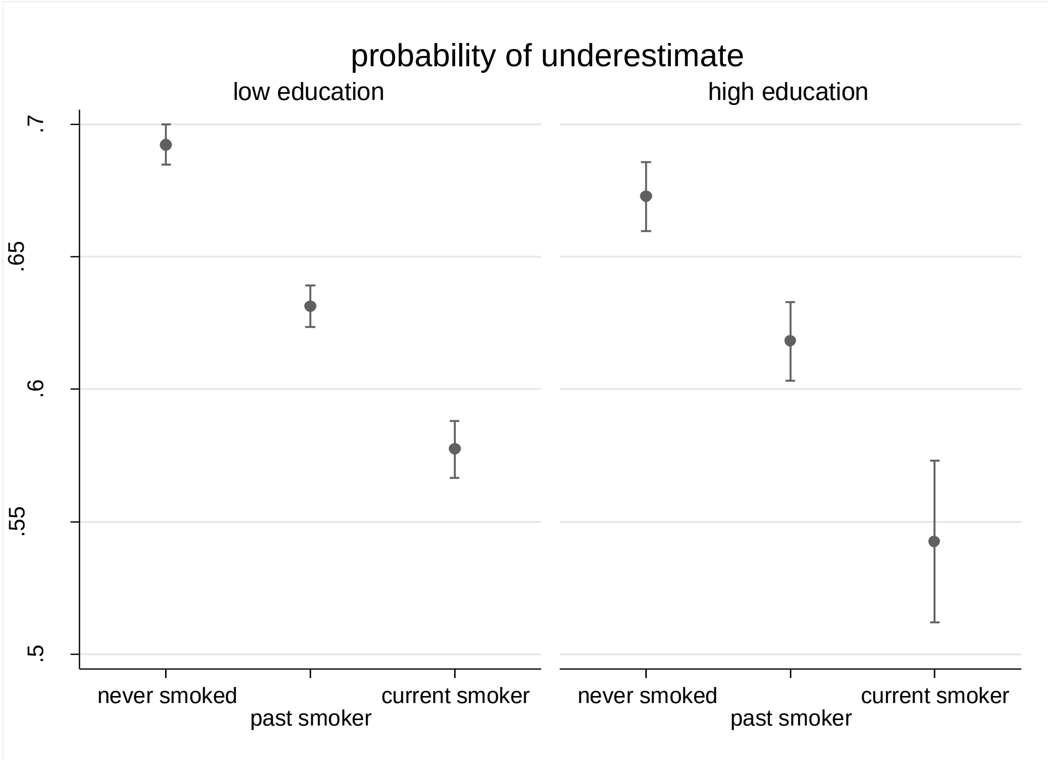
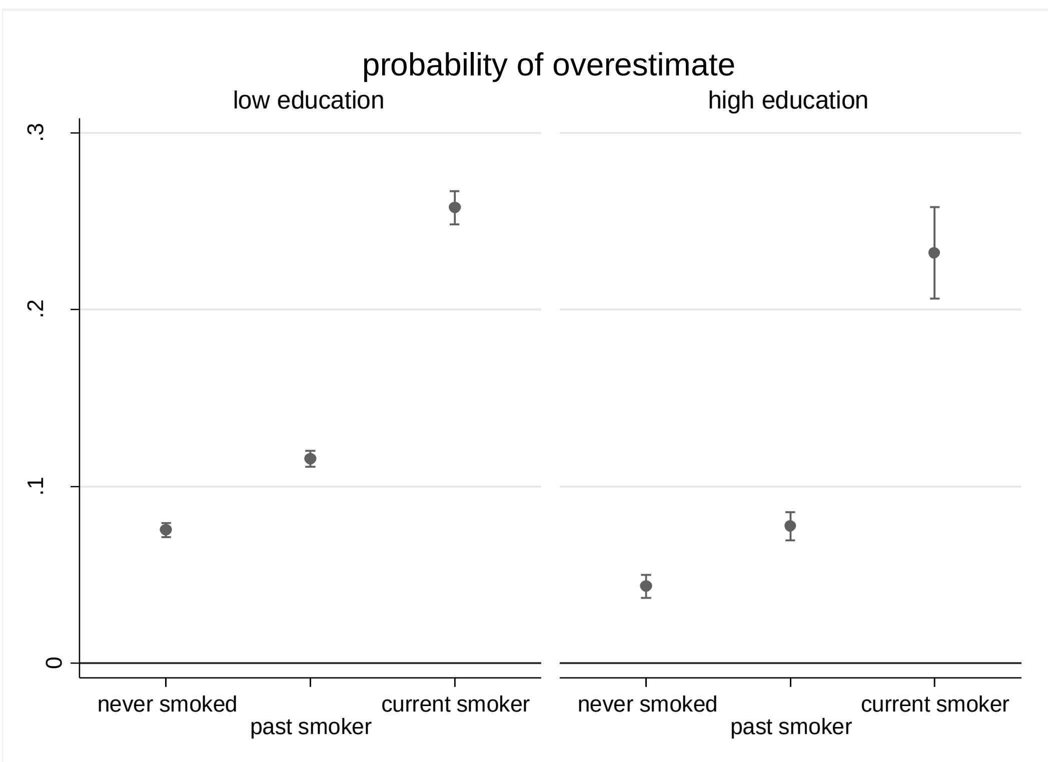


Figure 5. Probabilities of overestimating survival probabilities



Conclusions

Subjective survival probabilities (SSP) have been found to be a good predictor of mortality. In this paper we analyzed whether smoking behaviour and education influence the ability of people to predict their survival probability. Using longitudinal data from the Health and Retirement Study we compared subjective survival probabilities obtained by education and smoking status. Second, we used a Gompertz survival model to study how educational attainment and smoking behaviour affect the objective survival probability. Third, we compared subjective and objective survival probabilities.

We found that people currently smoking report lower survival probabilities especially if they are low educated. This is consistent with real mortality data that show higher mortality among these groups. When comparing subjective and objective survival probabilities we found that irrespectively of the smoking status, high educated people are more likely to correctly predict their survival probabilities than their low educated counterparts. Within education groups, people who smoked in the past are the best at predicting their mortality. Interestingly, those who currently smoke show the highest probability to incorrectly overestimate their survival probability (i.e., to underestimate the negative effect of smoking on mortality).

References

1. Manski CF. Measuring expectations. Econometrica. 2004;72(5):1329–76.

2. Perozek M. Using Subjective Expectations to Forecast Longevity: Do Survey Respondents Know Something We Don't Know? Demography. 2008 Feb;45(1):95–113.

3. Hurd MD, McGarry K. Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study. J Hum Resour. 1995 Jan 1;30:S268–92.

4. Hurd MD, McGarry K. The Predictive Validity of Subjective Probabilities of Survival*. Econ J. 2002 Oct 1;112(482):966–85.

5. Siegel M, Bradley EH, Kasl SV. Self-Rated Life Expectancy as a Predictor of Mortality: Evidence from the HRS and AHEAD Surveys. Gerontology. 2003;49(4):265–71.

6. Doorn C van, Kasl SV. Can Parental Longevity and Self-Rated Life Expectancy Predict Mortality among Older Persons? Results from an Australian Cohort. J Gerontol B Psychol Sci Soc Sci. 1998 Jan 1;53B(1):S28–34.

7. Smith V, Taylor DH, Sloan F. Longevity Expectations and Death: Can People Predict Their Own Demise? Am Econ Rev. 2001;91(4):1126–34.

8. Elder TE. The Predictive Validity of Subjective Mortality Expectations: Evidence From the Health and Retirement Study. Demography. 2012 Nov 15;50(2):569–89.

9. Hamermesh D. Expectations, Life Expectancy, and Economic Behavior. Q J Econ. 1985;100(2):389–408.

10. Salm M. Subjective mortality expectations and consumption and saving behaviours among the elderly. Can J Econ Can Déconomique. 2010 Aug 1;43(3):1040–57.

11. Carbone JC, Kverndokk S, Røgeberg OJ. Smoking, health, risk, and perception. J Health Econ. 2005;24(4):631–53.

12. Scott-Sheldon LAJ, Carey MP, Vanable PA, Senn TE. Subjective Life Expectancy and Health Behaviors among STD Clinic Patients. Am J Health Behav. 2010;34(3):349–61.

13. Sanderson WC, Scherbov S. The characteristics approach to the measurement of population aging. Popul Dev Rev. 2013 Dec 1;39(4):673–85.

14. Sanderson WC, Scherbov S. Measuring the Speed of Aging across Population Subgroups. PLoS ONE. 2014 May 7;9(5):e96289.