SHARE ERIC
SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

SHARE WORKING PAPER SERIES

The educational gradient in life expectancy in Europe: preliminary evidence from SHARE

Radim Boháček, Laura Crespo, Pedro Mira, Josep Pijoan-Mas

Working Paper Series 22-2015

SHARE-ERIC | Amalienstr. 33 | 80799 Munich | Germany | share-eric.eu

mea | MAX-PLANCK-INSTITUT FÜR SOZIALRECHT UND SOZIALPOLITIK
Munich Center for the Economics of Aging
Horizon 2020
Federal Ministry of Education and Research
DFG
NIH National Institute on Aging

The Educational Gradient in Life Expectancy in Europe: Preliminary Evidence from SHARE*

Radim Boháček
CERGE-EI

Laura Crespo
Bank of Spain

Pedro Mira
CEMFI

Josep Pijoan-Mas
CEMFI and CEPR

May 19, 2015

1 Introduction

Economic inequality manifests itself very prominently in inequality of health outcomes in general, and inequality of lifespans in particular.[1] In this paper we measure the mortality gradient of education in several SHARE countries. In particular, we compute survival functions from age 50 for men and women distinguishing between high and low educated individuals. We complement the study by computing the associated differences in life expectancies by education, and we compare the results to the ones obtained with ELSA for England.

We focus on education for several reasons. First, it is arguably the best approximation to lifetime income and the only measure of socio-economic status that does not change over the life cycle. Second, it has been shown (at least in the US) that it is the measure of

*ELSA data were made available through the UK Data Archive. ELSA was developed by a team of researchers based at the NatCen Social Research, University College London and the Institute for Fiscal Studies. The data were collected by NatCen Social Research. The funding is provided by the National Institute of Aging in the United States, and a consortium of UK government departments co-ordinated by the Office for National Statistics. The developers and funders of ELSA and the Archive do not bear any responsibility for the analyses or interpretations presented here.

[1] Kitagawa and Hauser (1973) were among the first to show that mortality rates in 1960 in the United States were inversely related to education and income.

socio-economic status that really matters for mortality.[2] Third, it is arguably the measure of socioeconomic status with lowest measurement error and the most widely available one.

We have relatively good measures of the education gradient in the US thanks to the death certificates, but also because of good survey data. In particular, the National Longitudinal Mortality Study (NLMS) tracks mortality of individuals observed in the Current Population Survey or the Census in the late 70's and early 80's. For more recent measures, the long panel dimension of the Health and Retirement Study (HRS) has also proved useful. Overall, the results show large life expectancy gaps at different ages between college educated and non college educated individuals.

There are however fewer studies for European countries because of data limitations. First, the death registers are less useful in Europe because they do not record data on education. A few studies have linked the death registers with census data to obtain education and sex-specific death rates in the 1980'a and 1990's (see Avendano, Kok, Glymour, Berkman, Kawachi, Kunst, and Mackenbach (2011) and Mackenbach, Stirbu, Roskam, Schaap, Menvielle, Leinsalu, and Kunst (2008)). This is possibly the best information we have today about the education gradient of mortality in Europe. However, samples are not always nationally representative and the resulting data are not homogenized across countries, which makes cross-country comparisons more problematic. Regarding survey data, only the European Community Household Panel (ECHP) –which covers the period between 1994 and 2000– has been used (see Majer, Nusselder, Mackenbach, and Kunst (2010)).

SHARE is an interesting source of data for this topic. First, it explicitly attempts to obtain end-of-life interviews so it is arguably less likely to under-report deaths than other survey data. Second, it is based on nationally representative country samples of people aged 50 and older, which allows to compare the gradient across countries using a harmonized dataset. Finally, SHARE provides a unique wealth of socio-economic and health variables that should allow researchers to broaden the analysis of the socioeconomic gradient of survival and healthy survival. The possible problems with SHARE are the small sample sizes for every country, and the potential biases in sample design, sample collection, or sample retention inherent in survey data. This paper is a first exploration of the potential of the SHARE data for this kind of analysis.

Table 1: Sample sizes

Country	# Waves	Males				Females			
		# obs	edu_1	edu_2	edu_3	# obs	edu_1	edu_2	edu_3
South									
Italy	5	1722	44.7	69.6	90.8	1985	57.7	76.2	92.0
Spain	5	1759	59.1	79.6	89.6	2079	66.7	84.8	93.2
East									
Czechia	4	2058	7.5	48.3	83.4	2798	24.9	50.5	87.4
Estonia	2	2278	6.2	34.3	67.4	3394	6.0	29.6	61.7
Poland	3	851	38.8	39.2	84.0	1039	51.2	51.5	90.7
North									
Denmark	5	1343	9.6	13.6	62.0	1496	18.3	27.9	60.9
Sweden	5	1448	37.4	52.8	70.3	1633	33.4	50.2	68.0
Center									
Austria	5	1998	9.0	14.4	66.6	2645	17.5	33.8	76.1
France	5	2431	34.8	41.1	78.4	2991	43.2	53.1	81.9
England	7	6380	33.5	57.6	65.55	6736	46.1	68.9	76.16

Notes: Columns labeled edu_x contain percentage of population with selected education category. edu_1 denotes ISCED-97 codes 0 and 1; edu_2 denotes ISCED-97 codes 0, 1, and 2; edu_3 denotes ISCED-97 codes 0, 1, 2, and 3.

2 Building a sample for survival analysis with SHARE

We compute Kaplan and Meier (1958) estimates of survival functions for high and low educated individuals, separately for men and women in SHARE countries and ELSA. We separate men and women because there is ample evidence of large gender-specific gradients in the US and elsewhere.

Our sample period is 2004 to 2013, which covers up to 5 waves although not all countries provide data for all of them. The sampling universe is the civilian non-institutionalized population aged 50 or older, although individuals are kept in the survey if they move into a nursing home. The data structure that we use is as follows. We need to observe individuals for at least two waves even if they are not consecutive waves. Then, every observation

[2]See Pijoan-Mas and Ríos-Rull (2014)

is an individual with the date at which we observe her for the first time, the date at which we observe her for the last time, and the alive/dead status the last time we observe her. We drop individuals with missing data for education and with initial age below 50 or above 90. Of course we also need the gender, and the age of the individual at the first observation.

There are two concerns when choosing the countries to use for this note. The first concern is that country samples are relatively small, which makes estimates of survival functions by education type quite noisy. The second concern is the quality of the country samples for survival analysis. Specifically, some country samples produce life tables that are substantially different from the population life tables obtained from Eurostat. Therefore, we will focus on countries that produce aggregate life tables that are reasonably close to the population tables.[3] Table 1 reports the sample sizes for each country included in this note.

We group respondents into just two education categories,"high" and "low" because our country-gender samples are too small for non-parametric estimation with more categories. We may consider different groupings depending on the ISCED-97 code that is used as threshold. Dummy variables edu_1, edu_2, and edu_3 are set to 1 if the respondent's schooling corresponds to ISCED-97 codes 0-1, 0-2 or 0-3 respectively. The thresholds are primary (ISCED-97 =1), lower-secondary (ISCED-97=2) o upper-secondary (ISCED-97=3). In Table 1 we report the fraction of low educated individuals according to each definition. As it can be seen, there is a large heterogeneity in the distribution of education across country-gender cells. In Southern countries around 3/4 of the population over the age of 50 has education corresponding to lower-secondary or less. In Austria and Denmark it is less than 1/4.

Given the heterogeneity of the education distribution across countries, we choose the threshold to be country and gender specific. Our choice reflects the fact that the partition of a country population in two socio-economic groups of roughly the same size happens at different levels of education in different countries. In order to maximize sample size across the two education categories in the estimates of survival functions obtained in the next section, we select the dummy variable such that the threshold is closest to the median of the distribution within country-gender cell. For instance, in cell Estonia-males "low" education corresponds to edu_2 whereas for for cell Estonia-females it corresponds to edu_3.

[3]The fit of SHARE-based life tables to Eurostat tables was good or very good for Spain, Poland, Italy (females), Estonia (males) and Denmark (females). At the other end, it was very poor for Germany, the Netherlands, Switzerland and Belgium. We exclude the latter countries from our analysis. At the margin we include France where the fit was quite poor for males but not so bad for females, so this caveat should be kept in mind. See Online Appendix for more details.

Table 2: Life expectancy differences at age 50

Country	Males		Females	
	edu_i	l.e.	edu_i	l.e.
South				
Italy	1	2.1	1	3.1
Spain	1	1.5	1	-1.5
East				
Czechia	2	1.7	2	4.3
Estonia	2	8.4	3	2.7
Poland	2	5.7	2	2.4
North				
Denmark	3	1.4	3	1.8
Sweden	2	1.9	2	1.4
Center				
Austria	3	5.3	2	-1.5
France	2	5.0	2	1.9
England	2	3.3	1	3.2

Notes: Column edu_i indicates the education definition used for the low education category, see footnote in Table tab:descriptive.

3 Education and survival

Life expectancy differences at age 50 computed for our samples from selected SHARE countries and England are presented in Table 2. The underlying KM survival functions with 95% confidence bands are shown in Figures 1 to 10. Although the life-expectancy gradients that we report would not be statistically significant at the 95% level for many country-gender cells, our results conform to the patterns found in the literature on education-related life expectancy differentials in general as well as with respect to gender and country areas.

First, we find that lower education attainment is associated with higher mortality rates. With the exception of Austrian and Spanish women, higher education carries

an important longevity premium.[4] The education advantage in life expectancy has been documented by Meara, Richards, and Cutler (2008) and Pijoan-Mas and Ríos-Rull (2014) for the United States as well as by Avendano, Kok, Glymour, Berkman, Kawachi, Kunst, and Mackenbach (2011) or Majer, Nusselder, Mackenbach, and Kunst (2010) for European countries.

Second, there is substantial heterogeneity across Europe. We find the largest inequalities in mortality in Eastern countries, which is consistent with some findings in the literature.[5] For instance, the education premium in Estonia is up to 8.4 years for males, but we also find large large gradients for Poland and the Czech Republic. Then, we find relatively small education-based life expectancy differentials in Northern Europe, that is in countries with egalitarian and generous welfare policies, smaller income inequalities and lower poverty rates. This is in contrast to the results in Mackenbach, Stirbu, Roskam, Schaap, Menvielle, Leinsalu, and Kunst (2008). However, the education gradients are also relatively small in Southern countries as noted in previous studies that used population data for selected sub-regions or cities.

Third, in general but not always, the educational differences in life expectancy are smaller for women than men. This matches the European Comission (2013) comparison of the educational gradient in life expectancies for a sample of 14 EU countries. Using Eurostat data, they report a larger education premium for men than for women, and larger in countries with shorter life expectancies. For Estonia and France, the large gender difference in the educational gradient is in line with Avendano, Kok, Glymour, Berkman, Kawachi, Kunst, and Mackenbach (2011). In contrast, we find a larger education premium in life expectancy for females than for males in the Czech Republic, Denmark, and Italy.

4 Conclusion

We have used education as a social stratification variable in this note, because it is relatively easy to measure and it can be made comparable across countries. Education attainment has been identified as a major determinant of general living conditions and occupation and related to different patterns of smoking, obesity, excessive alcohol consumption, access and use of health care, and other important lifestyle choices which have an impact on life-expectancy. In spite of our small sample sizes we obtain non-parametric

[4]Looking at Figure 2 we see that the survival advantage for less educated women in Spain appears only after age 75, which may reflect some cohort effect.

[5]See Avendano, Kok, Glymour, Berkman, Kawachi, Kunst, and Mackenbach (2011) and European Comission (2013)

estimates of survival functions by education and of the education premia in life-expectancy which are in line with the findings from non-harmonized data in the existing literature.

On the other hand, it should be noted that internationally comparable education levels based on the ISCED classification do not take into account the quality of education in individual countries. Most importantly, education captures only one dimension of socio-economic characteristics while a more comprehensive set of measures would also include marital and labor market status, occupation, income, wealth or poverty index that might be more relevant for mortality at older ages and be more sensitive to health care policies across different countries. We plan to address these issues in our future work, together with the dynamic evolution of individual characteristics over time, in parametric hazard survival models with time varying stochastic endogenous covariates.

As the magnitudes of life expectancy differentials are very variable across countries and regions, we believe that SHARE data provides a unique opportunity for identifying their determinants. Policy interventions targeting these determinants might increase longevity and quality of lives of the European population.

References

Avendano, M., R. Kok, M. Glymour, L. Berkman, I. Kawachi, A. Kunst, and J. Mackenbach (2011): "Do Americans Have Higher Mortality Than Europeans at All Levels of the Education Distribution?: A Comparison of the United States and 14 European Countries," in *International Differences in Mortality at Older Ages: Dimensions and Sources*, ed. by E. M. Crimmins, S. H. Preston, and B. Cohen, chap. 11. National Academies Press.

European Comission (2013): *Health inequalities in the EU - Final report of a consortium. Consortium lead: Sir Michael Marmot*. European Commission Directorate-General for Health and Consumers.

Kaplan, E., and P. Meier (1958): "Nonparametric estimation from incomplete observations," *Journal of the American Statistical Association*, 53, 457–481.

Kitagawa, E. M., and P. M. Hauser (1973): *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. Harvard University Press, Cambridge.

Mackenbach, J. P., I. Stirbu, A.-J. R. Roskam, M. M. Schaap, G. Menvielle, M. Leinsalu, and A. E. Kunst (2008): "Socioeconomic Inequalities in Health in 22 European Countries," *The New England Journal of Medicine*, 358, 2468–2481.

Majer, I., W. Nusselder, J. Mackenbach, and A. Kunst (2010): "Socioeconomic Inequalities in Life and Health Expectancies Around Official Retirement Age in 10 Western-European Countries," *Journal of Epidemiology and Community Health*.

Meara, E., S. Richards, and D. Cutler (2008): "The Gap Gets Bigger: Changes In Mortality And Life Expectancy, By Education, 1981? 2000," *Health Affairs*, 27(2), 350–360.

Pijoan-Mas, J., and J. V. Ríos-Rull (2014): "Heterogeneity in Expected Longevities," *Demography*, 6(51), 2075–2102.

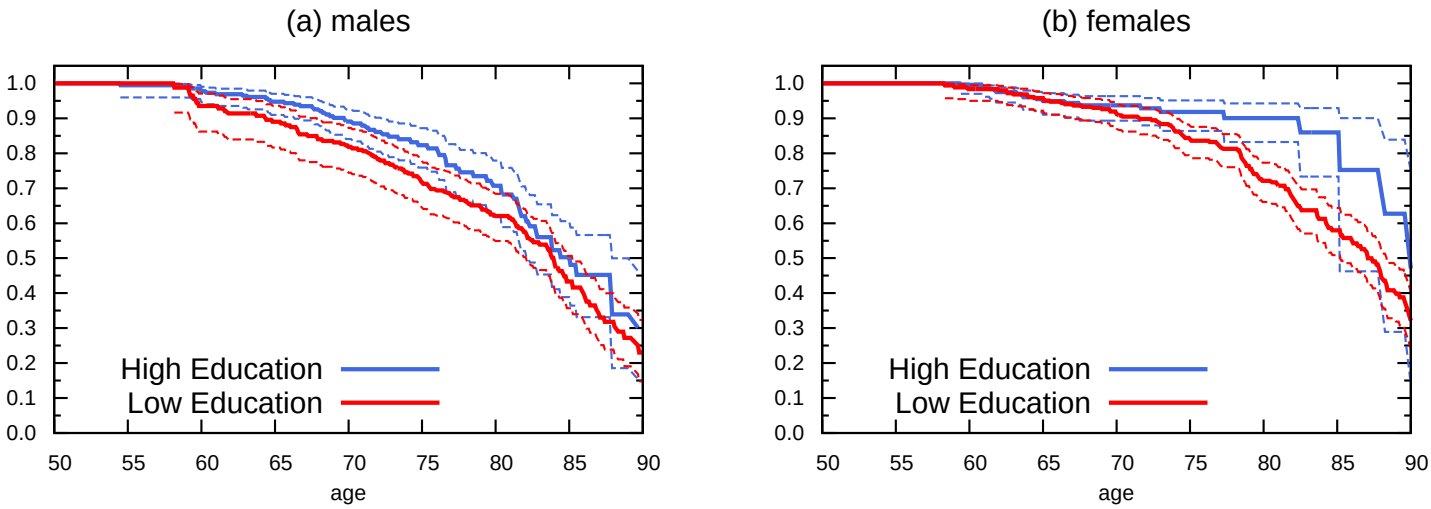

Figure 1: Survival function. Italy.

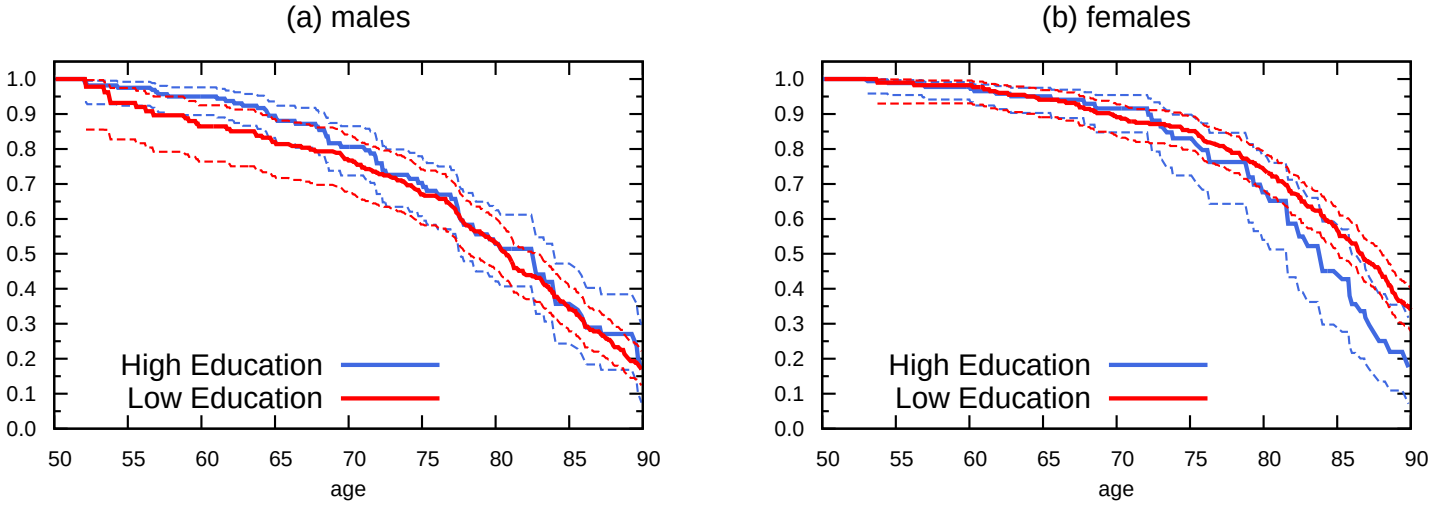

Figure 2: Survival function. Spain.

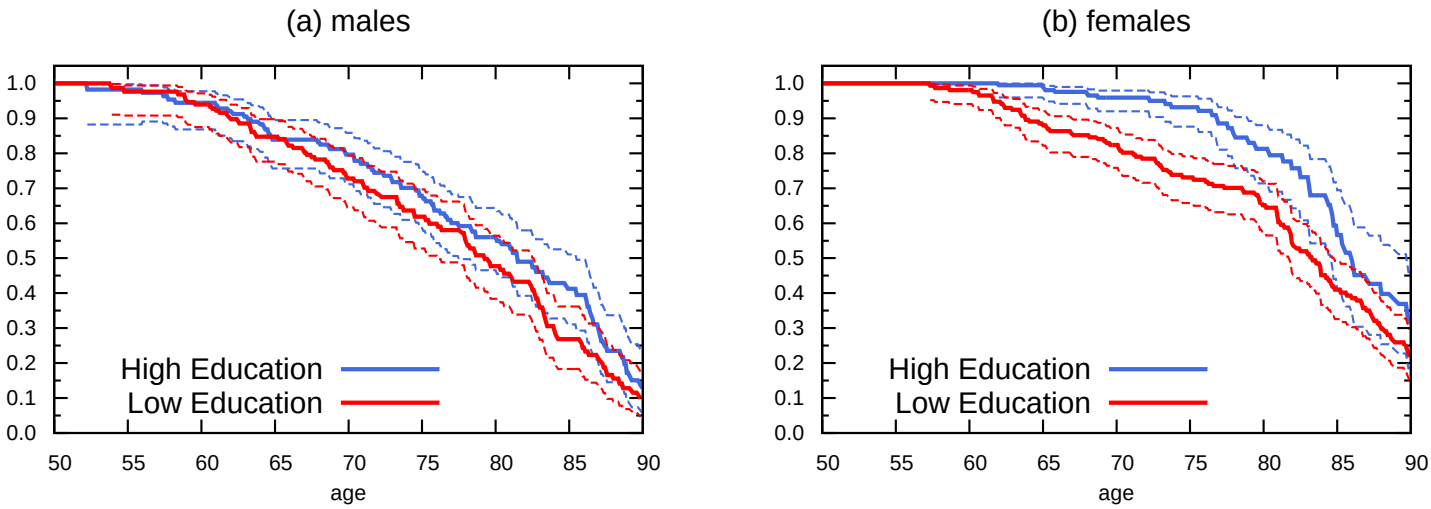

Figure 3: Survival function. Czechia.

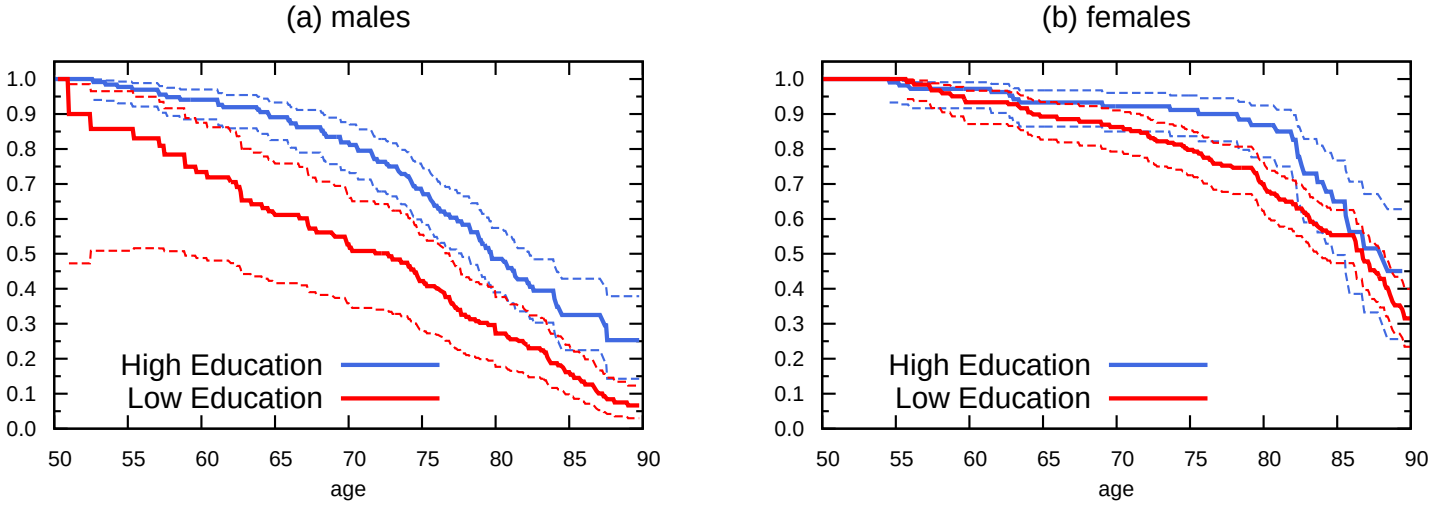

Figure 4: Survival function. Estonia.

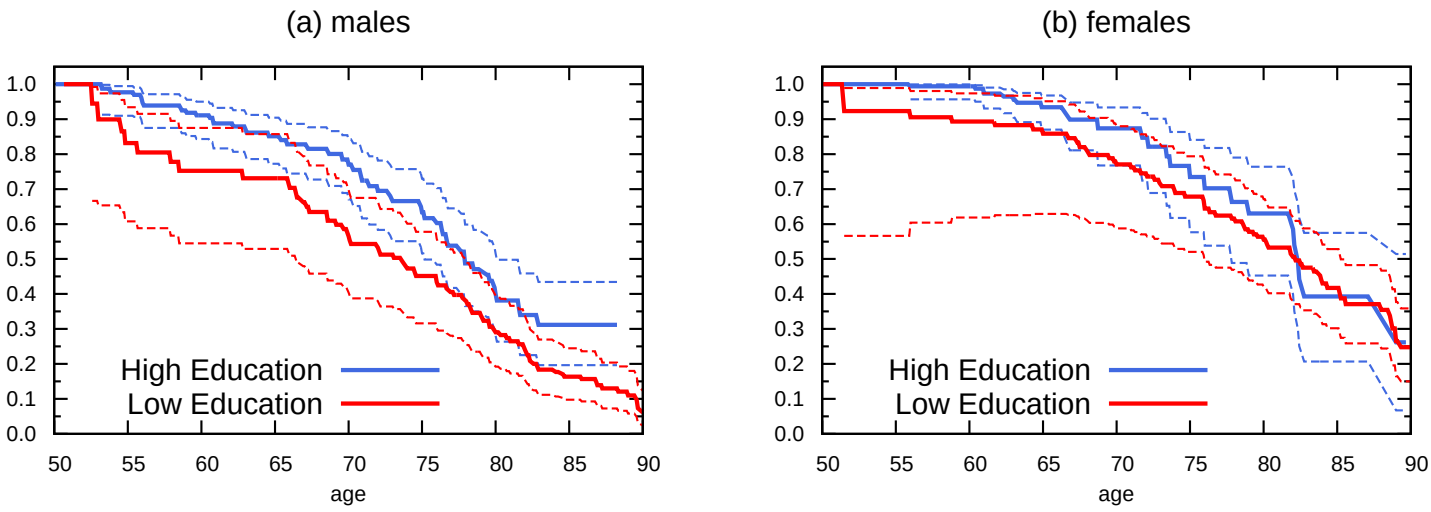

Figure 5: Survival function. Poland.

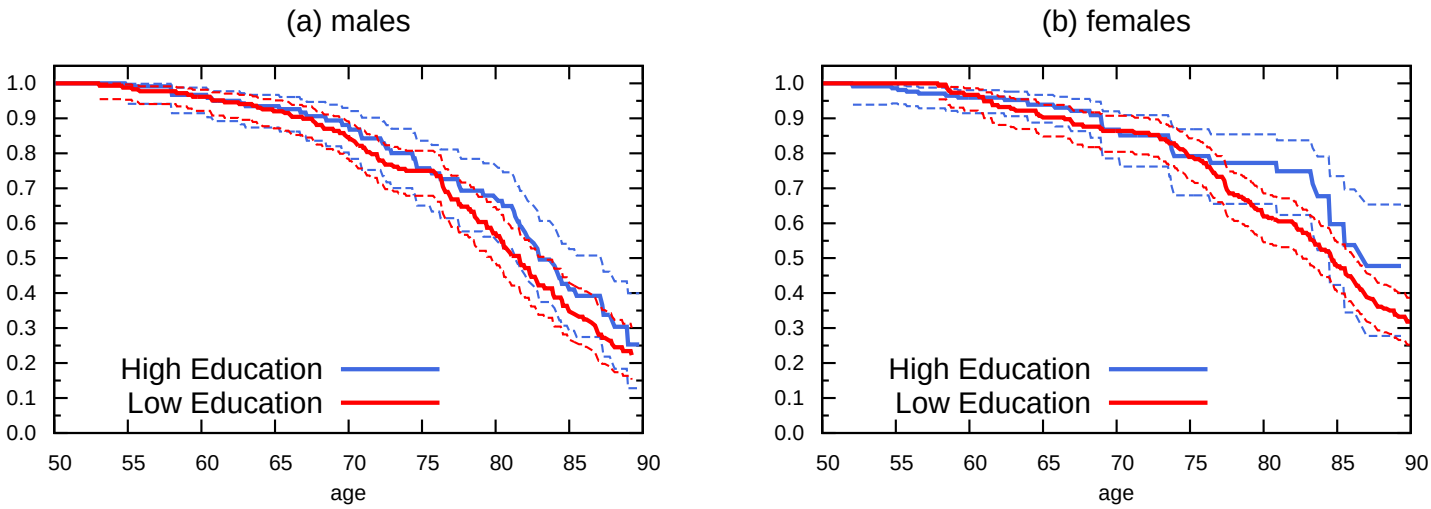

Figure 6: Survival function. Denmark.

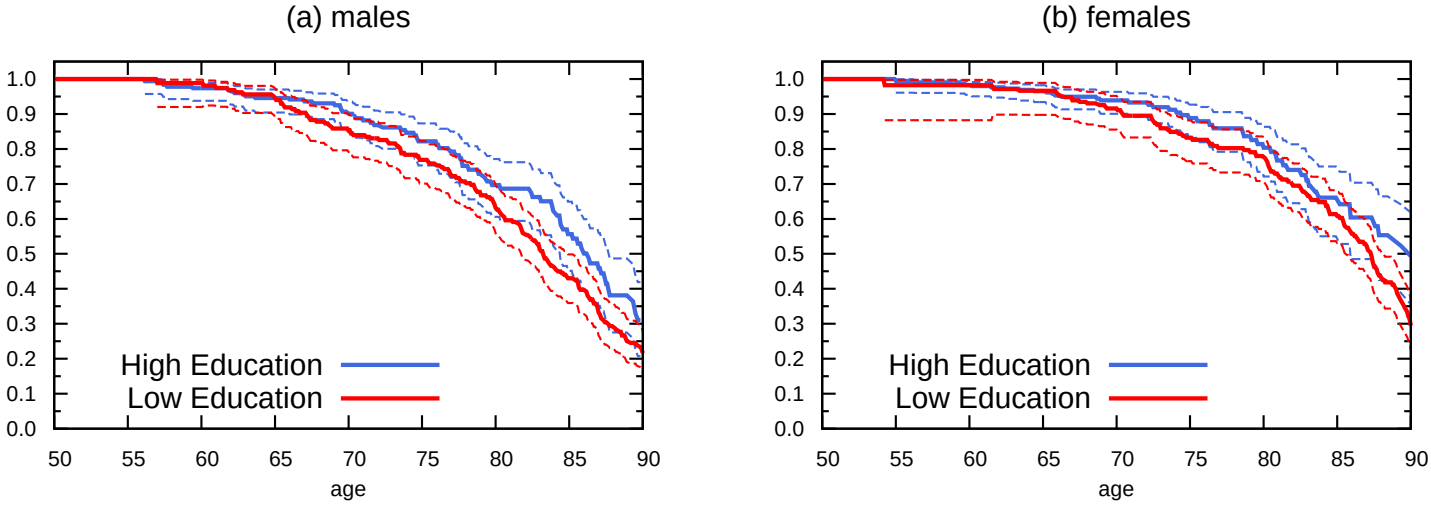

Figure 7: Survival function. Sweden.

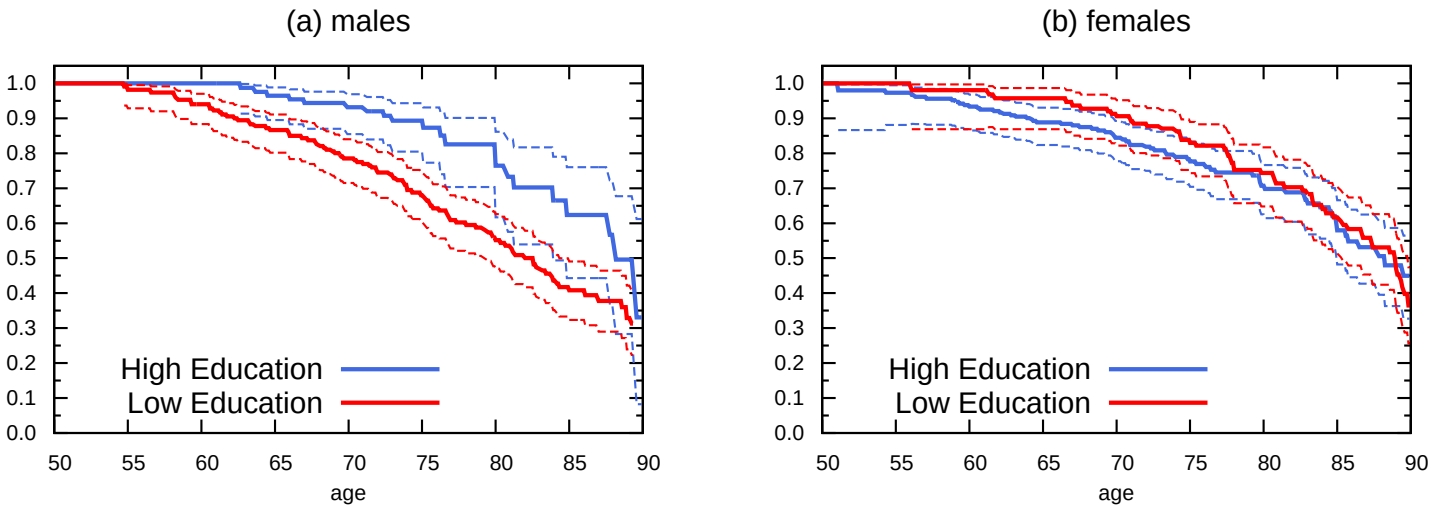

Figure 8: Survival function. Austria.

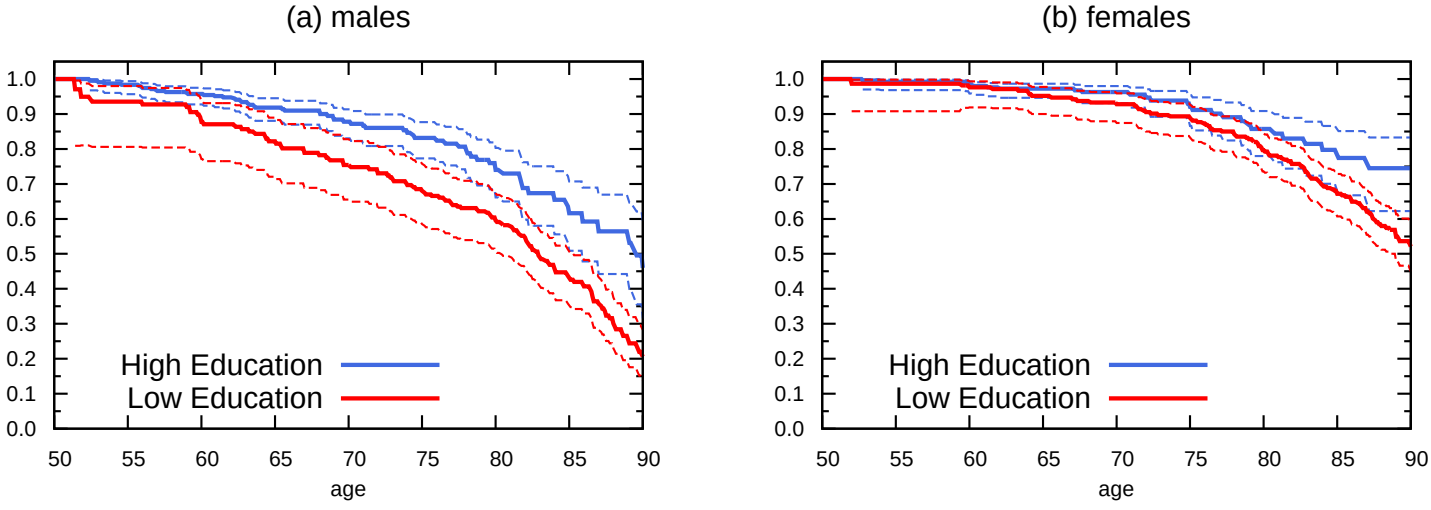

Figure 9: Survival function. France.

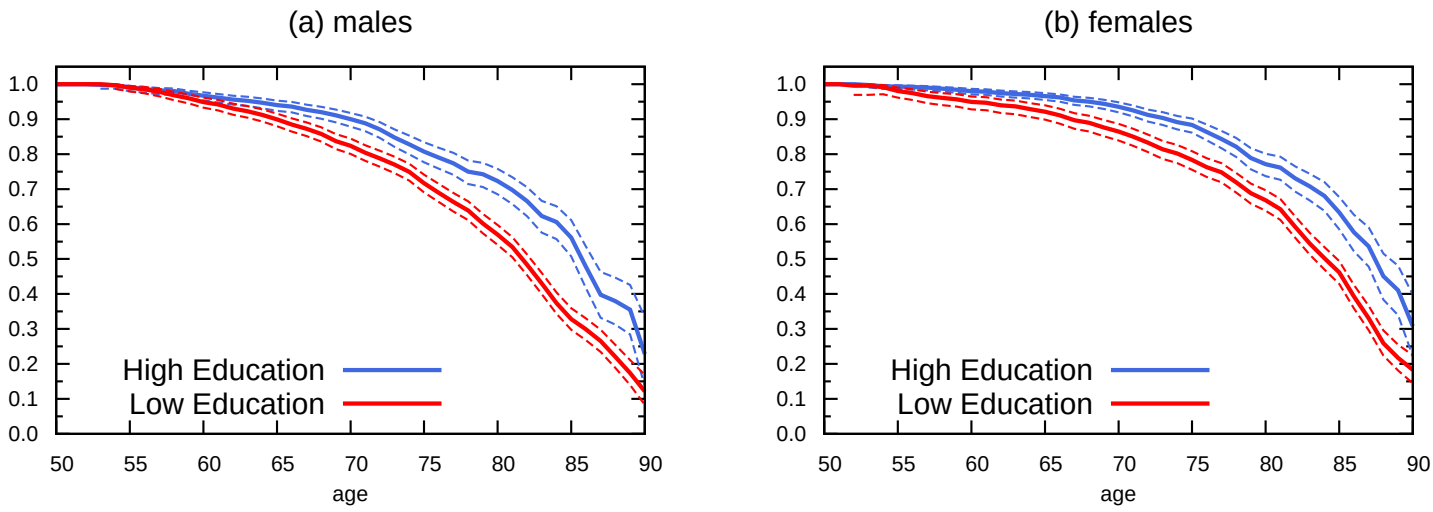

Figure 10: Survival function. England.

Appendix

5 Appendix A: Data

For all countries, we use data for individuals aged between 50 and 90.9 years in the SHARE and the English Longitudinal Study of Ageing (ELSA) longitudinal samples. Eligible individuals were interviewed in at least two waves and we use exit interviews to construct living status indicators (alive/dead) in each wave for each individual. We include those individuals with available information for the variables used in the analysis: highest level of education achieved (ISCED-97 code), age at the first interview, age at the date of death for those who died during the period under study, and age at the last interview for those who are still alive.

5.1 SHARE Data

SHARE longitudinal countries including Austria, Sweden, the Netherlands, Spain, Italy, France, Denmark, Belgium, the Czech Republic, Poland, Slovenia, and Estonia from Wave 1 (2004) to Wave 5 (2011). The sample consists of 61,203 individuals, 26,895 males and 34,308 females. Among males, 2,664 individuals (9.91 percent) died between Wave 1 and Wave 5 whereas among women, 2,262 (6.59 percent) did. Low education includes no schooling (ISCED-97 code 0), primary education (ISCED-97 code 1), and lower secondary education (ISCED-97 code 2).

5.2 ELSA Data

For the UK sample, we use ELSA data from waves 0 (1998) to 6 (2013).[6] There are 7,390 female respondents who were interviewed/contacted at least in two waves of which one can be the exit interview.[7] Between Wave 0 and 6, 1,004 females (13.6%) have died. Of age-eligible male 6,626 respondents, 1,118 have died (16.9%).

Education classification of ELSA respondents is based on Schneider (2008), Table 4 for revised ISCED-97 codes. Educational attainment of at most ISCED 1 includes UK categories of "No qualification", ISCED 2 includes additionally the NVQ1/CSE and other grade equivalent and NVQ2/GCE O level equivalent categories, and ISCED level 3 contains NVQ3/GCE A level equivalent category.

[6] ELSA does not provide information on the month of birth and on the month of death of a respondent. Respondents older than 90 years are grouped for confidentiality reasons.

[7] Contact considered as a proof that respondent is alive if it is coded as: Personal refusal, Proxy refusal, Broken appointment, Ill at home, Ill in hospital, Senile/incapacitated, Inadequate English or language difficulties, Refusal before interview, Refusal during interview, Broken appointment, Ill at home during survey period, Physically/mentally unable/incompetent, or In institution.

6 Appendix B: A comparison of Eurostat and SHARE-based life tables

The graphs that follow show KM survival functions obtained from SHARE samples and Eurostat survival functions for country-gender cells. In each graph:

- KM short: SHARE samples, individuals observed in consecutive waves.
- KM filling: SHARE samples, includes individuals recovered after skipping one or more waves ("Sample C").
- ucsvl_estat: Eurostat