A Appendix A: Data construction and details

The English Longitudinal Study of Ageing (ELSA) is available for download from the UK Data service (UKDS) at the following link:
`https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=200011#!/access-data`.
The data in this study required Special License Access and so cannot be shared but is accessible by request to ELSA.

A.1 Data construction

Construction and description of Frailty Index

We use the complete ELSA respondent sample over waves 1-9 as the basis for the sample in this study. Following Abeliansky, Erel, and Strulik (2020), we restrict the sample to observations fulfilling the following criteria:

- Place of birth must be in the United Kingdom
- Must be respondent in a given survey wave
- Age must be non-missing and between 50 and 90
- At least 35 out of the 50 items in the frailty index must be non-missing

The frailty index is generated following Rogers et al. (2017); we end up with 50 items. Summary statistics for these items are provided in Table A.1.

Additional variables

In addition, we include the following variables in our analysis:

- **Wealth**: Following Marshall et al. (2015), the wealth variable is the natural logarithm of the sum of financial and housing wealth for a given household. Negative and zero values are coded as 1 for the logarithm to be defined. We define sample splits based on the tertile in the wave-5-year-age-group wealth distribution in which an individual is in the first time they are recorded in the survey.
- **Education**: The analysis uses the internationally standardized education variable that records three levels of education: Less than secondary (38% in our sample), upper secondary and vocational training (47%), and tertiary education (15%).
- **Regions**: We obtained confidential data on individuals' region of residence, which we aggregated up to the NUTS-1 level (comprising 9 regions in England)

Table A.1: Summary statistics for items on frailty index

		count	mean	sd
	Mobility			
1	Some difficulty walking 100 yards	78,850	0.125	0.331
2	Some difficulty sitting for 2 hours	78,850	0.131	0.337
3	Some difficulty getting up from chair	78,850	0.249	0.433
4	Some difficulty climbing several flights of stairs	78,850	0.341	0.474
5	Some difficulty climbing one flight of stairs	78,850	0.145	0.352
6	Some difficulty kneeling/stooping/crouching	78,850	0.372	0.483
7	Some difficulty reaching/extending arms	78,850	0.108	0.310
8	Some difficulty pushing/pulling large object	78,850	0.173	0.378
9	Some difficulty lifting 10lbs	78,850	0.229	0.420
10	Some difficulty picking up a dime	78,850	0.057	0.231
	ADLs/IADLs			
1	Some difficulty getting dressed	78,855	0.128	0.334
2	Some difficulty walking across room	78,855	0.034	0.181
3	Some difficulty taking a bath or shower	78,855	0.100	0.300
4	Some difficulty eating	78,855	0.023	0.149
5	Some difficulty going into/out of bed	78,855	0.060	0.237
6	Some difficulty using toilet	78,855	0.035	0.185
7	Some difficulty reading map	78,855	0.048	0.214
8	Some difficulty preparing hot meals	78,855	0.047	0.211
9	Some difficulty grocery shopping	78,855	0.089	0.284
10	Some difficulty using telephone	78,855	0.025	0.155
11	Some difficulty taking medications	78,855	0.023	0.149
12	Some difficulty managing money	78,855	0.031	0.173
13	Some difficulty doing work around house or garden	78,855	0.152	0.359
1	Poor self-reported health	78,855	0.152	0.359
	Depression			
1	CESD: Felt depressed	75,681	0.138	0.344
2	CESD: Felt everything was an effort	75,681	0.198	0.399
3	CESD: Sleep was restless	75,683	0.397	0.489
4	CESD: Not happy	75,506	0.099	0.299
5	CESD: Felt lonely	75,676	0.120	0.324
6	CESD: Does not enjoy life	75,509	0.092	0.289
7	CESD: Felt sad	75,645	0.192	0.394
8	CESD: Could not get going	75,644	0.200	0.400
	Self-reported conditions			
1	Ever had High blood pressure	78,854	0.416	0.493
2	Ever had heart problem	78,854	0.194	0.395
3	Ever had Diabetes	78,854	0.100	0.300
4	Ever had Stroke	78,854	0.046	0.209
5	Ever had Lung Disease	78,855	0.064	0.246
6	Ever had asthma	78,855	0.130	0.337
7	Ever had Arthritis	78,854	0.366	0.482
8	Ever had Cancer	78,855	0.096	0.295
9	Ever had Parkinson's disease	78,855	0.007	0.086
10	Ever had Psychiatric problem	78,855	0.100	0.300
11	Ever had Alzheimers	78,855	0.004	0.061
12	Ever had Dementia	78,855	0.011	0.106
	Cognitive measures			
1	Remembers day incorrectly	75,606	0.171	0.377
2	Remembers month incorrectly	75,712	0.022	0.147
3	Remembers year incorrectly	75,715	0.021	0.145
4	Remembers day of week incorrectly	75,729	0.019	0.135
5	Delayed recall: <=2/10 words	75,894	0.158	0.365
6	Immediate recall: <=4/10 words	75,792	0.196	0.397

A.2 STROBE-Flowchart: Sample Selection and sample size

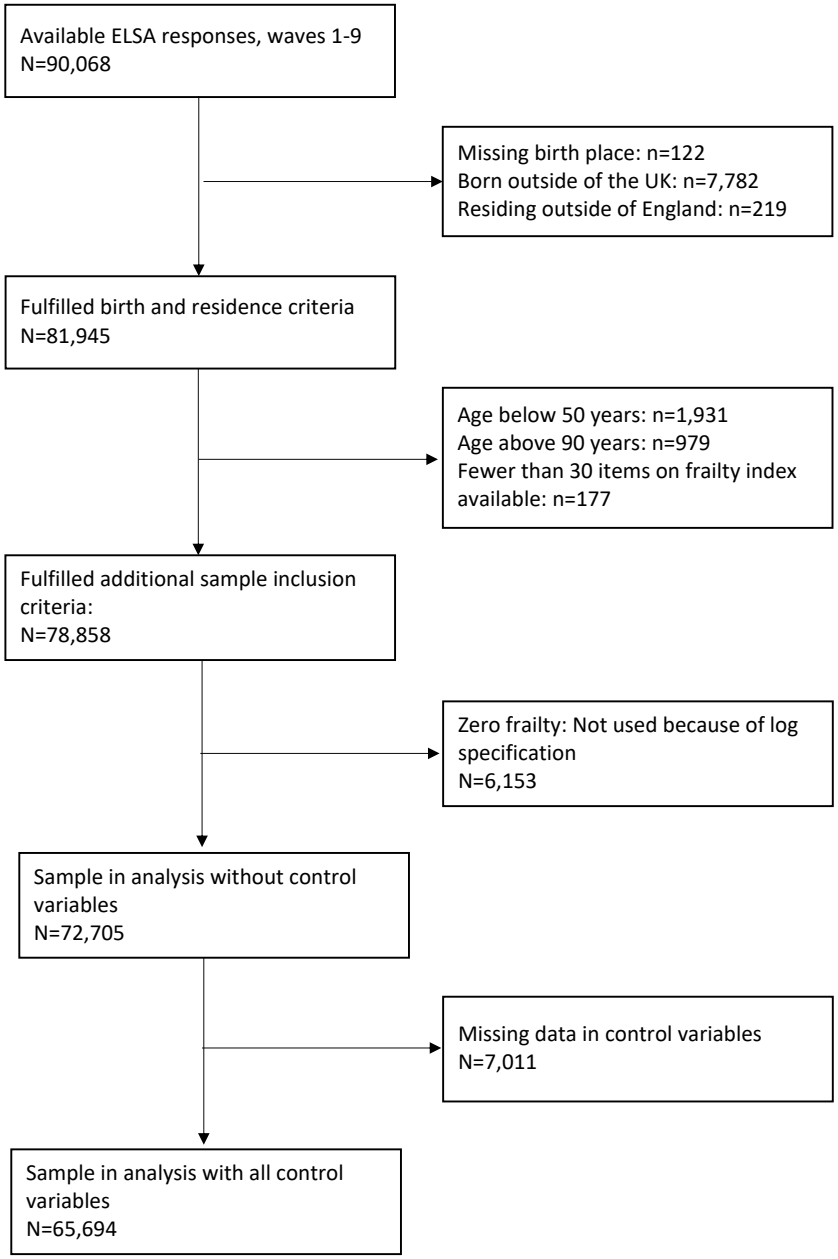

Figure A.1: Flow chart showing sample inclusion criteria and their successive impact on the sample size.

A.3 Frailty Index – Availability over time

Figure A.2 shows the availability of the 50 items in our frailty index for the nine waves of ELSA included in this study. Most of the items have very few missing values throughout all waves. The self-reported health measure is not included in wave 3. For the items on depression and on cognitive function, the share of missing observations is increasing over time, with around 6% missing values in waves 7, 8, and 9.

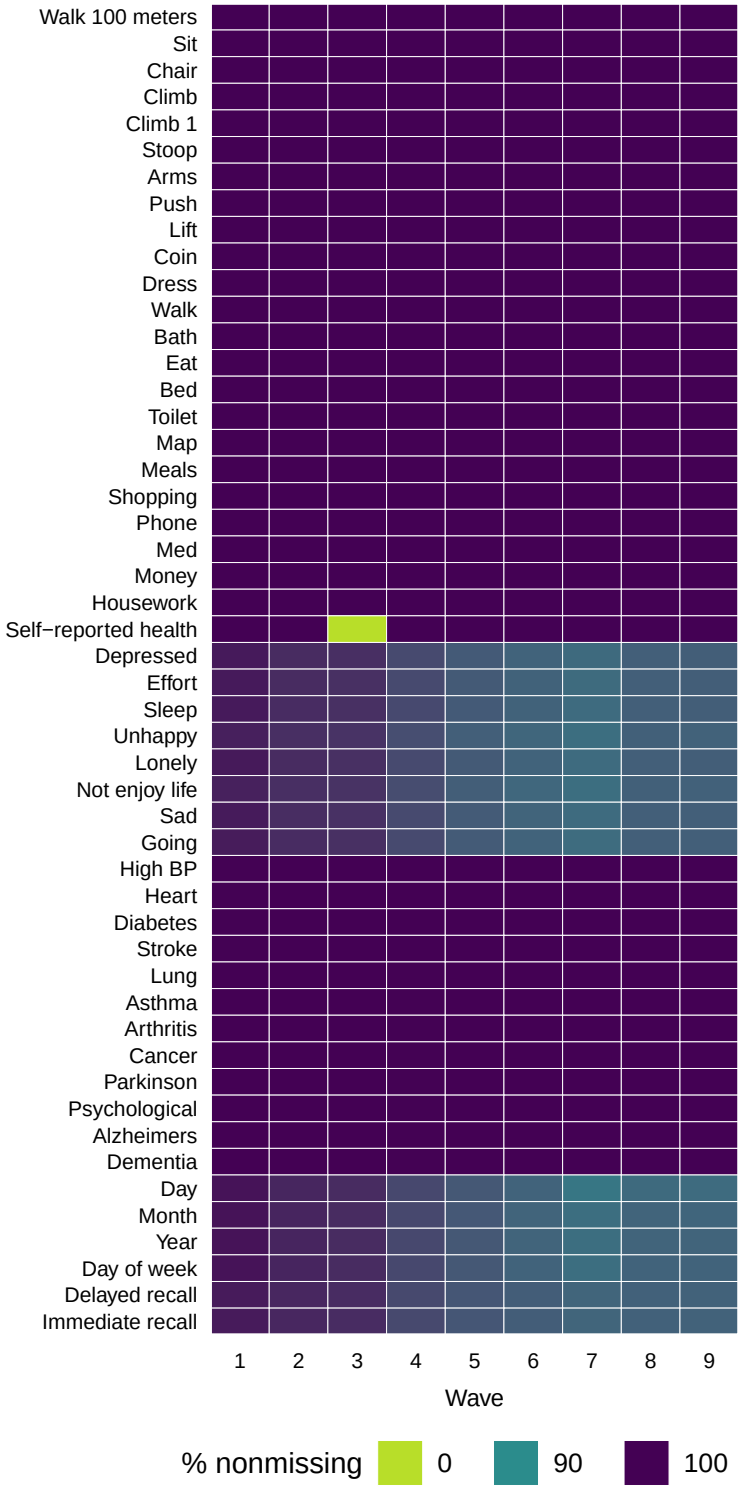

Figure A.2: Data availability of items in the frailty index.

B Appendix B: Robustness and additional analysis

B.1 Full regression results for determinants and speed of ageing

Table B.1 shows the full regression results from estimating equations 1 and 2. The corresponding results are visualized in the main text in Figure 1a and 1b.

Table B.1: Determinants of frailty and speed of ageing

	All		Gender		Education			Wealth		
	OLS	FE	Females	Males	Low	Middle	High	Low	Middle	High
Age	0.028***	0.039***	0.036***	0.043***	0.042***	0.038***	0.033***	0.039***	0.040***	0.040***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Female	0.159***									
	(0.013)									
Upper Secondary /	−0.201***									
Vocational Training	(0.016)									
Tertiary education	−0.358***									
	(0.021)									
Log HH wealth	−0.070***	−0.018***	−0.018***	−0.018***	−0.019***	−0.017***	−0.013**	−0.012***	−0.002	−0.054***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.006)	(0.002)	(0.027)	(0.017)
East Midlands	0.121***	0.088	−0.000	0.181*	0.102	0.210*	0.025	0.327	−0.045	0.011
	(0.027)	(0.077)	(0.104)	(0.109)	(0.210)	(0.111)	(0.140)	(0.202)	(0.151)	(0.100)
East of England	0.020	−0.006	0.002	−0.029	0.039	0.110	−0.230	0.142	−0.062	−0.003
	(0.024)	(0.061)	(0.084)	(0.087)	(0.103)	(0.100)	(0.161)	(0.168)	(0.118)	(0.101)
London	0.036	0.049	0.090	−0.005	0.136	0.106	−0.034	0.204	0.096	0.044
	(0.029)	(0.053)	(0.066)	(0.086)	(0.101)	(0.086)	(0.096)	(0.138)	(0.098)	(0.072)
North East	0.168***	0.277*	0.001	0.615***	0.517*	0.364	0.010	0.312	0.215	−0.125
	(0.030)	(0.143)	(0.141)	(0.208)	(0.280)	(0.257)	(0.165)	(0.419)	(0.133)	(0.134)
North West	0.099***	−0.071	−0.044	−0.091	0.077	0.138	−0.336**	0.263	−0.239	−0.034
	(0.024)	(0.091)	(0.107)	(0.147)	(0.189)	(0.162)	(0.132)	(0.269)	(0.205)	(0.132)
South West	0.041*	0.025	0.027	0.015	−0.025	0.084	0.024	0.004	0.016	0.039
	(0.025)	(0.048)	(0.068)	(0.068)	(0.081)	(0.079)	(0.087)	(0.163)	(0.086)	(0.068)
West Midlands	0.068***	0.030	0.043	0.005	0.014	0.131	−0.088	−0.060	−0.124	0.124
	(0.026)	(0.075)	(0.092)	(0.122)	(0.128)	(0.108)	(0.145)	(0.205)	(0.161)	(0.098)
Yorkshire & Humber	0.083***	0.142	0.082	0.197	0.202	0.139	0.255	0.413*	−0.011	0.096
	(0.025)	(0.091)	(0.114)	(0.137)	(0.214)	(0.140)	(0.160)	(0.221)	(0.163)	(0.125)
Observations	65,694	71,872	40,021	31,851	23,530	31,953	10,211	24,824	24,007	23,041
R-squared	0.216	0.118	0.109	0.130	0.141	0.111	0.089	0.118	0.108	0.106
Method	OLS	FE	FE	FE	FE	FE	FE	FE	FE	FE

Notes: Column (1) reports pooled OLS estimates. Other columns report fixed-effects estimates. Robust standard errors clustered at respondent level in parentheses. Significance levels: *10%, **5%, ***1%.
Age is demeaned at individual level. Education levels are: below upper secondary, upper secondary/vocational training, and tertiary education, with the first being the reference level in column (1). Reference region is South East. Wealth is split in tertiles relative to five-year age group in the first occurrence in the survey.

B.2 Comparing with literature

To provide further evidence for our finding that in England later cohorts are ageing better than earlier cohorts in contrast to Marshall et al. (2015) we replicate across three different waves some of the raw summary statistics that they present. Table B.2 repeats the message from Table 2 that for each age band, there have been improvements in mean frailty across waves. Table B.3 repeats, for our data, Table 1 in Marshall et al. (2015). Reading down the columns shows how frailty increases with age so that those who were 50-54 in wave 1 with mean frailty of 0.091 have mean frailty of 0.101 in wave 5, by which point they are aged 60-64. However, whilst this group have seen their frailty rise with age their mean frailty is below that of 60-64-year-olds in wave 1. Similar comparisons for other ages show the same cohort improvement.

Table B.2: Mean frailty index values by age group, waves 1-9

Wave	50-54	55-59	60-64	65-69	70-74	75-79	80+
1	0.091	0.106	0.118	0.126	0.149	0.171	0.223
5	0.094	0.097	0.108	0.122	0.142	0.170	0.241
9	0.083	0.096	0.115	0.110	0.131	0.160	0.216

Table B.3: Mean frailty index values by 2002 cohort group, as in Marshall, waves 1-9

Wave	50-54	55-59	60-64	65-69	70-74	75-79	80+
1	0.091	0.106	0.118	0.126	0.149	0.171	0.223
5	0.101	0.117	0.136	0.152	0.199	0.248	0.297
9	0.117	0.136	0.163	0.201	0.265		

B.3 Robustness

In order to test the robustness of our results, we performed the analysis: 1) using only individuals who survived throughout the sample, 2) restricting the analysis to individuals with at least 35 or 3) at least 40 non-missing items when constructing the frailty index, 3) restricting the analysis to observations between 50 and 80 years of age, or 5) between 55 and 90 years of age, 6) calculating the frailty index only using items with at least 99% coverage throughout the sample, 7) restricting the analysis to individuals which never had any missing value on the frailty index from the first to the last time they were surveyed, 8) excluding individuals with fewer than non-missing frailty index observations, and 9) using, for every individual, the largest set of non-missing items available throughout their inclusion in ELSA to calculate their individual frailty index. This removes any potential bias from the inclusion of new items on the frailty index over time, and thus only looks at changes in the frailty index coming from changes in the items available in all survey waves for a given individual.

The results are presented in Table B.4. The coefficients for age and year of birth are very similar across these robustness checks. This shows that neither longitudinal attrition and death nor missing items in the frailty items, nor age inclusion criteria substantively change our main result.

Table B.4: Robustness check: Varying sample inclusion criteria

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Age	0.038***	0.038***	0.036***	0.032***	0.041***	0.045***	0.039***	0.038***	0.038***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Year of birth	−0.014***	−0.019***	−0.018***	−0.019***	−0.021***	−0.018***	−0.012***	−0.030***	−0.019***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.007)	(0.002)
Mean Age	0.012***	0.008***	0.007***	0.005**	0.006**	0.006***	0.014***	−0.009	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.008)	(0.002)
Observations	63,971	65,693	63,167	58,317	59,802	59,079	52,149	48,101	65,122
Sample	Survived	35 items	40 items	Below 80	Above 55	>99% items	Complete indiv.	>=5 records	Common items
What is the new 70?	74.4	75.3	75.4	75.9	75.4	74.5	73.8	77.0	75.4

Notes: All columns show results from Mundlak regressions. Robust standard errors clustered at year of birth level in parentheses. Age is demeaned at individual level. Significance levels: *10%, **5%, ***1%. Included controls: Sex, education, log wealth (+ mean), NUTS-1 region dummies (+ mean).
The last row displays the predicted age in 2018 in which frailty equals the frailty level of a 70-year-old in 2002.

B.4 Fully flexible estimation: year of birth fixed effects

Figure B.1 implements a Mundlak specification where one dummy variable for each year of birth cohort is added instead of a continuous year-of-birth variable. The estimated coefficients are fitted with a quadratic fit, weighted by the number of observations in each year-of-birth cohort. The figure confirms the quadratic trend observed in the data: the level of frailty is declining with more recent cohorts, and this is slowing down at a moderate rate.

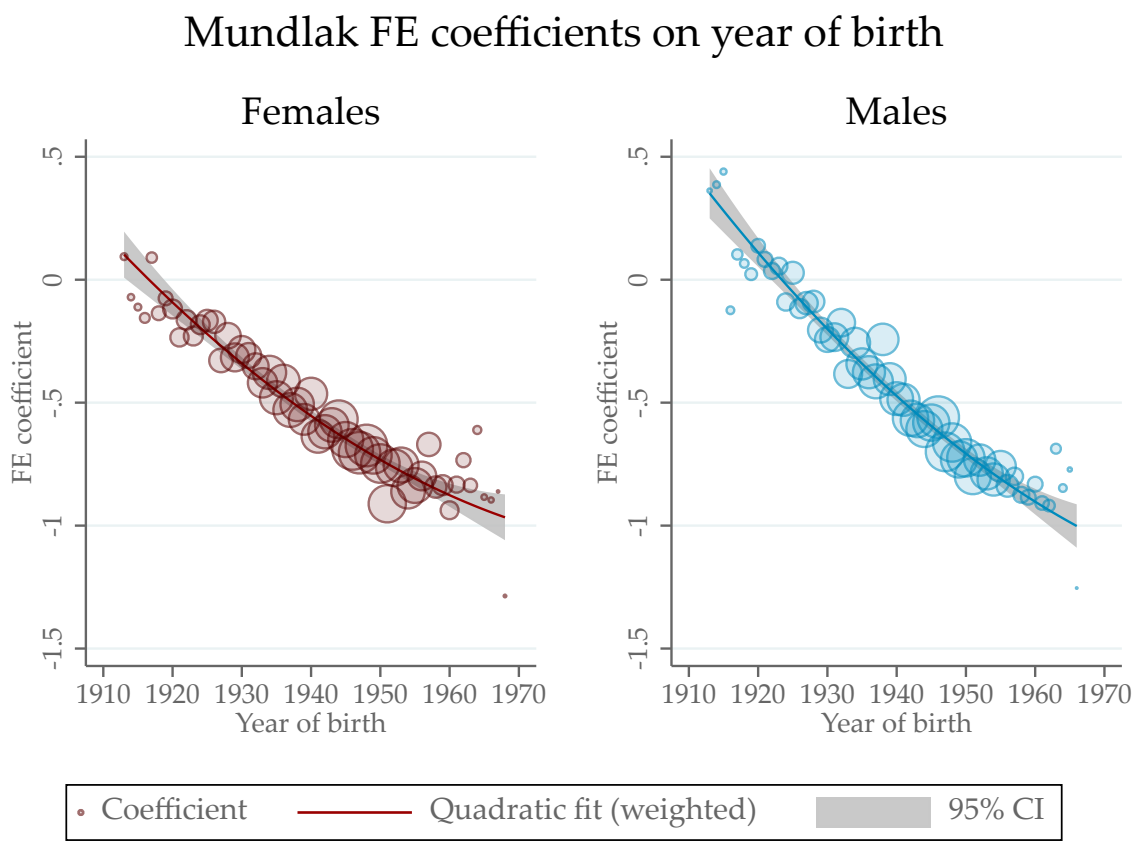

Figure B.1: Coefficients on year of birth fixed effects from Mundlak regression with full set of controls, with quadratic fit (weighted by number of observations). The area of the markers is proportional to the number of observations

B.5 Additional results

Table B.5 estimates specifications on the five sub-components of the frailty index. Compared to the main text (Table 5), the log of the sub-components +1 is used as dependent variable. In this way, observations with zero frailty in any of the sub-components are not dropped, leading to a larger sample size. Direction and statistical significance of the coefficients are unchanged.

Table B.6 estimates Mundlak specifications by sex, education, and baseline wealth tertile. In addition to the main text (Table 6), quadratic terms are included. The results poiint to a general slowdown in the rate of frailty improvement, but high SES groups (high education, high initial wealth) see a larger slowdown in the rate of frailty improvement than groups with low education or initial wealth.

Table B.5: Cohort trends by sub-components of frailty index – ln(x+1)

	(1) Full Index	(2) Mobility (10)	(3) ADL/IADL (13)	(4) Depression (8)	(5) Conditions (12)	(6) Cognitive (6)
Age	0.004***	0.005***	0.003***	0.001	0.006***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Year of birth	−0.002***	−0.003***	−0.002***	−0.001***	0.000	−0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mean Age	0.001***	0.002***	0.000	−0.001	0.002***	0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	71,289	71,285	71,286	68,551	71,288	68,794
Mean of DV	0.132	0.192	0.061	0.179	0.128	0.101

Notes: All columns show Mundlak regressions. Robust standard errors clustered at year of birth level in parentheses. Significance levels: *10%, **5%, ***1%. Age is demeaned at individual level. Controls: Sex, NUTS1 region dummies (+ mean), education, log wealth (+ mean). Number of items in each sub-component listed in parentheses in the column headers.

Table B.6: Longevity trends: Heterogeneous effects with quadratic year of birth

	Sex		Education			Wealth		
	(1) Males	(2) Females	(3) Low	(4) Middle	(5) High	(6) Low	(7) Middle	(8) High
Age	0.041***	0.036***	0.042***	0.038***	0.033***	0.036***	0.040***	0.038***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Year of birth	−0.625**	−0.609***	−0.482*	−1.071***	−1.600***	−0.111	−0.509*	−1.176***
	(0.272)	(0.183)	(0.273)	(0.225)	(0.356)	(0.294)	(0.265)	(0.287)
(Year of Birth^2)/1000	0.155**	0.153***	0.120*	0.270***	0.408***	0.026	0.126*	0.296***
	(0.070)	(0.047)	(0.070)	(0.058)	(0.091)	(0.076)	(0.068)	(0.074)
Mean Age	0.003	0.011***	0.009**	0.003	0.017***	0.010*	0.008**	0.005
	(0.003)	(0.002)	(0.004)	(0.003)	(0.005)	(0.005)	(0.003)	(0.004)
Observations	30,046	35,648	23,530	31,953	10,211	20,934	21,754	23,006
Method	Mundlak	Mundlak	Mundlak	Mundlak	Mundlak	Mundlak	Mundlak	Mundlak
What is the new 70 in 2018?	75.8	75.1	74.5	76.6	75.9	73.6	75.7	76.8
What is the new 70 in 2034?	75.0	74.0	73.6	75.1	72.0	73.4	74.9	75.1
p-value Interaction term	0.082		<0.001			<0.001		

Notes: Robust standard errors clustered at year of birth level in parentheses. Significance levels: *10%, **5%, ***1%.
Education levels are: Low: below upper secondary, Middle: upper secondary/vocational training, High: tertiary education. Wealth is split in tertiles relative to five-year age group in the first occurrence in the survey. Age is demeaned at individual level. Controls: Sex, NUTS1 region dummies (+ mean), education, log wealth (+ mean). In the specifications for sex and education, the respective variable is not included as a control.
The p-value is from a Wald-test on the interaction term of year of birth squared and the heterogeneity variable.