# Early retirement and cognitive decline.
# A longitudinal analysis using SHARE data

Martina Celidoni
Department of Economics and Management, University of Padua
Via del Santo 33, 35122 Padua, Italy
martina.celidoni@unipd.it

Chiara Dal Bianco
Department of Economics, Ca' Foscari University of Venice
Cannaregio 873, 30121 Venice, Italy
chiara.dalbianco@unive.it

Guglielmo Weber*
Department of Economics and Management, University of Padua
Via del Santo 33, 35122 Padua, Italy
guglielmo.weber@unipd.it

July 9th, 2015

**Abstract**

We show that a new measure of cognitive decline, that can be computed in longitudinal surveys where respondents perform the same recall memory tests over the years, is highly predictive of the onset of dementia. Using SHARE data, we investigate the association between cognitive decline and years in retirement controlling for age, physical health, early life conditions and socio-economic status. We find a positive association and an even stronger causal effect. However, the causal effect can be established for individuals who retire at the statutory eligibility age, not for those who retire on an early retirement scheme. The evidence we produce confirms the 'mental retirement' hypothesis and suggests its relevance for the onset of dementia, but suggests that a distinction must be made for those who retire as soon as possible (early retirement) and as late as possible (statutory retirement).

Keywords: Ageing, cognition, retirement, instrumental variable estimation, SHARE

JEL: I12, I1, J26

* Corresponding author: Guglielmo Weber, Department of Economics and Management, University of Padua, via del Santo 33 - 35122 Padua, Italy – tel.: +39 049 8274271, email: guglielmo.weber@unipd.it.
This paper was presented at the EUGMS 2013 conference, the 4th SHARE user conference (Liège, November 2013), the Netspar International Workshop (Venice, June 2014), VIU Summer Institute (Venice, June 2015). We are grateful for comments and suggestions made by participants. This research was financed by the UE, SHARE-M4 project, and by the University of Padua, POPA_EHR project.

## 1 Introduction

Population ageing in Europe and other developed countries challenges the sustainability of the health care and long term care systems. One of the key reasons individuals require long term care in old age is cognitive decline, leading to dementia when it interferes with independent functioning (American Psychiatric Association, 2000). According to a recent study about the United States (Hurd et al., 2013), dementia affects a large and growing number of older adults and represents a substantial financial burden for the society with estimated costs similar to those related to heart disease and cancer.

Cognitive abilities later in life have been widely investigated in epidemiology and gerontology (Dixon et al., 2004; Schaie, 1994): the related literature documents a decline of cognition at older ages with strong effects on fluid abilities such as memory when recalling specific past events (Peterson et al., 2002; Bäckman et al., 2005). [1] Individual heterogeneity in cognition levels and changes with age are likely to be associated with individuals' engagement in mentally stimulating activities (Salthouse, 2006; Maguire et al. 2000).

Particular attention has been devoted (since the seminal paper by Adams et al, 2007) to the effects of retirement on cognition since this transition marks a change in individuals' life-style. According to Rohwedder and Willis (2010), retirees are engaged in less mental exercise than workers: the latter are exposed to environments that are more cognitively challenging and stimulating compared to the non-work condition (the so-called "unengaged lifestyle hypothesis"). If they are right, the spate of recent pension reforms increasing retirement age would also reduce long-term care expenditure (Dave et al., 2008; Bonsang et al., 2012).

When assessing the role of retirement on cognition, endogeneity issues have to be taken into account. There could be a reverse causal link - individuals who experienced a bad health shock retire as soon as possible (see Insler, 2014). Also, there is likely to be a selection problem (Coe and Zamarro, 2011): people self-select into retirement based on their gains from retirement - those with the most physically demanding jobs, or who enjoy their jobs the least, retire earlier to relieve themselves of the daily strain.

In this paper we use individual panel data from a host of European countries to investigate the relation between cognitive decline and years from retirement, controlling for age, education and other confounding factors. We estimate the causal effect of retirement on cognitive decline by using eligibility ages for early retirement and statutory (old-age) pension in several European countries over time as instruments for retirement and retirement-related variables (such as years from retirement). The variability across individuals in public pension eligibility reflects gender, time of retirement and country of residence, and this ensures that the instruments we construct are informative. The existence of two different types of pension eligibility criteria is a key feature of our data that allows us to investigate whether retirement has heterogeneous effects on cognition.

The key novel features of our analysis are:

- we exploit the longitudinal dimension of SHARE data after showing that non-random attrition does not impair the analysis;[2]

[1] Episodic memory is traditionally considered an information processing system that receives and stores information about temporally dated episodes or events, retains various aspects of this information and, upon instructions, transmits specific retained information (Tulving, 1972).

[2] Almost all papers on European data use cross sections rather than the available longitudinal information. To our knowledge, the only exceptions are Bianchini and Borella (2014) and Mazzonna and Peracchi (2014), that present fixed effect estimates on SHARE data. However, fixed effect estimation controlling for age relies on transitions into retirement for identification, and is therefore unsuitable to estimate the cumulative effect of retirement in a short panel. For this reason, we prefer to control for the lagged endogenous variable, instead (Imbens and Wooldridge, 2009).

- we adopt a sharp measure of cognitive decline, based on a 20% drop in the number of words recalled between waves and show that it predicts well the onset of dementia in a commonly used US data set, the Aging, Demographics and Memory Study (ADAMS);
- we investigate heterogeneity in retirement effects related to the existence of early and statutory retirement ages.

Exploiting the longitudinal dimension of the data is important because it allows us to construct individual-level indicators of cognitive decline, such as the measure that is based on a 20% drop in words recalled. We argue that this is more appropriate in presence of re-testing effects, typically found in longitudinal data (Ferrer et al., 2004): respondents tend to improve their performance in memory tests across waves, particularly the second time they are interviewed. Dal Bianco et al. (2013) present descriptive evidence on this issue using SHARE data. They argue that high decreases in words recalled are more informative about actual declines, as opposed to straight changes in the score. In this paper, we show that the 20% decline measure is highly predictive of the onset of dementia: using a sample of 432 individuals aged 70 or more who took part in ADAMS and were later medically assessed for dementia, we find that our measure correctly classifies 70% of individuals according to their later dementia status.

In our baseline results retirement status per-se has no effect on cognitive decline, but years in retirement has a significant, positive effect, after controlling for age and education, in line with Bonsang et al. (2012). Our evidence therefore supports the cumulative negative effect hypothesis. Retiring at younger ages increases the probability of experiencing cognitive decline at older ages - and this reinforces the view that retirement can be detrimental to the well-being of individuals.

However, we find that there are significant differences in estimated parameters when we focus on two distinct groups of retirees: those who retire as soon as possible (that is, at the time when they become eligible for an early retirement pension) and those who instead retire as late as possible (at the time when they qualify for an old-age pension - when in most cases retirement from the job is mandatory or anyhow employment protection ceases to operate). For the former, we find that retirement has no negative effects on cognition – for the latter, that retirement has strong negative cumulative effects. To the extent that early retirement is a free choice, but statutory retirement is not, this finding is consistent with consumer rationality.

The paper is organized as follows. Section 2 presents the data; the validation analysis is described in section 3 and the empirical strategy in section 4. Section 5 comments the results, and presents some robustness checks and section 6 concludes.

## 2 The Data

In our empirical analysis we use data drawn from the Survey of Health, Ageing and Retirement in Europe (SHARE)[3] which collects information on health, socio-economic status and social and family networks. The SHARE target population are individuals aged fifty or over who speak the official language(s) of their

[3] This paper uses data from SHARE wave 4 release 1.1.1, as of March 28th 2013 or SHARE wave 1 and 2 release 2.5.0, as of May 24th 2011 or SHARELIFE release 1, as of November 24th 2010. The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N211909, SHARE-LEAP, N227822 and SHARE M4, N 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions.

country, plus their partner regardless of age. The baseline study, which took place in 2004, involved a balanced representation of the various regions in Europe, ranging from Scandinavia (Denmark and Sweden) through Central Europe (Austria, France, Germany, Switzerland, Belgium, and the Netherlands) to the Mediterranean (Spain, Italy and Greece). To this first set of 11 countries several others have been added in the following waves.

### 2.1 Sample selection

In our study we restrict the sample of analysis to respondents, aged 50 or over, taking part in the first wave, or those interviewed for the first time in the second wave (refreshment sample).

Among these, we keep only individuals re-interviewed both in the third wave, called SHARELIFE given its retrospective nature, and in the fourth one. Since we are interested in studying the effect of retirement on cognition, we select respondents who were working or retired from work in the baseline (i.e. the first or the second wave depending on when respondents entered the sample).[4] For the most part we pool males and females. However, we also investigate gender specific effects to understand whether the more interrupted careers of females play a role.

While Rohwedder and Willis (2010), Mazzonna and Peracchi (2012) and Bonsang et al. (2012) define an individual as retired if he or she reports not to be working, for the purpose of our analysis we define a respondent as retired if he or she declares to be retired from work and has work experience higher than or equal to 15 years. By doing this, we avoid to include unemployed or disabled individuals among the retirees so that we can strictly focus on the consequences of retirement from work on cognitive abilities, reducing the influence of long inactivity periods on cognition for other reasons.

Coe and Zamarro (2011) observe that there are individuals still working but who declare themselves as retired simply because they left their *career job*. In our sample 94% of individuals that in baseline declare to be retired did not do any work for pay in the previous four weeks. Even if we reclassify that 6% of individuals as not retired, results do not change.

Further we exclude proxy interviews because in those cases individuals do not perform cognitive tests and we do not consider interviewees with missing values in tests' scores in at least one of the two measurement occasions (i.e. baseline and wave4). The final sample is a balanced panel with two time periods for 8221 individuals (6125 observed between wave 1 and 4, 2096 refresher of wave 2).[5]

### 2.2 Covariates

[4] The number of released interviews with coverscreen information in SHARELIFE is 26769. We drop 501 individuals who participated only to SHARELIFE (typically new spouses), 1596 individuals born after 1956, 3988 individuals who participated to SHARELIFE but to only one regular wave - for whom we cannot compute our cognitive decline measure. We do not include in our sample 2625 individuals who are neither retired nor employed/self-employed; 604 retired individuals for whom we do not have information about the year of retirement. We then drop individuals who participated only to w1, w2 and SHARELIFE (4605, including the whole Greek sample) and individuals belonging to countries which do not participate continuously in SHARE (Czech Republic and Poland - 2109 individuals). We keep individuals with age of retirement between 40 and 70 and job experience higher than 15 (151 and 786 observations dropped). For 288 individuals we were not able to compute our cognitive decline measure, since they did not do the word recalling test in both waves. We drop 921 observations due to missing information about covariates in the richest specification (e.g. contact with children, depression, early life conditions). Finally, we drop individuals belonging to cohorts for whom early and statutory retirement eligibility ages coincide (537) because we will investigate heterogeneous effects related to the two types of pension.

[5] Our final sample includes the following countries: Austria, Germany, Sweden, Netherlands, Spain, Italy, France, Denmark, Switzerland and Belgium.

In addition to retirement, another key variable in our analysis is years spent in retirement. This variable is computed on the basis of the year when the SHARELIFE respondent reports to have retired from work. When this information is missing or the respondent retired between waves 3 and 4, we use a comparable question from wave 4 questionnaire. In Figures B1-4 in Appendix B we show the proportions of individuals who self-report being retired in each country by gender as a function of the years from/to the eligibility age separately for early and statutory retirement (in Appendix B we describe the eligibility criteria used for each country). It is interesting to notice that there are sizeable jumps in the proportion of retired at both eligibility rates in most countries. In Table 1 we report descriptive statistics by country for our set of retirement variables (retired, transition into retirement and years in retirement) and eligibility age for early and normal retirement.

In our baseline specification we include education, expressed according to the International Standard Classification of Education (ISCED), the logarithm of age, gender, quartiles of household equivalent income, self-reported health and participation in activities. We include in the model also a dummy, *low cognition*, that takes value one if the baseline memory score is lower than the median value by wave and country, as a 20% decrease is less likely to be observed if the initial value is already low. We capture the role of activities, by defining a dummy for whether the individual performs a vigorous physical activity at least weekly (*physicalact*), and another dummy (*dailyact*) that takes value 1 if the individual performed daily, during the last month, other activities such as voluntary work, training course, participation in religious or political organizations. For time-varying covariates, health and engagement in activities, we include variables capturing changes between waves. *Drophealth* identifies individuals that declare to be in poor health in wave 4, given that in baseline they answered their health was at least fair; *increasehealth* denotes individuals that reported being in poor health in baseline but report an improved health status in wave 4. For changes in activities: a *drop* means that the individual used to perform the activity and stops it in wave 4, while *increase* means that the respondent starts performing the activity in question in wave 4. We control also for income in our specification, by adding quartiles computed on household equivalent income by country and wave, *eqincomeQ1* is the lowest quartile and the reference category, the equivalence scale used is the square root of the household size. Our income variable is meant to capture differences in living standards, rather than day-to-day income variations. Finally, we have also country dummies, a control for the wave 2 refreshment sample (to take into account the shorter time distance from wave 4) and *less repetition* which identifies individuals who performed the test only twice (because they enter the sample in wave 2 or because they were interviewed in wave 1 and wave 4 but not in wave 3).

In some of our analyses, we enrich our specification with additional controls related to early-life conditions, contextual effects, job characteristics, other health measures, and contacts with children (see Appendix A for a detailed description and Table A.1 for summary statistics).

### 3 Outcome variable: definition and validation

Data about cognitive abilities are collected in each regular wave of SHARE [6]: a series of brief tests are included in the CAPI questionnaire. Among them, only verbal fluency and verbal learning tests are performed in all three waves.[7] It is worth noting that, since cognitive decline is a multidimensional phenomenon, each test usually measures a different aspect of cognition.

In agreement with the economics literature, we focus on memory scores (based on a modified version of the Rey's Auditory Verbal Learning Test-RAVLT). In this ten-word-list learning test the respondent is asked to

[6] We refer to waves 1, 2 and 4 as regular.

[7] The verbal learning test has the same technical features in all three waves, the only exception is that words used in the fourth wave of SHARE are different from those used in the previous waves, for details see Malter and Boersch-Supan (2013).

learn a list of ten common words and recall them immediately (immediate recall or first trial) as well as after an interference period (delayed recall or second trial), roughly 5 minutes later. As in Rohwedder and Willis (2010) and Bonsang et al. (2012), we measure cognitive abilities as the sum of words remembered in the immediate and delayed recalls with a score ranging from 0 to 20. This test is preferred to verbal fluency (whereby respondents list as many animal names as possible within a short period of time) because memory is particularly affected by ageing and, in addition, it does not suffer from floor and ceiling effects (Bonsang et al., 2012). On the basis of the memory score, we shall compute our main outcome of interest, "high decrease".

Figure 1 here

Figure 2 here

The standard negative association between cognitive abilities and age, that most papers on cognition refer to, is confirmed in Figure 1: looking at the cross-sectional variability in memory scores in SHARE it is possible to notice how the total number of words recalled decreases almost linearly with age. This is the kind of relation that almost all the previous studies about the effect of retirement on cognition have exploited, further highlighting a drop around pension eligibility ages. To better understand cognitive decline, however, the longitudinal information should be exploited as following the same individuals over time is a way to fully control for cohort and other individual-specific time-invariant effects.

The left panel of Figure 2 provides a graphical representation of the average memory score, computed as the sum of words recalled in the first and second trial by cohort and wave. Following cohorts across waves we notice that the average memory score falls with time/age only for individuals born before 1935. The right hand panel of Figure 2 shows that the same pattern occurs when we control for country, cohort, time, gender, education and retirement status: the test performances of oldest individuals fall over time as one might expect, whereas the number of words recalled improves over time for the youngest individuals (born in the 1950s). This latter finding is probably due to the fact that we are looking at individuals who were interviewed more than once.

As emphasized in the literature, in fact, longitudinal analysis suffers from learning or retesting effects (Ferrer et al., 2004). The phenomenon of practice, or re-test, is well known in the area of cognitive abilities (e.g. McArdle and Woodcock, 1997; Schaie, 1996): measures of cognitive decline over time are plagued by the fact that individuals might learn from tests performed in the previous waves and this implies an upward bias in cognitive ability measurement. There is no guarantee that this bias is constant over age or education.

The literature has suggested some strategies to tackle learning effects issues, an example could be to include in the same analytical model separate terms for age and measurement occasion. We follow this strategy, but in addition, we focus on a particular, dichotomous outcome variable that takes value 1 if the number of words recalled has fallen by more than 20% across waves (as put forward in Dal Bianco et al, 2013). This variable has the advantage of being a conservative measure of decline in the presence of re-testing effects.

Table 2 here

In the literature there is no standard threshold to discriminate different levels of cognitive declines. However, as Table 2 reveals, by focusing on drops higher than 20% we select those changes that are in the bottom quarter of the distribution of memory score variations between waves. This is a fairly stable result: in Table 2 we show the distribution of decreases in memory scores respectively between wave 1 (2004) and wave 4 (2011) and, for the refreshment sample of wave 2, between wave 2 (2006) and wave 4 (2011). We also show the same statistics for a restricted sample where very old individuals are excluded. In all cases the 25$^{th}$

percentile corresponds to falls between 20% and 23%.[8] In Figure 3 we show the percentage of high decrease by age in wave 4.

Figure 3 here

Our conservative strategy of analysing high decreases in cognitive abilities should identify the most vulnerable individuals whose drop in cognition wipes out any learning effects. Focusing on high decreases rather than any decrease should also help in reducing measurement issues, since mild decreases might not reveal true cognitive deterioration but only measurement errors.

To understand if our measure of cognitive decline can be considered symptomatic of a pathological impairment related to dementia, we use data drawn from HRS where, for a sub-sample of individuals aged 70+, we have both a memory test, similar to that proposed in SHARE, with immediate and delayed recall[9], and a clinical assessment for dementia. To our knowledge this is the only source of information that allows the comparison that we are interested in. The HRS sub-sample of respondents is the basis for ADAMS, whose purpose is to gather additional information on cognitive status and assign a diagnosis for dementia to a group of respondents who are particularly at risk of developing it. We are especially interested in cognitive declines between waves that we will compare with the clinical assessment of dementia provided by a nurse and a neuropsychology technician specifically trained in data collection for dementia evaluation. Since in ADAMS we can observe, for the same individuals, many transitions between waves, to avoid individuals' replications, we selected the longest transition at our disposal. Our final sample size consists of 432 individuals, who potentially transit into dementia, and for whom we can compute our high decline in memory score indicator.

Table 3 here

In order to understand whether a high decline in cognition can be a symptom of dementia, we follow these steps. First we estimate a logit model where the outcome is dementia (whether the individual has been diagnosed with 'mild/moderate/severe dementia') controlling for high decrease, gender, interval length and test features (if there was any distracting factor during the test). Predictions are then compared with the diagnosed cognitive status.[10] Table 3 summarises our results: if we are particularly interested in correctly identifying individuals with any sign of dementia, we see that predictions based on high decline in cognition is a good measure as in the 70% of cases it corresponds to a clinical assessment of the pathology.

To validate our measure, we provide also results coming from a widely used approach in the field of disease diagnosis, the Receiver Operating Characteristic (ROC) curve. This curve measures the extent to which a given signal - predictions based also on high decreases in cognition in our framework - can detect an underlying condition, i.e. dementia. By varying the probability threshold that classifies individuals between predicted demented or not, the curve provides a graphical representation of the signal identification power. In order to draw the curve, we need the following information for every possible probability threshold of our predictions: those predicted and assessed demented (the so called true positives - TP), those predicted but not assessed demented (false positives - FP), those not predicted but assessed demented (false negatives - FN) and finally those neither predicted nor assessed demented (true negatives - TN). The ROC curve exploits that

[8] In Table 2, we notice that for some percentiles, individuals observed between wave 2 and 4 exhibit a larger drop than those observed in wave 1 and 4. This might be explained by selection (attrition) issues, that we investigate in the robustness section, or by re-testing effects.

[9] Differently from SHARE, the verbal learning and memory test in ADAMS consists of three immediate and one delayed recalls. In order to maintain the comparability with SHARE, which proposes to respondents only one immediate and one delayed recall, we selected the number of words listed in the first immediate and the delayed recall.

[10] In order to compare predicted and actual classification we used a cutoff of 0.55. When we control also for a low level cognition in baseline, the percentage of correct predictions rises to 74%.

classification to plot, on the vertical axis, the sensitivity or TP rate, TP/(TP+FN), against 1-the specificity or TN rate, 1-TN/(FP+TN), on the horizontal axis, for all possible values of the probability threshold. The more correlated are predicted and assessed dementia, the higher will be sensitivity and specificity, the nearer will be the curve to the upper-left corner in Figure 4. For a more intuitive summary of the extent to which predictions are correlated with assessed dementia, we compute also the area under the curve (AUC), which is estimated to be 0.78 (95% Confidence interval: 0.73-0.82), a value that is considered good in the literature.

Figure 4 here

After presenting all the key variables in this analysis and the related issues, we now turn to the description of the empirical strategy

## 4 Empirical strategy

The aim of the empirical analysis is to estimate the role of years spent in retirement on cognitive decline accounting for re-testing effects, i.e. looking at the within-person drop in cognition over time as a function of the years spent in retirement, controlling for age and other covariates and taking into account the endogenous nature of retirement. We base our measure of cognitive decline on the percentage change in words recalled between waves. Formally, if $score_i$ denotes the number of words recalled in both immediate and delayed test, we define $y^*_i$, the percentage change in memory score, as follows:

$$y^*_{i,t} = (\mathrm{score}_{i,t} - \mathrm{score}_{i,t-1})/\mathrm{score}_{i,t-1} \qquad (1)$$

We further define our sharper measure of high cognitive decline, $y_i$, as a dummy variable that takes value 1 if $y^*_{i,t}$ is lower than -0.2 and zero otherwise.[11] We adopt the following linear specification for $y_{i,t}$:

$$y_{i,t} = \beta_1\, retired_{i,t-1} + \beta_2 \log(1+yearinR_{i,t-1}) + \beta_3\, fromWtoR_{i,t} + X^T_i\, \beta_4 + \epsilon_i \qquad (2)$$

where we assume that the probability of observing a decline in cognitive abilities, measured on the basis of memory scores, depends on retirement status in t − 1, $retired_{i,t-1}$, the logarithm of years spent in retirement, $\log(1+yearsinR_{i,t-1})$, and whether we observe the transition from working to retirement between waves, $fromWtoR_{i,t}$. We include in the model also a vector of covariates, $X^T_i$ as described above (that always includes age, gender and education[12]).

The first retirement-related variable captures the status effect of retirement on cognitive decline. For instance, individuals who are retired from work may be more often depressed, and this may lead to faster loss of cognition over time, or instead feel relieved. The second retirement-related variable captures the progressive loss of fluid memory induced by a less engaged life style, for a given age. One might expect this loss to be zero at the beginning of the retirement period, and to build up over the years. The third and last retirement-related variable instead captures the immediate effects of retiring from work – that may be beneficial if work had become a psychological or physical burden ("honeymoon effect").

[11] We stress that we exclude proxy interviews because individuals in those cases did not perform cognitive tests; we also delete interviews with missing values in tests scores in either baseline or wave 4. This selection could affect our estimates, since there is a high probability of not observing a cognitive drop for individuals cognitively impaired that did a proxy interview or did not participate in wave 4 due to poor health conditions. Our results do not change if we include in the cognitive decline group individuals who performed the memory test in baseline but did not in wave 4.

[12] Bingley and Martinello, 2013, show that omitting education can seriously bias the estimated effect of retirement on cognition, because low-education individuals, who perform poorly in cognitive tests, tend to retire earlier.

The advantage of the (log) linear specification is that we can easily account for the potential endogeneity of the retirement decision. For this, we need instruments that are both relevant, i.e. directly related to retirement decisions, and exogenous - that have an effect on cognition only through their impact on retirement. As by now standard in the literature (following Battistin et al. 2009), retirement decisions are instrumented by legislated ages of eligibility for early retirement and old-age pension. Differently from other studies, that adopted the same instrumental variables strategy (Rohwedder and Willis, 2010), we exploit not only the cross-country variability in eligibility ages, but also variations over time as in Angelini et al. (2009). As Mazzonna and Peracchi (2012) observe, in fact, SHARE data offer a substantial within-country variability in eligibility rules arising from the pension reforms of the 1990s, which contributes substantially to the European heterogeneity of pension entitlements.

We use the following instruments for $retired_{i,t-1}$, log $(1+yearsinR_{i,t-1})$ and $fromWtoR_{i,t}$: two dummy variables that take value 1 if the individual is eligible for early and normal retirement, two variables indicating the logarithm of years since eligibility for the two types of retirement and two dummies that equal 1 if we observe the transitions from not being eligible to being eligible between waves. Therefore in our two-stage least squares (TSLS) specification, we have three endogenous variables and six instruments. As we show in the next section, the instruments are relevant but the over-identification tests are rejected (and this may suggest heterogeneity between early and statutory retirement).

Differently from Coe and Zamarro (2011) and Rohwedder and Willis (2010) that consider only the binary treatment of retirement, we also take into account the cumulative role of years spent in retirement. As argued in Bonsang et al. (2012), Coe et al. (2012) and Mazzonna and Peracchi (2012), in fact, the effect of retirement may not be instantaneous. According to Atchley (1976, 1982), individuals might experience, right after retirement, a so-called *honeymoon phase* in which they can engage in different activities that were set aside because of work-related constraints. This engagement in desired activities may attenuate the negative effects of retirement on cognition. We might also expect that changes in activities would translate only progressively into changes in cognitive abilities. If this is true, considering in the empirical model only the retirement status could provide just a partial description of changes in cognition and be uninformative for policy purposes. There could be, in fact, a cumulative effect of years spent in retirement: the longer the period of time since the individual retired the more likely that he or she experiences a high decrease in cognition. As Bonsang et al. (2012) argue, if the impact is cumulative, there could be gains in terms of lower long-term care expenditures coming from policies that increased retirement eligibility ages, since they might delay the appearance of cognitive impairments at older ages.

## 5 Results and robustness analysis

In this section we present our estimation results when the dependent variable is the cognitive decline measure described above.[13] We also show the results of robustness analysis in a number of directions.

Table 4 here

### 5.1 Estimation results – main specification

In Table 4, we reports OLS and TSLS estimates for two different specifications, one with a limited set of controls (first and second column) and the other one with a larger set of controls (third and fourth column). Estimated standard errors are robust to clustering at the country, gender and cohort level. Focusing on OLS estimates (first column), we can see that the dummy indicator "being retired" is significant with negative sign, the indicator for transiting from "work to retirement" has a negative but insignificant effect while the variable (logarithm of) "years spent in retirement" is significant with a positive sign. The positive sign on the key variable of interest (years in retirement) is consistent with the notion that a long period in retirement is associated to an increased probability of experiencing a large decline. Given that we control for age (that has a positive, significant effect), we can infer that retiring early is statistically associated with cognitive decline (the longer you have been retired for a given age, the earlier you must have retired). To the extent that early retirement can be induced by negative health shocks, or is anyhow taken by individuals who no longer enjoy working, this association cannot be given a causal interpretation. Other controls include low cognition in baseline (that has a strong negative effect), gender (women are less likely to suffer from cognitive decline[14]), education, income quartiles in base year, poor health, physical activity, daily activities (and their changes), as well as a set of country dummies and a dummy that controls for the shorter period in the panel for those who were first interviewed in wave 2 (the wave 2 dummy). High education and income are seen to exert the expected protective role against cognitive decline, and so do activities in the base year. Deteriorations in health are positively and significantly associated to cognitive decline – changes in activities (both positive and negative) also have strong effects of the expected sign.

The role of variables associated with activities is particularly interesting as activities have been proposed as determinants of cognitive aging patterns in the psychological literature (Salthouse, 2006). We find that physical activity has a protective role for cognitive abilities (according the ancient Roman say "*mens sana in corpore sano*", a healthy mind in a health body). Also an increase or a drop in such activities has a significant effect on the probability of a high decrease. Not only physical activity but also practicing daily any activity in baseline seems to be beneficial, that is they are associated with a reduction in the probability of experiencing a high decline in cognition. A change in the daily activity behaviour, especially if the individual between waves starts practicing daily any activity, is associated with a reduction in the probability of observing a high decline in memory score.[15]

We now compare these results with TSLS estimates (second column) where we account for endogeneity of retirement decision as explained above. We can observe that the transition from work into retirement again has insignificant effects, whereas being retired and years spent in retirement remain significant with a

[13] We are able to replicate previous findings with different dependent variables on SHARE and HRS data; results are available upon request.

[14] It should be remembered that women in our analysis are those who worked or were retired from work in baseline, and this might be a positively selected sample (home makers are excluded). However, it is also possible that women are less likely to experience memory losses because of their more engaged life style (Steeves et al., 2014).

[15] Boersch-Supan and Schuth (2013) highlight the role played by social networks in forging the link between early retirement and cognition, and between early retirement and depression. Unfortunately questions on the social network were not asked in SHARE until wave 4, but the information we have on daily activities may capture the size and quality of the social network.

negative and positive sign respectively, supporting the so-called *honeymoon phase* and the negative cumulative hypothesis. Retirement duration therefore, according to our estimates, plays a role in the evolution of cognitive decline at older ages, over and above the pure age effect that is well documented in the medical literature. In this specification, we control for the same set of other variables as in column (1).

We notice that OLS suffers from attenuation bias compared to TSLS. A possible reason is that retirement is an error-ridden measurement of the disengagement associated to the end of a working career: some people mentally retire in their last years on the job, others remain active even after official retirement.

At the bottom of Table 4 we report the Sanderson-Windmeijer (2015) version of Angrist-Pischke weak instruments F-tests, and the Sargan-Hansen test of the over-identifying restrictions (Table A.3 in Appendix A reports the first stage regressions). On the basis of the test results, we conclude that the instruments used are not weak, but the over-identification restrictions are rejected at the conventional 5% significance level. We shall discuss later a possible reason for this rejection.

It is interesting also to control for other variables that can affect the probability of experiencing a high decline in cognition, especially those related to how the test has been carried out, that are normally ignored in the literature (with the notable exception of Mazzonna and Peracchi, 2012), or variables associated with depression and mental health, or with childhood circumstances.

Columns (3) – OLS – and (4) – TSLS – in Table 4 show that changes in test characteristics between waves have a significant effect on the probability to observe a high decrease in words recalled. Especially, the fact that the respondent in baseline carried out the test while alone in the room is associated with a lower probability of a high drop in cognition. Distracting factors in baseline and changes between waves are associated with the probability of observing a drop in memory scores as expected. The effects of depression are less strong: only the transitions from non-depression in baseline to depression in wave 4 has a positive significant effect on cognitive decline (when we estimate by OLS), but also in this case the direction of causality is less than obvious. We find that early life conditions affect the probability of experiencing cognitive decline: especially doing well in math at age 10 lowers the probability, whereas living in rural area during childhood increases it. This result confirms the importance of early life conditions on late life outcomes, as stressed in recent research by Heckman and collaborators (see, for instance, Heckman and Cunha, 2007). We also find that daily activities exert a protective role on cognition – estimated coefficients on these variables are extremely close to those obtained by OLS. These effects of course could be due to reverse causality – for this reason we report in Appendix A (Table A.2) parameter estimates for specifications that do not include any activities-related variable.

Point parameter estimates of retirement-related variables in column (2) and (4) are almost identical (and so are their standard errors), and this confirms that the causal effect of retirement on cognition is robust to the inclusion of a richer set of controls.

Table 5

Following Bonsang at al. (2012), our baseline specification includes the logarithm of years spent in retirement and the logarithm of age to account for possible non-linear effects. In Table 5 we address the issue of functional form specification, focusing on the years in retirement and age effects. Column (1) in Table 5 reproduces column (2) of Table 4 – columns (2), (3) and (4) show what happens when an age polynomial of first or second or third order is introduced together with years spent in retirement. The age terms are jointly significant in columns (3) and (4), and in these columns the point estimates on *yearsinR* are significant and of comparable size. No significant effects are estimated for ‘being retired’ and the transition from work to retirement. We conclude that our key results are not driven by the choice of a logarithmic functional form.

### 5.2 Estimation results: early and statutory retirement heterogeneity

The evidence we provided so far confirms that work has in the long run a protective role on cognition – the later individuals retire the less likely they are to develop cognitive decline, given age, gender, education, income and general health. To arrive at this conclusion, we exploited (presumably) exogenous variability across time and space of pension eligibility rules. To clarify, we estimated the causal effect of cognition of retirement for those individuals who were induced or forced to retire by these rules. If the effect is the same across the population, one can extend this result to all retirees. If however, the effect varies across individuals, what we estimate is the average effect for those who are close to the eligibility thresholds (Local Average Treatment Effect – LATE).

So far we have grouped together different individuals who retire as soon as possible (as soon as they qualify for an early retirement pension) and individuals who instead retire as late as possible (when they reach normal or old age retirement pension eligibility, that is normally associated with mandatory retirement from the job). If the effects of retirement on cognition are heterogeneous, the estimates we presented therefore are a weighted average of causal effects for instrument-specific compliers where the weights depend on the strength of each instrument in the first stage (see Angrist and Pischke, 2009).[16]

We might expect retirement to have different effects for such widely different groups of individuals, and this is consistent with the rejection of the over-identifying restrictions (Sargan-Hansen) test. In this section we separate these two groups by using one set of instruments at the time: early retirement pension eligibility on the one hand, statutory retirement age on the other.

Table 6 here

Table 7 here

We therefore report in Tables 6 (and 7) TSLS estimates using early (statutory) retirement eligibility ages as instruments for the four different specifications of the age effect that we considered in Table 5 (first stage estimates for the logarithmic specification are reported in the appendix, Table A.4).

Table 6 (columns 3 to 6) shows that for individuals who retire as soon as possible retirement is beneficial: as seen in Table 4, being retired significantly reduces the probability of experiencing a high decline in cognitive abilities. However, in this case years into retirement do not have the positive, significant effect that we have seen in Tables 4 and 5, while transiting into retirement has a negative, significant effect. At the bottom of the table we report the Angrist Pischke weak identification tests – that are all above ten. Given that the number of instruments equals the number of endogenous explanatory variables (just identification) we cannot compute the Hansen test. Also, we know that in this case weak instruments are less of a problem, because the IV estimator is median-unbiased. Table 7 (columns 3 to 6) instead reports estimates using statutory retirement eligibility instruments. For individuals who retire as late as possible, the fact of being retired or transiting from work to retirement has a positive effect on the probability of experiencing cognitive decline, in contrast to what we reported in Table 4. There is also a strong, positive cumulative effect on cognitive

[16] In a simple case where there are two instruments, $z_{1i}$ and $z_{2i}$, if the population first stage fitted values for TSLS are given by $D_i=\gamma_{11}z_{1i} + \gamma_{12}z_{2i}$, then the TSLS coefficient to be estimated can be expressed as follows

$$\rho = w\,\rho_1 + (1-w)\,\rho_2,$$

where $\rho_1$ and $\rho_2$ are the instrument-specific LATE using $z_1$ and $z_2$ respectively, and

$$w = \gamma_{11}cov(D_i, z_{1i})/[\gamma_{11}cov(D_i, z_{1i}) + \gamma_{12}cov(D_i, z_{2i})].$$

decline. The instruments in this case are much stronger compared to the previous table (see also Table A.5 in Appendix A for first stage estimates).

Table 8 here

In Table 8 we address the issue of whether the parameter estimates reported in Tables 6 and Table 7 are significantly different for our baseline logarithmic specification. To compute the standard errors of the difference we follow a bootstrap approach. We draw 1000 bootstrap samples stratified by country, gender and the initial wave. Table 8 shows that the differences between the estimated coefficients on all three retirement-related variables are statistically different from zero.

Additionally in Appendix A (Table A.6, A.7 and A.8) we report estimation results from a sensitivity analysis: we drop one country at a time to see whether results change according to the country excluded. Even if in some cases we lose precision, the signs and the magnitude of the key parameters do not change.

### 5.3 Robustness analysis

In this section, we present estimation results for some of the many robustness analyses that we performed. We later discuss the important issue of non-random attrition and its consequences on our analysis.

Table 9 here

In Table 9 we show robustness analyses related to gender. In this table as in the following ones the first column presents parameter estimates based on the full set of instruments. The second column reports estimates when only the early retirement instruments are used (ER) and the third column estimates when only the statutory retirement instruments are used (SR).

We see from the first column of Table 9 that all interaction terms of the retirement variables with the female dummy are insignificant. This suggests that cognitive decline is similarly affected by retirement for men and women (at least, for those women with at least fifteen years of work experience). Column (2) and (3) do not reveal significant gender differences, and confirm the pattern of effects already discussed.

Table 10 here

In Table 10 we address the issue of whether selective mortality may somehow influence our results. It is well known that the rich survive longer, and – to the extent that they are eligible to retire earlier or later than the rest of the population – this could affect our TSLS estimates of the parameters of interest. The table reports estimates restricting the sample to individuals aged 50-80 and shows that results are not affected by dropping the oldest old from the estimation sample.

Table 11 here

We additionally investigate the effect of adopting a different definition of retirement in Table 11. In our analysis, we have so far relied on a self-reported retirement status, without any correction for work activities in the four weeks prior to the interview. Such corrections are instead normally implemented in the literature. We report in Table 11 TSLS estimates when we consider as retired those individuals who are permanently out of the labour force, i.e. not only those who declare to be retired from work, but additionally also those who did not do any paid work in the four weeks before the baseline interview (in wave 1 or 2) and in the year before the interview in wave 4. With this alternative definition we do not include among the retired those individuals who might report being retired, simply because they left their *career job*, even if they work full-or part-time (Coe and Zamarro, 2011). We can see that results point to the same conclusions: *log (1+yearsinR)* continues to be highly significant with positive effect on the probability of observing a high decrease in cognition (column (1)), with no noticeable differences in magnitude. In this case, too, this result is driven by those who retire as late as possible (column (2) and (3)).

Last, but by no means least, we address the point of non-random attrition and selection. Mazzonna and Peracchi (2012) and Zamarro et al (2008) notice that panel attrition might be a problem because people in poor health and with low cognitive abilities are more likely to exit the panel, and this may lead to invalid inference. In the remainder of this section we show that this is not an issue in our case.

Table 12 here

Table 13 here

If there is non-random attrition, we would expect differences in memory score at baseline between those staying in the panel and those dropping out. In Table 12 we present the evidence on this score. The first column refers to individuals who entered the sample in wave 1, have not been coded as dead between wave 1 and 4, have done any paid work, are retired or employed/self-employed and aged 50+. The same selection applies in column (2) for those entering the sample in the second wave. We regress memory score in baseline (the sum of words recalled in the first and second trial) on country dummies and a binary indicator taking value 1 if we observe the individual also in wave 4, and zero otherwise. In column (1) we can see that individuals dropping out have a lower memory score than those staying in the panel (as noted in Zamarro et al, 2008); remarkably, no such difference emerges from column (2). The selection issue therefore seems to affect only wave 1 respondents. Similar conclusions can be drawn when looking at the Kolmogorov-Smirnov test for equality of memory score distributions between individuals observed and not observed in wave 4 (Table 13).

Given the evidence of sample selection affecting wave 1 respondents in our longitudinal sample, we next show the results of estimating the model correcting for endogenous selectivity by the standard two-step Heckman procedure (see Wooldrige, 2002). We exploit information from a final section of the questionnaire that should be completed by the interviewer him/herself, known as IV section, providing information on how the interview proceeded and on the general surroundings (living area and type of building). In particular, we use information on the interviewer average time of completion of the IV section to generate additional exogenous variables that determine selection. This set of variables includes the length of the IV section[17] (*length_iv_m*), its square and a dummy variable indicating whether this information is missing[18]. The average time that interviewers need to complete the IV section should capture interviewers' characteristics such as the extra burden in terms of interview length or characteristics which are unobserved but might play a key role in gaining survey cooperation (Korbmacher and Schroeder, 2013; Krosnick, 1991). It is reasonable to assume that interviewer specific information, captured by the IV module, does not affect respondents' cognitive impairment. As De Luca and Peracchi (2012) observe, these variables are external to the individuals under investigation and are not under their control, they are therefore expected to be irrelevant for the performance in the memory tests.

We include among covariates of the selection equation demographic controls (age, gender, education), country dummies, a poor health indicator, variables capturing activities, quartile household income dummies,

[17] The variable is defined at the interviewer level (average time of completion in seconds).

[18] This information can be computed from the so-called keystroke files. In those text files, every time a key is pressed on the keyboard of the laptop, this is registered and stored by the software. Missing information is due to missing interviewer id or too few completed interviews (less than five) that would not provide a meaningful value for the interviewer time of completion.

the set of instruments we use for retirement and the length of IV section variables described above that we will exclude from the cognitive decline analysis. Focusing on column (1) of Table 14, we can see that the IV module length variables are highly significant capturing a non-linear effect.

In column (2) we report TSLS estimates for the effect of retirement on cognitive decline where the Inverse of the Mills Ratio is included. Compared to column (2) of Table 4, the set of controls is reduced, as variables capturing changes in health and activities - that should be also included in the selection equation - cannot be defined for those dropping out the panel. We report therefore in column (3), for comparability, TSLS estimates for individuals entering the sample in wave 1, where we do not control for selection. Comparing columns (2) and (3) of Table 14, we can see that results do not change: years in retirement has a positive and statistically significant effect on the probability of experiencing a large decline in cognition.

Table 15 here

Table 16 here

In Table 15 and 16 we report estimates when using separately early and statutory retirement instruments respectively. Results also in this case are in line with Tables 6 and 7 estimates, even if we have weak instruments in the early retirement case.

## 6 Conclusions

In this paper we have used a new measure of cognitive decline that is highly predictive of the onset of dementia and can be computed in standard surveys where recall memory tests are administered to the same individuals over the years. It is based on a 20% drop in words recalled and is arguably more appropriate in presence of re-testing effects, typically found in longitudinal data: respondents tend to improve their performance in memory tests across waves, particularly the second time they are interviewed. We argued that high decreases in words recalled are informative about actual declines. In fact, using a small sample of individuals aged 70 or more who took part in the US Aging, Demographics and Memory Study and were later medically assessed for dementia, we showed that the 20% decline measure is highly predictive of the onset of dementia. Our test correctly classifies 70% of individuals according to their later dementia status.

We have used SHARE data, that cover ten different European countries, to show that there is a strong, positive association between cognitive decline and years in retirement after controlling for age, physical health, income, education and early-life conditions. Using a plausible identification strategy that exploits country and time variability in pension eligibility to instrument retirement, we have estimated an even stronger causal effect of years in retirement on cognitive decline. We therefore support on European data the cumulative negative effect hypothesis documented in US data: retirement duration plays a role in the evolution of cognitive decline at older ages, over and above the pure age effect, even in the presence of learning effects.

In our analysis we investigate heterogeneity of the effect of retirement on the risk of cognitive decline at older ages. We distinguish those who retire as soon as possible (early retirement) from those instead who retire as late as possible (statutory retirement). For the former group, we find that retirement has beneficial effects on cognition; for the latter instead it has a detrimental effect that gets worse over time. We conclude that the overall estimated effect is largely driven by the group of individuals who are forced to retire as they reach statutory retirement age. To the extent that early retirement is a free choice, but statutory retirement is not, this finding is consistent with consumer rationality.

## References


Aartsen, M. J., Smiths, C. H. M., Van Tilburg, T., Knopscheer, K. C. P. M. and Deeg, D. J. H. (2002) Activity in older adults: Cause or consequence of cognitive functioning? a longitudinal study on everyday activities and cognitive performance in older adults, Journal of Gerontology: Psychological Science, 57, 153–162.

Adam, S., Bonsang, E., Germain, S. and Perelman, S. (2007) Retirement and cognitive reserve: a stochastic frontier approach applied to survey data, Tech. rep., HEC-ULg, CREPP working papers 2007/04.

American Psychiatric Association (2000) Diagnostic and statistical manual of mental disorders, Washington DC: American Psychiatric Association, http://www.psychiatry.org/practice/dsm.

Angelini, V., Brugiavini, A. and Weber, G. (2009) Ageing and unused capacity in Europe: is there an early retirement trap?, Economic Policy, 24, 463–508.

Angrist, J., and Pischke, J. (2009)  Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press.Atchley, R. C. (1976) The Sociology of Retirement, Halsted Press.

Atchley, R. C. (1982) Retirement as a sociological institution, Annual Review of Sociology, 8, 263–287.

Bäckman, L., Jones, S., Berger, A. K., Laukka, E. J. and Small, B. J. (2005) Cognitive impairment in preclinical Alzeimer's disease: a meta-analysis, Neuropsychology, 19, 520–531.

Battistin, E., Brugiavini, A., Rettore, E. and Weber, G. (2009) The retirement consumption puzzle: Evidence from a regression discontinuity approach, American Economic Review, 99, 2209–26.

Baum, C., Schaffer, M., and S. Stillman (2007) Enhanced Routines for Instrumental Variables/GMM Estimation and Testing. Boston College Economics Working Paper, No. 667.

Berkel B., and Börsch-Supan A. (2004) Pension reform in Germany: The impact on retirement decision, FinanzArchiv: Public Finance Analysis, 60, 393-421.

Bingley, P. and Martinello, A. (2013) Mental retirement and schooling, European Economic Review, 63, 292298.

Börsch-Supan, A. and M. Schuth (2013) Early retirement, mental health and social networks, in Active Ageing and Solidarity between Generations in Europe: First Results from SHARE after the Economic Crisis (Eds.) H. L. A. Boersch-Supan, M. Brandt and G. Weber, De Gruyter, Berlin, 337-348.

Bohannon, R. W. (2008) Hand-grip dynamometry predicts future outcomes in aging adults, Journal of Geriatric Physical Therapy, 31, 3–10.

Bonsang, E., Adam, S. and Perelman, S. (2012) Does retirement affect cognitive functioning?, Journal of Health Economics, 31, 490–501.

Bianchini L. and M. Borella, 2014. Cognitive Functioning and Retirement in Europe, CeRP Working Papers 139.

Brugiavini, A., Trevisan, E., and Weber G. (2014) Workability of the young old. Presented at the Kick off conference 'Population aging: economics, health, retirement and the welfare state', Padua, 23rd June 2014.

Brunello, G., Weber, G. and Weiss, C. T. (2012) Books are forever: Early life conditions, education and lifetime income, IZA Discussion Papers 6386. Forthcoming: Economic Journal.

Charles, K. K. (2004) Is retirement depressing? Labour force inactivity and psychological wellbeing in later life, Research in Labour Economics, 23, 269–299.

Chow, G. D. (1960) Test of equality between sets of coefficients in two linear regressions, Econometrica, 28, 591–603.

Coe, N. B., von Gaudecker, H.-M., Lindeboom, M. and Maurer, J. (2012) The effect of retirement on cognitive functioning, Health Economics, 21, 913927.

Coe, N. B. and Zamarro, G. (2011) Retirement effects on health in Europe, Journal of Health Economics, 30, 77–86.

Dal Bianco, C., Garrouste, C. and Paccagnella, O. (2013) Early-life circumstances and cognitive functioning dynamics in later life, in Active Ageing and Solidarity between Generations in Europe: First Results from SHARE after the Economic Crisis (Eds.) H. L. A. Boersch-Supan, M. Brandt and G. Weber, De Gruyter, Berlin, 209-224.

Dave, D., Rashad, I. and Spasojevic, J. (2008) The effects of retirement on physical and mental health outcomes, Southern Economic Journal, 75, 497–523.

Dixon, R., Backman, L. and Nilsson, L. G. (2004) New Frontiers in Cognitive Aging, Oxford University Press.

Drentea, P. (2002) Retirement and mental health, Journal of Aging and Health, 14, 167–194.

Ekerdt, D. J., Bosse, R. and Locastro, J. S. (1983) Claims that retirement improves health, Journal of Gerontology, 38, 231–236.

Estevez-Gonzalez, A., Kulisevsky, J., Boltes, A., Otermin, P. and Garcia-Sanchez, C. (2003) Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: comparison with mild cognitive impairment and normal aging, International Journal of Geriatric Psychiatry, 18, 1021–1028.

Ferrer, E., Salthouse, T. A., Stewart, W. F. and Schwartz, B. S. (2004) Modeling age and retest processes in longitudinal studies of cognitive abilities, Psycology and Aging, 19, 243259.

Fratiglioni, L., Von Strauss, E. and Qiu, C. X. (2008) Epidemiology of dementias of old age, in The Oxford Textbook of Old Age Psychiatry (Eds.) T. Dening, R. Jacoby, C. Oppenheimer and A. Thomas, Oxford University Press, New York.

Gall, L. T., Evans, D. R. and Howard, J. (1997) The retirement adjustment process: changes in the well-being of male retirees across time, Journal of Gerontology, 52B, 110–117.

Grossman, M. (1972) On the concept of health capital and the demand for health, The Journal of Political Economy, 80, 223–255.

Gruber J., and Wise D.A. (1999), Social Security Programs and Retirement around the World. University of Chicago Press, Chicago.

Gruber J., and Wise D.A. (2007), Social Security Programs and Retirement around the World. Fiscal implications of the reform. University of Chicago Press, Chicago.

Heckman, J. and Cunha, F. (2007) The technology of skill formation, American Economic Review, 97, 31–47.

Hurd, M. D., Martorell, P., Delavande, A., Mullen, K. J. and Langa, K. M. (2013) Monetary costs of dementia in the united states, The New England Journal of Medicine, 368, 1326–1334.

Imbens, G. W., and J. M. Wooldridge (2009) Recent Developments in the Econometrics of Program Evaluation. Journal of Economic Literature, 47, 5-86.

Insler, M., 2014. "The Health Consequences of Retirement," Journal of Human Resources, University of Wisconsin Press, vol. 49(1), 195-233.

Maguire, E., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J. and Frith, C. D. (2000) Navigation-related structural change in the hippocampi of taxi drivers, Tech. rep., Proceedings of the National Academy of Sciences of the United States of America 97: 4398-4403.

Malter, F. and Boersch-Supan, A. (2013) Share wave 4: Innovations & methodology, Tech. rep., Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Mazzonna, F. and Peracchi, F. (2012) Ageing, cognitive abilities and retirement, European Economic Review, 56, 691–710.

Mazzonna, F. and Peracchi, F. (2014) Unhealthy Retirement?, EIEF Working Paper 09/14.

McArdle, J. J. and Woodcock, J. R. (1997) Expanding test-retest designs to include developmental time-lag components, Psychological Methods, 2, 403–435.

Mein, G., Martikainen, P., Hemingway, H., Stansfeld, S. and Marmot, M. (2003) Is retirement good or bad for mental and physical health functioning? whitehall ii longitudinal study of civil servants, Journal of Epidemiology and Community Health, 57, 46–49.

Minkler, M. (1981) Research on the health effects of retirement: An uncertain legacy, Journal of Health and Social Behavior, 22, 117–130.

Mojon-Azzi, S., Sousa-Poza, A. and Widmer, R. (2007) The effect of retirement on health: a panel analysis using data from the swiss household panel, Swiss Medical Weekly, 137, 581–585.

Peterson, M., Kramer, A. and Colcombe, A. (2002) Contextual guidance of attention in younger and older adults, Tech. rep., 2002 Cognitive Aging Conference, Atlanta, GA.

Rabbitt, P., Diggle, P., Smith, D., Holland, F. and McInnes, L. (2001) Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents, Neuropsychologia, 39, 532–543.

Rohwedder, S. and Willis, R. J. (2010) Mental retirement, Journal of Economic Perspectives, 24, 119–38.

Salthouse, T. A. (2006) Mental exercise and mental aging, Perspectives on Psychological Science, 1, 68–87.

Salthouse, T. A. (2009) When does age-related cognitive decline begin?, Neurobiology of Aging, 30, 507–514.

Sanderson E. and F. Windmeijer (2015) A weak instrument F-test in linear IV models with multiple endogenous variables, Journal of Econometrics, forthcoming.

Satz, P. (2009) Brain reserve capacity on symptom onset after brain injury: A formulation and review of evidence for threshold theory, Neuropsychologia, 7, 273–295.

Schaie, K. W. (1989) Individual differences in rate of cognitive change in adulthood, in The Course of Later Life: Research and Reflections (Eds.) V. L. Bengston and K. W. Schaie, Springer, New York.

Schaie, T. A. (1994) The course of adult intellectual development, American Psychologist, 49, 304–313.

Schaie, T. A. (1996) Intellectual development in adulthood, Cambridge University Press.

Staubli S., and Zweimueller J. (2012), Does raising the early retirement age increase employment of older workers?, Journal of Public Economics, 108, 17-32.

Stern, Y. (2009) Cognitive reserve, Neuropsychologia, 47, 2015–2028.

Steeves, J., Murphy, R., Zippunikov, V., Strath, S. and Harris, T. (2014) Women Workers and Women at home are equally inactive: NHANES 2003-2006, Medicine and science in sports and exercise, *forthcoming*

Stock J. and M. Yogo. (2005) Testing for Weak Instrument in Linear IV Regression. Andrews DWK Identification and Inference for Econometric Models, 80-108.

Tulving, E. (1972) Episodic and semantic memory, in Organization of memory (Eds.) E. Tulving and W. Donaldson, Academic Press, New York.

Westerlund, H., Vahtera, J., Ferrie, J. E., Singh-Manoux, A., Pentti, J., Melchior, M., Leineweber, C., Jokela, M., Siegrist, J., Goldberg, M. and Kivimaki, M. (2010) Effect of retirement on major chronic conditions and fatigue: French GAZEL occupational cohort study. In: British Medical Journal, Vol. 341, c6149

Wooldrige, J.M. (2002) Econometric analysis of cross section and panel data, Cambridge, MA: MIT Press.

# Tables

**Table 1. Summary statistics by country**

| | Age | Retired | Transitions into retirement | Years in retirement (if retired) | Early ret. age Males | Early ret. age Females | Normal ret. age Males | Normal ret. age Females |
|---|---|---|---|---|---|---|---|---|
| | Mean | % | % | Mean | Mean | Mean | Mean | Mean |
| Austria | 63 | 0.72 | 0.18 | 8.2 | 61 | 56 | 65 | 60 |
| Germany | 63 | 0.57 | 0.18 | 7.2 | 62 | 61 | 65 | 65 |
| Sweden | 64 | 0.52 | 0.22 | 6.7 | 61 | 61 | 65 | 65 |
| Netherlands | 62 | 0.48 | 0.21 | 5.5 | 60 | 60 | 65 | 65 |
| Spain | 65 | 0.62 | 0.14 | 8.1 | 61 | 61 | 65 | 65 |
| Italy | 61 | 0.60 | 0.19 | 6.0 | 56 | 56 | 64 | 59 |
| France | 63 | 0.60 | 0.19 | 9.2 | 55 | 55 | 60 | 60 |
| Denmark | 64 | 0.50 | 0.18 | 5.9 | 60 | 60 | 66 | 66 |
| Switzerland | 60 | 0.32 | 0.21 | 3.5 | 63 | 62 | 65 | 63 |
| Belgium | 63 | 0.59 | 0.20 | 8.0 | 60 | 60 | 65 | 62 |

**Table 2. Percentage decrease in memory score between waves**

| | Males and Females | | | | | | |
|---|---|---|---|---|---|---|---|
| Waves | Percentiles | | | | | | |
| | 5 | 10 | 25 | 50 | 75 | 90 | 95 |
| W1-W4 | -0.615 | -0.429 | -0.214 | 0.000 | 0.300 | 0.714 | 1.000 |
| W2-W4 | -0.600 | -0.444 | -0.231 | 0.000 | 0.250 | 0.625 | 1.000 |
| | Males and Females 80- | | | | | | |
| | Percentiles | | | | | | |
| | 5 | 10 | 25 | 50 | 75 | 90 | 95 |
| W1-W4 | -0.584 | -0.417 | -0.200 | 0.000 | 0.300 | 0.714 | 1.000 |
| W2-W4 | -0.571 | -0.429 | -0.222 | 0.000 | 0.273 | 0.625 | 1.000 |

**Table 3. Predicted probability of cognitive status versus assessed cognitive status**

| | | Predicted | | |
|---|---|---|---|---|
| | | Normal | With dementia | Total |
| Assessed | Normal | 62.15% | 37.85% | 100 % |
| | With dementia | 24.71% | 75.29% | 100 % |
| | Total | 40.98% | 59.02% | 100 % |
| | Correctly classified | | | 70 % |

**Table 4. Cognitive decline - OLS and TSLS estimates**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Model 1 | | Model 2 | |
| | OLS | TSLS | OLS | TSLS |
| retired | -0.041** | -0.089** | -0.041** | -0.089** |
| | (0.017) | (0.037) | (0.018) | (0.038) |
| fromWtoR | -0.010 | -0.012 | -0.006 | -0.005 |
| | (0.013) | (0.055) | (0.013) | (0.056) |
| log (1+yearsinR) | 0.025*** | 0.056* | 0.030*** | 0.064** |
| | (0.009) | (0.031) | (0.009) | (0.032) |
| log age | 0.588*** | 0.510** | 0.541*** | 0.441* |
| | (0.071) | (0.238) | (0.075) | (0.250) |
| low_cognition | -0.173*** | -0.174*** | -0.180*** | -0.182*** |
| | (0.010) | (0.010) | (0.010) | (0.010) |
| female | -0.029*** | -0.030*** | -0.026*** | -0.026** |
| | (0.009) | (0.009) | (0.010) | (0.010) |
| wave 2 | -0.009 | -0.008 | -0.014 | -0.014 |
| | (0.020) | (0.021) | (0.020) | (0.021) |
| less repetitions | 0.002 | -0.000 | -0.002 | -0.005 |
| | (0.019) | (0.019) | (0.019) | (0.019) |
| eqincomeQ2 | -0.033** | -0.032** | -0.032** | -0.032** |
| | (0.016) | (0.016) | (0.016) | (0.016) |
| eqincomeQ3 | -0.040*** | -0.041*** | -0.038*** | -0.039*** |
| | (0.015) | (0.015) | (0.015) | (0.015) |
| eqincomeQ4 | -0.036** | -0.037** | -0.036** | -0.037** |
| | (0.015) | (0.015) | (0.015) | (0.015) |
| highschool | -0.035*** | -0.036*** | -0.022* | -0.021* |
| | (0.011) | (0.011) | (0.011) | (0.011) |
| college | -0.074*** | -0.074*** | -0.043*** | -0.042*** |
| | (0.013) | (0.013) | (0.013) | (0.014) |
| poorhealth_bl | -0.003 | -0.003 | -0.009 | -0.009 |
| | (0.034) | (0.034) | (0.034) | (0.034) |
| drophealth | 0.065** | 0.062** | 0.049* | 0.047* |
| | (0.026) | (0.026) | (0.026) | (0.026) |
| increasehealth | -0.006 | -0.008 | -0.012 | -0.014 |
| | (0.035) | (0.035) | (0.034) | (0.034) |
| physicalact_bl | -0.036*** | -0.034*** | -0.031** | -0.029** |
| | (0.012) | (0.012) | (0.012) | (0.012) |
| increasephysicalact | -0.035** | -0.033** | -0.032** | -0.030** |
| | (0.015) | (0.015) | (0.015) | (0.015) |
| dropphysicalact | 0.045*** | 0.045*** | 0.040*** | 0.040*** |
| | (0.013) | (0.013) | (0.012) | (0.013) |
| dailyact_bl | -0.091*** | -0.089*** | -0.071*** | -0.068*** |
| | (0.018) | (0.019) | (0.018) | (0.018) |
| increasedailyact | -0.075*** | -0.074*** | -0.060*** | -0.058*** |
| | (0.013) | (0.013) | (0.013) | (0.013) |
| dropdailyact | 0.072** | 0.072** | 0.054 | 0.054 |
| | (0.035) | (0.036) | (0.035) | (0.035) |
| jobexperience40 | | | 0.020** | 0.024* |
| | | | (0.010) | (0.013) |
| missingincome_bl | | | 0.045*** | 0.045*** |
| | | | (0.013) | (0.013) |
| missingincomeW4 | | | 0.053*** | 0.053*** |
| | | | (0.015) | (0.015) |
| nomissingincomeW4 | | | -0.041*** | -0.041*** |
| | | | (0.012) | (0.012) |
| fewbooks | | | 0.017 | 0.017 |
| | | | (0.011) | (0.012) |
| mathskills | | | -0.038*** | -0.038*** |
| | | | (0.009) | (0.009) |
| ruralarea | | | 0.025*** | 0.024*** |
| | | | (0.009) | (0.009) |
| public | | | -0.018 | -0.020 |
| | | | (0.011) | (0.012) |
| selfemployed | | | 0.003 | 0.003 |
| | | | (0.015) | (0.016) |
| partner | | | 0.011 | 0.015 |
| | | | (0.012) | (0.012) |
| nogstest_bl | | | 0.018 | 0.016 |
| | | | (0.055) | (0.055) |
| dropgstest | | | 0.087*** | 0.086*** |

| | | | | |
|---|---|---|---|---|
| | | | (0.029) | (0.030) |
| increasegstest | | | 0.017 | 0.019 |
| | | | (0.064) | (0.064) |
| alonecftest_bl | | | -0.076*** | -0.077*** |
| | | | (0.029) | (0.029) |
| dropalonecftest | | | 0.033* | 0.032* |
| | | | (0.017) | (0.017) |
| increasealonecftest | | | -0.048 | -0.047 |
| | | | (0.031) | (0.031) |
| contextcftest_bl | | | 0.130 | 0.128 |
| | | | (0.084) | (0.084) |
| increasecontextcftestW4 | | | -0.153* | -0.151* |
| | | | (0.085) | (0.085) |
| dropcontextcftestW4 | | | 0.148*** | 0.148*** |
| | | | (0.030) | (0.030) |
| eurodcat_bl | | | 0.022 | 0.022 |
| | | | (0.017) | (0.017) |
| increaseeurodcat | | | -0.012 | -0.013 |
| | | | (0.021) | (0.021) |
| dropeurodcat | | | 0.025* | 0.023 |
| | | | (0.015) | (0.015) |
| dailycontactchild_bl | | | -0.016 | -0.018 |
| | | | (0.012) | (0.013) |
| increase dailycontactchild | | | -0.009 | -0.010 |
| | | | (0.017) | (0.017) |
| dropdailycontactchild | | | 0.010 | 0.010 |
| | | | (0.014) | (0.014) |
| SE | -0.012 | -0.010 | 0.009 | 0.012 |
| | (0.019) | (0.021) | (0.019) | (0.021) |
| DK | 0.005 | 0.010 | 0.019 | 0.024 |
| | (0.020) | (0.022) | (0.020) | (0.022) |
| NL | 0.001 | 0.004 | 0.008 | 0.011 |
| | (0.022) | (0.022) | (0.022) | (0.023) |
| BE | -0.074*** | -0.076*** | -0.074*** | -0.075*** |
| | (0.020) | (0.019) | (0.020) | (0.020) |
| FR | -0.089*** | -0.094*** | -0.089*** | -0.092*** |
| | (0.018) | (0.019) | (0.018) | (0.019) |
| CH | -0.042* | -0.039 | -0.025 | -0.021 |
| | (0.024) | (0.026) | (0.024) | (0.026) |
| AT | 0.055* | 0.050 | 0.052* | 0.047 |
| | (0.031) | (0.033) | (0.031) | (0.033) |
| ES | -0.042 | -0.042 | -0.034 | -0.034 |
| | (0.026) | (0.026) | (0.026) | (0.026) |
| IT | -0.089*** | -0.090*** | -0.078*** | -0.078*** |
| | (0.022) | (0.023) | (0.023) | (0.024) |
| constant | -1.917*** | -1.599* | -1.733*** | -1.330 |
| | (0.286) | (0.933) | (0.300) | (0.973) |
| | | | | |
| Observations | 8,221 | 8,221 | 8,221 | 8,221 |
| R-squared | 0.091 | 0.089 | 0.107 | 0.105 |
| Adj R-squared | 0.087 | 0.086 | 0.101 | 0.099 |
| | | | | |
| Sargan-Hansen (p-value) | | 0.019 | | 0.023 |
| F-test joint significance (p-value) | | | | |
| retirement variables | 0.011 | 0.000 | 0.003 | 0.000 |
| | | | | |
| Angrist and Pischke, first-stage F stat (Weak identification test) | | | | |
| retired | | 153.831 | | 144.887 |
| fromWtoR | | 51.484 | | 50.028 |
| log(1+yearsinR) | | 51.982 | | 51.890 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table 5. Cognitive decline - TSLS estimates - Different functional forms for age

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | TSLS - Model 1 | | | |
| | | Age polynomial | | |
| | Logarithm | Degree: 1 | Degree: 2 | Degree: 3 |
| retired | -0.089** | -0.003 | -0.057 | -0.018 |
| | (0.037) | (0.061) | (0.074) | (0.076) |
| fromWtoR | -0.012 | -0.005 | -0.036 | 0.038 |
| | (0.055) | (0.050) | (0.054) | (0.069) |
| log (1+yearsinR) | 0.056* | | | |
| | (0.031) | | | |
| yearsinR | | 0.007 | 0.012** | 0.012** |
| | | (0.005) | (0.006) | (0.006) |
| age | | 0.006 | 0.040* | -0.186 |
| | | (0.005) | (0.021) | (0.118) |
| age^2/10 | | | -0.003* | 0.030* |
| | | | (0.002) | (0.017) |
| age^3/100 | | | | -0.002* |
| | | | | (0.001) |
| log age | 0.510** | | | |
| | (0.238) | | | |
| low cognition | -0.174*** | -0.174*** | -0.175*** | -0.175*** |
| | (0.010) | (0.010) | (0.010) | (0.010) |
| female | -0.030*** | -0.031*** | -0.032*** | -0.032*** |
| | (0.009) | (0.009) | (0.009) | (0.009) |
| wave 2 | -0.008 | -0.006 | -0.010 | -0.000 |
| | (0.021) | (0.021) | (0.021) | (0.022) |
| less_repetitions | -0.000 | 0.001 | 0.001 | -0.001 |
| | (0.019) | (0.019) | (0.019) | (0.019) |
| eqincomeQ2 | -0.032** | -0.032** | -0.032** | -0.034** |
| | (0.016) | (0.016) | (0.016) | (0.016) |
| eqincomeQ3 | -0.041*** | -0.039*** | -0.041*** | -0.044*** |
| | (0.015) | (0.015) | (0.015) | (0.015) |
| eqincomeQ4 | -0.037** | -0.035** | -0.037** | -0.038** |
| | (0.015) | (0.015) | (0.015) | (0.015) |
| highschool | -0.036*** | -0.035*** | -0.035*** | -0.036*** |
| | (0.011) | (0.011) | (0.011) | (0.011) |
| college | -0.074*** | -0.072*** | -0.074*** | -0.072*** |
| | (0.013) | (0.013) | (0.014) | (0.014) |
| poorhealth_bl | -0.003 | -0.007 | -0.007 | -0.005 |
| | (0.034) | (0.034) | (0.034) | (0.035) |
| drophealth | 0.062** | 0.062** | 0.061** | 0.060** |
| | (0.026) | (0.026) | (0.026) | (0.026) |
| increasehealth | -0.008 | -0.008 | -0.010 | -0.009 |
| | (0.035) | (0.035) | (0.035) | (0.035) |
| physicalact_bl | -0.034*** | -0.032*** | -0.033*** | -0.033*** |
| | (0.012) | (0.012) | (0.012) | (0.012) |
| increasephysicalact | -0.033** | -0.031** | -0.030** | -0.032** |
| | (0.015) | (0.015) | (0.015) | (0.015) |
| dropphysicalact | 0.045*** | 0.045*** | 0.046*** | 0.044*** |
| | (0.013) | (0.013) | (0.013) | (0.013) |
| dailyact_bl | -0.089*** | -0.090*** | -0.090*** | -0.088*** |
| | (0.019) | (0.019) | (0.019) | (0.019) |
| increasedailyact | -0.074*** | -0.073*** | -0.073*** | -0.074*** |
| | (0.013) | (0.013) | (0.013) | (0.013) |
| dropdailyact | 0.072** | 0.071** | 0.071** | 0.072** |
| | (0.036) | (0.035) | (0.036) | (0.036) |
| SE | -0.010 | -0.006 | -0.004 | -0.003 |
| | (0.021) | (0.022) | (0.022) | (0.021) |
| DK | 0.010 | 0.011 | 0.017 | 0.017 |
| | (0.022) | (0.023) | (0.023) | (0.022) |
| NL | 0.004 | 0.005 | 0.009 | 0.008 |
| | (0.022) | (0.023) | (0.023) | (0.023) |
| BE | -0.076*** | -0.078*** | -0.077*** | -0.079*** |
| | (0.019) | (0.020) | (0.020) | (0.020) |
| FR | -0.094*** | -0.095*** | -0.097*** | -0.100*** |
| | (0.019) | (0.020) | (0.020) | (0.020) |
| CH | -0.039 | -0.034 | -0.035 | -0.030 |
| | (0.026) | (0.027) | (0.027) | (0.027) |
| AT | 0.050 | 0.046 | 0.042 | 0.040 |

| | | | | |
|---|---|---|---|---|
| | (0.033) | (0.035) | (0.035) | (0.035) |
| ES | -0.042 | -0.041 | -0.040 | -0.042 |
| | (0.026) | (0.026) | (0.026) | (0.026) |
| IT | -0.090*** | -0.091*** | -0.091*** | -0.092*** |
| | (0.023) | (0.024) | (0.024) | (0.023) |
| Constant | -1.599* | 0.084 | -0.934 | 4.079 |
| | (0.933) | (0.244) | (0.667) | (2.667) |
| Observations | 8,221 | 8,221 | 8,221 | 8,221 |
| R-squared | 0.089 | 0.090 | 0.086 | 0.088 |
| Adj R-squared | 0.086 | 0.086 | 0.083 | 0.084 |
| Sargan-Hansen (p-value) | 0.019 | 0.029 | 0.013 | 0.032 |
| F-test joint significance (p-value) | | | | |
| retirement variables | 0.000 | 0.010 | 0.043 | 0.108 |
| age | | | 0.124 | 0.041 |
| Angrist and Pischke, first-stage F stat (Weak identification test) | F(4,687) | F(4,687) | F(4,687) | F(4,687) |
| retired | 153.831 | 36.853 | 46.354 | 38.688 |
| fromWtoR | 51.484 | 58.062 | 59.105 | 42.601 |
| log(1+yearsinR) | 51.982 | 37.843 | 40.324 | 41.858 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table 6. Cognitive decline - Early retirement estimates

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | OLS | | ER -TSLS - Model 1 | | | |
| | | | | Age polynomial: | | |
| | | Reduced form | Logarithm | Degree: 1 | Degree: 2 | Degree: 3 |
| retired | -0.041** | | -0.182*** | -0.169* | -0.216** | -0.237** |
| | (0.017) | | (0.056) | (0.093) | (0.097) | (0.120) |
| fromWtoR | -0.010 | | -0.310* | -0.193* | -0.230** | -0.263* |
| | (0.013) | | (0.186) | (0.105) | (0.089) | (0.140) |
| log (1+yearsinR) | 0.025*** | | -0.122 | | | |
| | (0.009) | | (0.119) | | | |
| yearsinR | | | | -0.010 | -0.009 | -0.009 |
| | | | | (0.010) | (0.011) | (0.012) |
| eligibleER | | -0.092*** | | | | |
| | | (0.020) | | | | |
| fromNEtoE_ER | | -0.063*** | | | | |
| | | (0.018) | | | | |
| log (1+sinceER) | | -0.004 | | | | |
| | | (0.013) | | | | |
| log age | 0.588*** | 0.866*** | 1.781** | | | |
| | (0.071) | (0.121) | (0.814) | | | |
| age | | | | 0.021** | 0.042 | 0.130 |
| | | | | (0.008) | (0.028) | (0.174) |
| age^2/10 | | | | | -0.002 | -0.014 |
| | | | | | (0.002) | (0.025) |
| age^3/100 | | | | | | 0.001 |
| | | | | | | (0.001) |
| low cognition | -0.173*** | -0.172*** | -0.169*** | -0.169*** | -0.171*** | -0.171*** |
| | (0.010) | (0.010) | (0.011) | (0.011) | (0.010) | (0.010) |
| female | -0.029*** | -0.028*** | -0.026** | -0.026** | -0.026*** | -0.026*** |
| | (0.009) | (0.009) | (0.010) | (0.010) | (0.010) | (0.010) |
| wave 2 | -0.009 | -0.015 | -0.040 | -0.040 | -0.029 | -0.033 |
| | (0.020) | (0.020) | (0.030) | (0.030) | (0.024) | (0.023) |
| less_repetitions | 0.002 | 0.004 | 0.009 | 0.009 | 0.005 | 0.005 |
| | (0.019) | (0.019) | (0.020) | (0.020) | (0.019) | (0.019) |
| eqincomeQ2 | -0.033** | -0.032** | -0.023 | -0.023 | -0.026 | -0.026 |
| | (0.016) | (0.016) | (0.017) | (0.017) | (0.016) | (0.016) |
| eqincomeQ3 | -0.040*** | -0.039*** | -0.026 | -0.026 | -0.029* | -0.029* |
| | (0.015) | (0.015) | (0.018) | (0.018) | (0.016) | (0.016) |
| eqincomeQ4 | -0.036** | -0.036** | -0.033** | -0.033** | -0.031** | -0.031** |
| | (0.015) | (0.015) | (0.015) | (0.015) | (0.016) | (0.016) |
| highschool | -0.035*** | -0.034*** | -0.036*** | -0.036*** | -0.036*** | -0.036*** |
| | (0.011) | (0.011) | (0.012) | (0.012) | (0.012) | (0.012) |
| college | -0.074*** | -0.073*** | -0.089*** | -0.089*** | -0.083*** | -0.085*** |
| | (0.013) | (0.013) | (0.016) | (0.016) | (0.014) | (0.014) |
| poorhealth_bl | -0.003 | -0.002 | 0.012 | 0.012 | 0.008 | 0.009 |
| | (0.034) | (0.034) | (0.037) | (0.037) | (0.035) | (0.035) |
| drophealth | 0.065** | 0.064** | 0.078*** | 0.078*** | 0.073*** | 0.073*** |
| | (0.026) | (0.026) | (0.028) | (0.028) | (0.027) | (0.027) |
| increasehealth | -0.006 | -0.003 | 0.005 | 0.005 | 0.000 | 0.000 |
| | (0.035) | (0.035) | (0.037) | (0.037) | (0.036) | (0.036) |
| physicalact_bl | -0.036*** | -0.035*** | -0.044*** | -0.044*** | -0.041*** | -0.042*** |
| | (0.012) | (0.012) | (0.014) | (0.014) | (0.013) | (0.013) |
| increasephysicalact | -0.035** | -0.033** | -0.033** | -0.033** | -0.034** | -0.034** |
| | (0.015) | (0.015) | (0.015) | (0.015) | (0.015) | (0.015) |
| dropphysicalact | 0.045*** | 0.046*** | 0.053*** | 0.053*** | 0.049*** | 0.050*** |
| | (0.013) | (0.013) | (0.014) | (0.014) | (0.013) | (0.013) |
| dailyact_bl | -0.091*** | -0.092*** | -0.092*** | -0.092*** | -0.091*** | -0.091*** |
| | (0.018) | (0.019) | (0.020) | (0.020) | (0.019) | (0.019) |
| increasedailyact | -0.075*** | -0.075*** | -0.071*** | -0.071*** | -0.072*** | -0.071*** |
| | (0.013) | (0.013) | (0.013) | (0.013) | (0.013) | (0.013) |
| dropdailyact | 0.072** | 0.070** | 0.062* | 0.062* | 0.067* | 0.067* |
| | (0.035) | (0.035) | (0.037) | (0.037) | (0.036) | (0.036) |
| SE | -0.012 | -0.008 | -0.050 | -0.050 | -0.036 | -0.038 |
| | (0.019) | (0.019) | (0.034) | (0.034) | (0.027) | (0.027) |
| DK | 0.005 | 0.010 | -0.045 | -0.045 | -0.027 | -0.027 |
| | (0.020) | (0.019) | (0.041) | (0.041) | (0.029) | (0.030) |
| NL | 0.001 | 0.009 | -0.023 | -0.023 | -0.015 | -0.015 |
| | (0.022) | (0.022) | (0.029) | (0.029) | (0.025) | (0.025) |
| BE | -0.074*** | -0.067*** | -0.060** | -0.060** | -0.065*** | -0.063*** |
| | (0.020) | (0.020) | (0.023) | (0.023) | (0.021) | (0.021) |

| | | | | | | |
|---|---|---|---|---|---|---|
| FR | -0.089*** | -0.060*** | -0.058* | -0.058* | -0.070*** | -0.069*** |
| | (0.018) | (0.021) | (0.031) | (0.031) | (0.023) | (0.023) |
| CH | -0.042* | -0.047** | -0.098** | -0.098** | -0.076** | -0.080** |
| | (0.024) | (0.023) | (0.045) | (0.045) | (0.034) | (0.033) |
| AT | 0.055* | 0.070** | 0.128** | 0.128** | 0.100** | 0.103** |
| | (0.031) | (0.031) | (0.059) | (0.059) | (0.043) | (0.042) |
| ES | -0.042 | -0.036 | -0.055** | -0.055** | -0.051* | -0.052* |
| | (0.026) | (0.025) | (0.027) | (0.027) | (0.026) | (0.027) |
| IT | -0.089*** | -0.068*** | -0.041 | -0.041 | -0.062** | -0.059** |
| | (0.022) | (0.023) | (0.040) | (0.040) | (0.027) | (0.026) |
| Constant | -1.917*** | -2.999*** | -6.563** | -6.563** | -0.627 | -1.306 |
| | (0.286) | (0.480) | (3.187) | (3.187) | (0.407) | (0.881) |
| Observations | 8,221 | 8,221 | 8,221 | 8,221 | 8,221 | 8,221 |
| R-squared | 0.091 | 0.092 | 0.034 | 0.070 | 0.064 | 0.057 |
| Adj R-squared | 0.087 | 0.089 | 0.030 | 0.067 | 0.060 | 0.054 |
| F-test joint significance (p-value) | | | | | | |
| retirement variables | 0.011 | 0.000 | 0.000 | 0.002 | 0.016 | 0.177 |
| age | | | | | 0.007 | 0.010 |
| Angrist and Pischke, first-stage F stat (Weak identification test) | | | F(1,687) | F(1,687) | F(1,687) | F(1,687) |
| retired | | | 27.292 | 44.144 | 96.359 | 42.877 |
| fromWtoR | | | 14.190 | 37.290 | 51.761 | 24.824 |
| log(1+yearsinR) | | | 13.426 | 33.811 | 48.270 | 31.306 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table 7. Cognitive decline - Statutory retirement estimates

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | OLS | | SR -TSLS - Model 1 | | | |
| | | | | Age polynomial: | | |
| | | Reduced form | Logarithm | Degree: 1 | Degree: 2 | Degree: 3 |
| retired | -0.041** | | 0.005 | 0.238** | 0.206* | 0.187 |
| | (0.017) | | (0.051) | (0.111) | (0.119) | (0.114) |
| fromWtoR | -0.010 | | 0.116* | 0.125* | 0.109 | 0.153* |
| | (0.013) | | (0.069) | (0.070) | (0.074) | (0.082) |
| log (1+yearsinR) | 0.025*** | | 0.131*** | | | |
| | (0.009) | | (0.040) | | | |
| yearsinR | | | | 0.027*** | 0.031*** | 0.026*** |
| | | | | (0.009) | (0.010) | (0.009) |
| eligibleSR | | -0.017 | | | | |
| | | (0.023) | | | | |
| fromNEtoE_SR | | 0.019 | | | | |
| | | (0.018) | | | | |
| log (1+sinceSR) | | 0.053*** | | | | |
| | | (0.013) | | | | |
| log age | 0.588*** | 0.348*** | -0.175 | | | |
| | (0.071) | (0.117) | (0.340) | | | |
| age | | | | -0.013 | 0.011 | -0.225* |
| | | | | (0.009) | (0.024) | (0.132) |
| age^2/10 | | | | | -0.002 | 0.034* |
| | | | | | (0.002) | (0.019) |
| age^3/100 | | | | | | -0.002* |
| | | | | | | (0.001) |
| low cognition | -0.173*** | -0.174*** | -0.177*** | -0.178*** | -0.179*** | -0.178*** |
| | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) |
| female | -0.029*** | -0.033*** | -0.032*** | -0.036*** | -0.036*** | -0.035*** |
| | (0.009) | (0.009) | (0.009) | (0.010) | (0.010) | (0.010) |
| wave 2 | -0.009 | -0.005 | 0.005 | 0.007 | 0.005 | 0.012 |
| | (0.020) | (0.020) | (0.022) | (0.022) | (0.023) | (0.023) |
| less_repetitions | 0.002 | 0.002 | -0.003 | 0.001 | 0.001 | -0.001 |
| | (0.019) | (0.019) | (0.019) | (0.019) | (0.019) | (0.019) |
| eqincomeQ2 | -0.033** | -0.032** | -0.036** | -0.035** | -0.035** | -0.036** |
| | (0.016) | (0.016) | (0.016) | (0.016) | (0.016) | (0.016) |
| eqincomeQ3 | -0.040*** | -0.040*** | -0.045*** | -0.041** | -0.042*** | -0.044*** |
| | (0.015) | (0.015) | (0.016) | (0.016) | (0.016) | (0.016) |
| eqincomeQ4 | -0.036** | -0.036** | -0.033** | -0.025 | -0.027 | -0.030* |
| | (0.015) | (0.015) | (0.016) | (0.017) | (0.017) | (0.017) |
| highschool | -0.035*** | -0.035*** | -0.035*** | -0.033*** | -0.033*** | -0.034*** |
| | (0.011) | (0.011) | (0.011) | (0.012) | (0.012) | (0.012) |
| college | -0.074*** | -0.073*** | -0.065*** | -0.057*** | -0.057*** | -0.060*** |
| | (0.013) | (0.013) | (0.013) | (0.015) | (0.015) | (0.015) |
| poorhealth_bl | -0.003 | -0.001 | -0.013 | -0.032 | -0.034 | -0.025 |
| | (0.034) | (0.034) | (0.035) | (0.038) | (0.038) | (0.037) |
| drophealth | 0.065** | 0.064** | 0.056** | 0.050* | 0.049* | 0.051* |
| | (0.026) | (0.026) | (0.027) | (0.027) | (0.027) | (0.027) |
| increasehealth | -0.006 | -0.005 | -0.014 | -0.020 | -0.021 | -0.018 |
| | (0.035) | (0.035) | (0.036) | (0.037) | (0.037) | (0.036) |
| physicalact_bl | -0.036*** | -0.036*** | -0.029** | -0.021 | -0.021 | -0.024* |
| | (0.012) | (0.012) | (0.012) | (0.013) | (0.013) | (0.013) |
| increasephysicalact | -0.035** | -0.035** | -0.032** | -0.023 | -0.023 | -0.027* |
| | (0.015) | (0.015) | (0.016) | (0.016) | (0.016) | (0.016) |
| dropphysicalact | 0.045*** | 0.044*** | 0.041*** | 0.041*** | 0.041*** | 0.040*** |
| | (0.013) | (0.013) | (0.013) | (0.013) | (0.013) | (0.013) |
| dailyact_bl | -0.091*** | -0.092*** | -0.091*** | -0.097*** | -0.097*** | -0.093*** |
| | (0.018) | (0.019) | (0.019) | (0.019) | (0.019) | (0.019) |
| increasedailyact | -0.075*** | -0.075*** | -0.075*** | -0.073*** | -0.073*** | -0.074*** |
| | (0.013) | (0.013) | (0.013) | (0.013) | (0.013) | (0.013) |
| dropdailyact | 0.072** | 0.072** | 0.078** | 0.078** | 0.079** | 0.078** |
| | (0.035) | (0.035) | (0.036) | (0.036) | (0.037) | (0.036) |
| SE | -0.012 | -0.014 | 0.013 | 0.035 | 0.039 | 0.030 |
| | (0.019) | (0.019) | (0.022) | (0.028) | (0.028) | (0.027) |
| DK | 0.005 | 0.002 | 0.038 | 0.059** | 0.066** | 0.055* |
| | (0.020) | (0.020) | (0.024) | (0.030) | (0.030) | (0.029) |
| NL | 0.001 | -0.005 | 0.017 | 0.031 | 0.035 | 0.028 |
| | (0.022) | (0.022) | (0.024) | (0.026) | (0.026) | (0.026) |
| BE | -0.074*** | -0.080*** | -0.084*** | -0.094*** | -0.093*** | -0.091*** |
| | (0.020) | (0.020) | (0.020) | (0.022) | (0.022) | (0.022) |

| | | | | | | |
|---|---|---|---|---|---|---|
| FR | -0.089*** | -0.112*** | -0.110*** | -0.122*** | -0.124*** | -0.122*** |
| | (0.018) | (0.020) | (0.020) | (0.024) | (0.024) | (0.023) |
| CH | -0.042* | -0.045* | -0.007 | 0.022 | 0.023 | 0.015 |
| | (0.024) | (0.024) | (0.029) | (0.036) | (0.036) | (0.035) |
| AT | 0.055* | 0.048 | 0.009 | -0.025 | -0.031 | -0.016 |
| | (0.031) | (0.032) | (0.037) | (0.046) | (0.046) | (0.045) |
| ES | -0.042 | -0.047* | -0.034 | -0.026 | -0.026 | -0.031 |
| | (0.026) | (0.026) | (0.026) | (0.028) | (0.028) | (0.028) |
| IT | -0.089*** | -0.098*** | -0.118*** | -0.130*** | -0.132*** | -0.123*** |
| | (0.022) | (0.022) | (0.026) | (0.030) | (0.030) | (0.029) |
| constant | -1.917*** | -0.956** | 1.078 | 1.037** | 0.308 | 5.434* |
| | (0.286) | (0.471) | (1.329) | (0.445) | (0.811) | (2.966) |
| Observations | 8,221 | 8,221 | 8,221 | 8,221 | 8,221 | 8,221 |
| R-squared | 0.091 | 0.092 | 0.072 | 0.044 | 0.039 | 0.055 |
| Adj R-squared | 0.087 | 0.088 | 0.069 | 0.041 | 0.035 | 0.051 |
| F-test joint significance (p-value) | | | | | | |
| retirement variables | 0.011 | 0.000 | 0.000 | 0.002 | 0.009 | 0.017 |
| age | | | | | 0.141 | 0.065 |
| Angrist and Pischke first-stage F stat (Weak identification test) | | | F(1,687) | F(1,687) | F(1,687) | F(1,687) |
| retired | | | 127.127 | 50.541 | 49.672 | 53.566 |
| fromWtoR | | | 87.476 | 64.902 | 63.477 | 81.902 |
| log(1+yearsinR) | | | 104.632 | 51.754 | 66.620 | 65.745 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

**Table 8. Bootstrap estimates - Difference between Early retirement (ER) and Statutory retirement (SR) coefficients (logarithmic specification)**

| | Coef. | Std. Err. | lower bound C.I. | upper bound C.I. |
|---|---|---|---|---|
| retired SR | 0.005 | 0.050 | -0.088 | 0.099 |
| retired ER | -0.182 | 0.056 | -0.306 | -0.077 |
| Difference | 0.187 | 0.061 | 0.079 | 0.321 |
| fromWtoR SR | 0.116 | 0.069 | -0.014 | 0.256 |
| fromWtoR ER | -0.31 | 0.177 | -0.709 | 0.021 |
| Difference | 0.427 | 0.188 | 0.075 | 0.841 |
| log (1+yearsinR) SR | 0.131 | 0.041 | 0.053 | 0.215 |
| log (1+yearsinR) ER | -0.122 | 0.115 | -0.367 | 0.088 |
| Difference | 0.253 | 0.121 | 0.038 | 0.521 |

## Table 9. Cognitive decline - TSLS estimates - Robustness analysis - Gender differences

| | (1) | (2) | (3) |
|---|---|---|---|
| | | Gender interaction | |
| | TSLS | TSLS - ER | TSLS – SR |
| retired | -0.078* | -0.168** | 0.023 |
| | (0.043) | (0.069) | (0.059) |
| fromWtoR | -0.026 | -0.320* | 0.092 |
| | (0.063) | (0.191) | (0.083) |
| log (1+yearsinR) | 0.057* | -0.122 | 0.128*** |
| | (0.032) | (0.121) | (0.042) |
| retired_female | -0.025 | -0.032 | -0.045 |
| | (0.052) | (0.094) | (0.057) |
| fromWtoR_female | 0.033 | 0.029 | 0.047 |
| | (0.060) | (0.101) | (0.076) |
| log (1+yearsinR)_female | 0.001 | 0.007 | 0.008 |
| | (0.025) | (0.049) | (0.027) |
| log age | 0.498** | 1.763** | -0.180 |
| | (0.236) | (0.810) | (0.341) |
| low cognition | -0.175*** | -0.170*** | -0.178*** |
| | (0.010) | (0.011) | (0.010) |
| female | -0.025 | -0.021 | -0.024 |
| | (0.023) | (0.028) | (0.027) |
| wave 2 | -0.007 | -0.039 | 0.006 |
| | (0.021) | (0.030) | (0.022) |
| less_repetitions | -0.001 | 0.009 | -0.004 |
| | (0.019) | (0.021) | (0.019) |
| eqincomeQ2 | -0.033** | -0.024 | -0.037** |
| | (0.016) | (0.017) | (0.016) |
| eqincomeQ3 | -0.041*** | -0.026 | -0.045*** |
| | (0.015) | (0.018) | (0.016) |
| eqincomeQ4 | -0.037** | -0.033** | -0.034** |
| | (0.015) | (0.015) | (0.016) |
| highschool | -0.036*** | -0.036*** | -0.035*** |
| | (0.011) | (0.012) | (0.011) |
| college | -0.075*** | -0.090*** | -0.067*** |
| | (0.013) | (0.015) | (0.013) |
| poorhealth_bl | -0.003 | 0.012 | -0.014 |
| | (0.034) | (0.037) | (0.035) |
| drophealth | 0.062** | 0.078*** | 0.056** |
| | (0.026) | (0.028) | (0.027) |
| increasehealth | -0.009 | 0.004 | -0.015 |
| | (0.035) | (0.037) | (0.036) |
| physicalact_bl | -0.034*** | -0.043*** | -0.030** |
| | (0.012) | (0.014) | (0.012) |
| increasephysicalact | -0.034** | -0.034** | -0.034** |
| | (0.015) | (0.015) | (0.016) |
| dropphysicalact | 0.044*** | 0.052*** | 0.040*** |
| | (0.013) | (0.014) | (0.013) |
| dailyact_bl | -0.088*** | -0.092*** | -0.091*** |
| | (0.019) | (0.020) | (0.019) |
| increasedailyact | -0.074*** | -0.071*** | -0.075*** |
| | (0.013) | (0.013) | (0.013) |
| dropdailyact | 0.072** | 0.062* | 0.078** |
| | (0.036) | (0.037) | (0.036) |
| SE | -0.009 | -0.049 | 0.013 |
| | (0.021) | (0.034) | (0.022) |
| DK | 0.011 | -0.044 | 0.039 |
| | (0.021) | (0.041) | (0.024) |
| NL | 0.003 | -0.024 | 0.016 |
| | (0.022) | (0.029) | (0.024) |
| BE | -0.076*** | -0.060** | -0.084*** |
| | (0.019) | (0.023) | (0.020) |
| FR | -0.094*** | -0.059* | -0.110*** |
| | (0.019) | (0.031) | (0.020) |
| CH | -0.039 | -0.098** | -0.007 |
| | (0.026) | (0.045) | (0.028) |
| AT | 0.051 | 0.128** | 0.011 |
| | (0.033) | (0.059) | (0.037) |
| ES | -0.043* | -0.056** | -0.036 |
| | (0.026) | (0.027) | (0.026) |
| IT | -0.091*** | -0.041 | -0.118*** |
| | (0.023) | (0.040) | (0.026) |

| | | | |
|---|---|---|---|
| constant | -1.552* | -6.495** | 1.095 |
| | (0.922) | (3.173) | (1.333) |
| | | | |
| Observations | 8,221 | 8,221 | 8,221 |
| | | | |
| R-squared | 0.089 | 0.035 | 0.072 |
| Adj R-squared | 0.085 | 0.031 | 0.068 |
| | | | |
| Sargan-Hansen (p-value) | 0.109 | | |
| F-test joint significance (p-value) | | | |
| retirement variables | 0.001 | 0.001 | 0.001 |
| interactions with female dummy | 0.519 | 0.821 | 0.369 |
| | | | |
| Angrist and Pischke, first-stage F stat | F(7,687) | F(1,687) | F(1,687) |
| (Weak identification test) | | | |
| retired | 101.620 | 37.494 | 128.165 |
| fromWtoR | 25.753 | 14.985 | 84.021 |
| log(1+yearsinR) | 31.255 | 13.925 | 111.349 |
| retired*female | 142.484 | 132.378 | 548.763 |
| fromWtoR*female | 74.821 | 149.695 | 274.774 |
| log(1+ yearsinR)*female | 123.528 | 137.785 | 499.027 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

**Table 10. Cognitive decline - TSLS estimates - Robustness analysis - Age: 50-80**

| | (1) | (2) | (3) |
|---|---|---|---|
| | TSLS | TSLS - ER | TSLS - SR |
| retired | -0.101** | -0.223*** | 0.021 |
| | (0.044) | (0.073) | (0.064) |
| fromWtoR | -0.022 | -0.358* | 0.119 |
| | (0.057) | (0.208) | (0.073) |
| log (1+yearsinR) | 0.050 | -0.148 | 0.131*** |
| | (0.031) | (0.132) | (0.043) |
| log age | 0.590** | 2.071** | -0.231 |
| | (0.263) | (0.939) | (0.397) |
| low cognition | -0.175*** | -0.168*** | -0.178*** |
| | (0.010) | (0.011) | (0.010) |
| female | -0.028*** | -0.023** | -0.030*** |
| | (0.009) | (0.011) | (0.010) |
| wave 2 | -0.008 | -0.046 | 0.006 |
| | (0.022) | (0.033) | (0.022) |
| less repetitions | 0.001 | 0.012 | -0.001 |
| | (0.019) | (0.021) | (0.019) |
| eqincomeQ2 | -0.027* | -0.016 | -0.031* |
| | (0.016) | (0.018) | (0.016) |
| eqincomeQ3 | -0.036** | -0.019 | -0.040** |
| | (0.015) | (0.019) | (0.016) |
| eqincomeQ4 | -0.033** | -0.028* | -0.028* |
| | (0.015) | (0.016) | (0.016) |
| highschool | -0.036*** | -0.036*** | -0.035*** |
| | (0.012) | (0.012) | (0.012) |
| college | -0.076*** | -0.094*** | -0.065*** |
| | (0.013) | (0.017) | (0.014) |
| poorhealth_bl | 0.003 | 0.022 | -0.011 |
| | (0.035) | (0.039) | (0.036) |
| drophealth | 0.073*** | 0.091*** | 0.065** |
| | (0.027) | (0.029) | (0.028) |
| increasehealth | -0.020 | -0.005 | -0.027 |
| | (0.035) | (0.038) | (0.036) |
| physicalact_bl | -0.033*** | -0.044*** | -0.026** |
| | (0.012) | (0.015) | (0.012) |
| increasephysicalact | -0.030** | -0.032** | -0.029* |
| | (0.015) | (0.016) | (0.016) |
| dropphysicalact | 0.045*** | 0.054*** | 0.041*** |
| | (0.013) | (0.015) | (0.013) |
| dailyact_bl | -0.086*** | -0.090*** | -0.089*** |
| | (0.019) | (0.020) | (0.019) |
| increasedailyact | -0.070*** | -0.067*** | -0.071*** |
| | (0.013) | (0.014) | (0.013) |
| dropdailyact | 0.077** | 0.066* | 0.083** |
| | (0.036) | (0.038) | (0.036) |
| SE | -0.011 | -0.058 | 0.016 |
| | (0.021) | (0.038) | (0.024) |
| DK | 0.011 | -0.053 | 0.043* |
| | (0.022) | (0.046) | (0.025) |
| NL | 0.009 | -0.022 | 0.025 |
| | (0.023) | (0.031) | (0.024) |
| BE | -0.077*** | -0.058** | -0.087*** |
| | (0.020) | (0.025) | (0.020) |
| FR | -0.090*** | -0.049 | -0.109*** |
| | (0.019) | (0.034) | (0.021) |
| CH | -0.041 | -0.109** | -0.003 |
| | (0.027) | (0.050) | (0.030) |
| AT | 0.057* | 0.148** | 0.009 |
| | (0.034) | (0.066) | (0.038) |
| ES | -0.035 | -0.051* | -0.024 |
| | (0.026) | (0.029) | (0.027) |
| IT | -0.086*** | -0.028 | -0.117*** |
| | (0.023) | (0.044) | (0.027) |
| constant | -1.921* | -7.713** | 1.283 |
| | (1.029) | (3.678) | (1.556) |
| Observations | 7,982 | 7,982 | 7,982 |
| R-squared | 0.086 | 0.005 | 0.067 |
| Adj R-squared | 0.082 | 0.002 | 0.064 |

| | | | |
|---|---|---|---|
| Sargan-Hansen (p-value) | 0.018 | | |
| F-test joint significance (p-value) | | | |
| retirement variables | 0.000 | 0.000 | 0.000 |
| | | | |
| Angrist and Pischke, first-stage F stat | F(4,601) | F(1,601) | F(1,601) |
| (Weak identification test) | | | |
| retired | 103.583 | 19.985 | 91.546 |
| fromWtoR | 46.437 | 12.091 | 76.134 |
| log(1+yearsinR) | 46.817 | 11.588 | 88.890 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

**Table 11. Cognitive decline - TSLS estimates - Robustness analysis - Alternative definition of retirement**

| | (1) | (2) | (3) |
|---|---|---|---|
| | TSLS | TSLS - ER | TSLS - SR |
| retired | -0.0841* | -0.234*** | 0.0192 |
| | (0.0434) | (0.0849) | (0.0580) |
| fromWtoR | 0.00104 | -0.317 | 0.105* |
| | (0.0535) | (0.199) | (0.0632) |
| log (1+yearsinR) | 0.0670** | -0.114 | 0.125*** |
| | (0.0293) | (0.125) | (0.0358) |
| log age | 0.434* | 1.810** | -0.148 |
| | (0.233) | (0.890) | (0.314) |
| low cognition | -0.175*** | -0.167*** | -0.178*** |
| | (0.00969) | (0.0117) | (0.00991) |
| female | -0.0305*** | -0.0220* | -0.0332*** |
| | (0.00924) | (0.0114) | (0.00946) |
| wave 2 | -0.00626 | -0.0424 | 0.00421 |
| | (0.0213) | (0.0321) | (0.0219) |
| less repetitions | -0.000223 | 0.00895 | -0.00161 |
| | (0.0188) | (0.0204) | (0.0192) |
| eqincomeQ2 | -0.0327** | -0.0223 | -0.0360** |
| | (0.0159) | (0.0174) | (0.0161) |
| eqincomeQ3 | -0.0408*** | -0.0284* | -0.0423*** |
| | (0.0151) | (0.0172) | (0.0157) |
| eqincomeQ4 | -0.0365** | -0.0388** | -0.0303* |
| | (0.0149) | (0.0154) | (0.0160) |
| highschool | -0.0360*** | -0.0342*** | -0.0362*** |
| | (0.0114) | (0.0125) | (0.0114) |
| college | -0.0734*** | -0.0914*** | -0.0650*** |
| | (0.0132) | (0.0169) | (0.0134) |
| poorhealth_bl | -0.00362 | 0.0148 | -0.0143 |
| | (0.0341) | (0.0381) | (0.0346) |
| drophealth | 0.0608** | 0.0834*** | 0.0531* |
| | (0.0264) | (0.0295) | (0.0271) |
| increasehealth | -0.00940 | 0.00306 | -0.0137 |
| | (0.0349) | (0.0378) | (0.0352) |
| physicalact_bl | -0.0336*** | -0.0483*** | -0.0272** |
| | (0.0121) | (0.0160) | (0.0123) |
| increasephysicalact | -0.0324** | -0.0334** | -0.0316** |
| | (0.0152) | (0.0157) | (0.0155) |
| dropphysicalact | 0.0447*** | 0.0524*** | 0.0416*** |
| | (0.0126) | (0.0144) | (0.0127) |
| dailyact_bl | -0.0884*** | -0.0961*** | -0.0889*** |
| | (0.0187) | (0.0209) | (0.0187) |
| increasedailyact | -0.0740*** | -0.0731*** | -0.0738*** |
| | (0.0127) | (0.0134) | (0.0129) |
| dropdailyact | 0.0708** | 0.0606 | 0.0763** |
| | (0.0356) | (0.0385) | (0.0356) |
| SE | -0.00711 | -0.0566 | 0.0146 |
| | (0.0212) | (0.0391) | (0.0226) |
| DK | 0.0118 | -0.0422 | 0.0334 |
| | (0.0212) | (0.0408) | (0.0228) |
| NL | 0.00473 | -0.0184 | 0.0139 |
| | (0.0222) | (0.0277) | (0.0227) |
| BE | -0.0770*** | -0.0494* | -0.0881*** |
| | (0.0197) | (0.0278) | (0.0205) |
| FR | -0.0965*** | -0.0493 | -0.114*** |
| | (0.0195) | (0.0378) | (0.0204) |
| CH | -0.0367 | -0.111** | -0.00433 |
| | (0.0269) | (0.0531) | (0.0294) |
| AT | 0.0465 | 0.133** | 0.00998 |
| | (0.0333) | (0.0654) | (0.0359) |
| ES | -0.0419 | -0.0453* | -0.0401 |
| | (0.0255) | (0.0268) | (0.0261) |
| IT | -0.0936*** | -0.0371 | -0.118*** |
| | (0.0233) | (0.0443) | (0.0257) |
| constant | -1.298 | -6.670* | 0.973 |
| | (0.913) | (3.482) | (1.228) |
| Observations | 8,221 | 8,221 | 8,221 |
| R-squared | 0.089 | -0.000 | 0.072 |

| | | | |
|---|---|---|---|
| Adj R-squared | 0.085 | -0.004 | 0.069 |
| | | | |
| Sargan-Hansen (p-value) | 0.019 | | |
| F-test joint significance (p-value) | | | |
| retirement variables | 0.000 | 0.000 | 0.000 |
| | | | |
| Angrist and Pischke, first-stage F stat | F(4,687) | F(1,687) | F(1,687) |
| (Weak identification test) | | | |
| retired | 96.46 | 16.23 | 126.1 |
| fromWtoR | 50.98 | 12.11 | 104.2 |
| log(1+yearsinR) | 51.59 | 11.51 | 125.5 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table 12. Memory score and attrition

| Memory score | (1)<br>Firstwave: 1 - dead not included | (2)<br>Firstwave: 2 - dead not included |
|---|---|---|
| observedW4 | 0.420*** | 0.130 |
| | (0.052) | (0.088) |
| SE | -0.001 | -0.337* |
| | (0.097) | (0.181) |
| DK | 0.161 | 0.341** |
| | (0.116) | (0.158) |
| NL | -0.312*** | -0.092 |
| | (0.108) | (0.176) |
| BE | -1.047*** | -0.815*** |
| | (0.100) | (0.254) |
| FR | -1.707*** | -1.525*** |
| | (0.101) | (0.173) |
| CH | 0.886*** | -0.461** |
| | (0.167) | (0.180) |
| AT | -0.573*** | 0.368 |
| | (0.113) | (0.385) |
| ES | -3.238*** | -3.314*** |
| | (0.121) | (0.178) |
| IT | -2.554*** | -1.528*** |
| | (0.108) | (0.160) |
| Constant | 9.047*** | 9.465*** |
| | (0.073) | (0.125) |
| | | |
| Observations | 17,313 | 6,891 |
| R-squared | 0.097 | 0.087 |

Notes: Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table 13. Two-sample Kolmogorov-Smirnov test for equality of distribution functions

| | Wave 1 – Wave 4 | | | Wave 2 – Wave 4 | | |
|---|---|---|---|---|---|---|
| | D | P-value | Corrected | D | P-value | Corrected |
| Leavers | 0.0464 | 0.000 | | 0.0261 | 0.109 | |
| Stayers | -0.0010 | 0.989 | | -0.0127 | 0.593 | |
| Combined K-S | 0.0464 | 0.000 | 0.000 | 0.0261 | 0.217 | 0.207 |

## Table 14. Selection and Cognitive decline TSLS estimates

| | (1) | (2) | (3) |
|---|---|---|---|
| Dep. Var. | Participation | TSLS | TSLS - drop between wave 1 and 4 - reduced set of covariates |
| retired | | -0.077 | -0.081* |
| | | (0.051) | (0.047) |
| fromWtoR | | 0.016 | 0.013 |
| | | (0.065) | (0.063) |
| log (1+yearsinR) | | 0.084** | 0.083** |
| | | (0.036) | (0.036) |
| log age | -1.109*** | 0.348 | 0.361 |
| | (0.368) | (0.293) | (0.290) |
| low cognition | -0.036* | -0.169*** | -0.169*** |
| | (0.021) | (0.012) | (0.011) |
| Female | 0.001 | -0.041*** | -0.041*** |
| | (0.022) | (0.011) | (0.011) |
| eqincomeQ2 | 0.019 | -0.051*** | -0.051*** |
| | (0.027) | (0.018) | (0.018) |
| eqincomeQ3 | -0.004 | -0.056*** | -0.055*** |
| | (0.030) | (0.017) | (0.017) |
| eqincomeQ4 | -0.042 | -0.051*** | -0.051*** |
| | (0.030) | (0.017) | (0.017) |
| highschool | 0.040 | -0.044*** | -0.045*** |
| | (0.026) | (0.014) | (0.014) |
| college | 0.194*** | -0.093*** | -0.095*** |
| | (0.027) | (0.016) | (0.014) |
| poorhealth_bl | -0.089* | -0.025 | -0.025 |
| | (0.048) | (0.034) | (0.033) |
| physicalact_bl | 0.093*** | -0.008 | -0.009 |
| | (0.020) | (0.012) | (0.012) |
| dailyact_bl | 0.102*** | -0.018 | -0.018 |
| | (0.029) | (0.015) | (0.014) |
| SE | 0.419*** | 0.004 | -0.000 |
| | (0.046) | (0.034) | (0.024) |
| DK | 0.662*** | 0.034 | 0.028 |
| | (0.052) | (0.041) | (0.027) |
| NL | 0.370*** | 0.003 | -0.000 |
| | (0.045) | (0.032) | (0.025) |
| BE | 0.565*** | -0.067** | -0.071*** |
| | (0.046) | (0.034) | (0.022) |
| FR | 0.526*** | -0.063* | -0.068*** |
| | (0.053) | (0.034) | (0.023) |
| CH | 0.707*** | 0.016 | 0.010 |
| | (0.070) | (0.048) | (0.034) |
| AT | 0.104** | 0.048 | 0.048 |
| | (0.051) | (0.035) | (0.035) |
| ES | 0.521*** | -0.011 | -0.015 |
| | (0.063) | (0.039) | (0.030) |
| IT | 0.528*** | -0.051 | -0.056* |
| | (0.055) | (0.038) | (0.029) |
| invmills | | 0.012 | |
| | | (0.073) | |
| eligibleER | 0.221** | | |
| | (0.096) | | |
| eligibleSR | 0.188** | | |
| | (0.077) | | |
| log (1+sinceER) | 0.016 | | |
| | (0.044) | | |
| log (1+sinceSR) | 0.030 | | |
| | (0.035) | | |
| fromNEtoE_ER | 0.169** | | |
| | (0.075) | | |
| fromNEtoE_SR | 0.119** | | |
| | (0.057) | | |
| length_iv_module | -0.003*** | | |
| | (0.001) | | |
| length_iv_module$^2$/1000 | 0.008*** | | |
| | (0.002) | | |
| length_iv_missing | -0.587*** | | |
| | (0.052) | | |
| constant | 4.096*** | -1.033 | -1.071 |
| | (1.455) | (1.133) | (1.131) |

| | | | |
|---|---|---|---|
| Observations | 17,313 | 6,125 | 6,125 |
| | | | |
| R-squared | | 0.081 | 0.081 |
| Pseudo R-squared | 0.036 | | |
| Adj R-squared | | 0.077 | 0.078 |
| | | | |
| Sargan-Hansen (p-value) | | 0.034 | 0.004 |
| | | | |
| Angrist and Pischke, first-stage | | F(7, 660) | F(4, 660) |
| F stat (Weak identification test) | | | |
| retired | | 49.869 | 77.923 |
| fromWtoR | | 25.948 | 41.017 |
| log(1+yearsinR) | | 24.208 | 38.217 |
| | | | |
| F-test joint significance | | | |
| (p-value) | 0.000 | | |
| length variables | | | |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table 15. Selection and Cognitive decline TSLS estimates – Early retirement

| | (1) | (2) | (3) |
|---|---|---|---|
| Dep. Var. | Participation | TSLS | TSLS - drop between wave 1 and 4 - reduced set of covariates |
| retired | | -0.211** | -0.261** |
| | | (0.096) | (0.125) |
| fromWtoR | | -0.240 | -0.351 |
| | | (0.222) | (0.299) |
| log (1+yearsinR) | | -0.055 | -0.124 |
| | | (0.145) | (0.194) |
| log age | -0.859*** | 1.456 | 1.973 |
| | (0.270) | (1.046) | (1.413) |
| low cognition | -0.036* | -0.164*** | -0.161*** |
| | (0.021) | (0.013) | (0.015) |
| female | 0.007 | -0.036*** | -0.034** |
| | (0.021) | (0.012) | (0.014) |
| eqincomeQ2 | 0.020 | -0.049*** | -0.049*** |
| | (0.027) | (0.018) | (0.018) |
| eqincomeQ3 | -0.003 | -0.049*** | -0.046** |
| | (0.030) | (0.018) | (0.019) |
| eqincomeQ4 | -0.043 | -0.057*** | -0.059*** |
| | (0.030) | (0.019) | (0.019) |
| highschool | 0.038 | -0.042*** | -0.042*** |
| | (0.026) | (0.014) | (0.015) |
| college | 0.193*** | -0.103*** | -0.111*** |
| | (0.027) | (0.018) | (0.020) |
| poorhealth_bl | -0.090* | -0.012 | -0.004 |
| | (0.048) | (0.038) | (0.041) |
| physicalact_bl | 0.094*** | -0.012 | -0.016 |
| | (0.020) | (0.013) | (0.013) |
| dailyact_bl | 0.102*** | -0.020 | -0.024 |
| | (0.029) | (0.016) | (0.016) |
| SE | 0.415*** | -0.029 | -0.053 |
| | (0.046) | (0.043) | (0.052) |
| DK | 0.647*** | -0.009 | -0.042 |
| | (0.051) | (0.054) | (0.065) |
| NL | 0.366*** | -0.018 | -0.036 |
| | (0.045) | (0.036) | (0.039) |
| BE | 0.560*** | -0.047 | -0.050* |
| | (0.044) | (0.042) | (0.030) |
| FR | 0.534*** | -0.028 | -0.023 |
| | (0.048) | (0.052) | (0.048) |
| CH | 0.726*** | -0.030 | -0.066 |
| | (0.070) | (0.062) | (0.073) |
| AT | 0.111** | 0.118 | 0.146 |
| | (0.050) | (0.075) | (0.093) |
| ES | 0.520*** | -0.022 | -0.038 |
| | (0.063) | (0.040) | (0.034) |
| IT | 0.536*** | 0.006 | 0.019 |
| | (0.052) | (0.068) | (0.071) |
| invmills | | 0.026 | |
| | | (0.077) | |
| eligibleER | 0.292*** | | |
| | (0.053) | | |
| log (1+sinceER) | 0.051 | | |
| | (0.033) | | |
| fromNEtoE_ER | 0.201*** | | |
| | (0.064) | | |
| length_iv_module | -0.003*** | | |
| | (0.001) | | |
| length_iv_module$^2$/1000 | 0.008*** | | |
| | (0.002) | | |
| length_iv_missing | -0.586*** | | |
| | (0.052) | | |
| constant | 3.106*** | -5.359 | -7.348 |
| | (1.065) | (4.103) | (5.518) |
| Observations | 17,313 | 6,125 | 6,125 |
| R-squared | | 0.058 | 0.017 |
| Pseudo R-squared | 0.035 | | |
| Adj R-squared | | 0.054 | 0.013 |

| | | | |
|---|---|---|---|
| Angrist and Pischke, first-stage F stat (Weak identification test) | | F(4, 660) | F(1, 660) |
| retired | | 4.156 | 6.718 |
| fromWtoR | | 3.364 | 5.883 |
| log(1+yearsinR) | | 3.196 | 5.574 |
| F-test joint significance (p-value) length variables | 0.000 | | |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

### Table 16. Selection and Cognitive decline TSLS estimates – Statutory retirement

| | (1) | (2) | (3) |
|---|---|---|---|
| Dep. Var. | Participation | TSLS | TSLS - drop between wave 1 and 4 - reduced set of covariates |
| retired | | 0.036 | 0.045 |
| | | (0.067) | (0.063) |
| fromWtoR | | 0.146* | 0.150** |
| | | (0.075) | (0.075) |
| log (1+yearsinR) | | 0.161*** | 0.163*** |
| | | (0.045) | (0.046) |
| log age | -0.555** | -0.383 | -0.412 |
| | (0.266) | (0.389) | (0.394) |
| low cognition | -0.036* | -0.172*** | -0.172*** |
| | (0.021) | (0.012) | (0.012) |
| female | 0.006 | -0.043*** | -0.043*** |
| | (0.021) | (0.011) | (0.011) |
| eqincomeQ2 | 0.019 | -0.052*** | -0.051*** |
| | (0.027) | (0.018) | (0.018) |
| eqincomeQ3 | -0.005 | -0.055*** | -0.055*** |
| | (0.030) | (0.018) | (0.018) |
| eqincomeQ4 | -0.045 | -0.040** | -0.040** |
| | (0.030) | (0.019) | (0.019) |
| highschool | 0.042 | -0.045*** | -0.045*** |
| | (0.026) | (0.014) | (0.014) |
| college | 0.195*** | -0.089*** | -0.086*** |
| | (0.027) | (0.016) | (0.015) |
| poorhealth_bl | -0.090* | -0.036 | -0.038 |
| | (0.048) | (0.035) | (0.035) |
| physicalact_bl | 0.093*** | -0.006 | -0.005 |
| | (0.020) | (0.012) | (0.012) |
| dailyact_bl | 0.102*** | -0.021 | -0.020 |
| | (0.029) | (0.015) | (0.015) |
| SE | 0.432*** | 0.018 | 0.027 |
| | (0.045) | (0.036) | (0.027) |
| DK | 0.683*** | 0.048 | 0.059** |
| | (0.049) | (0.044) | (0.030) |
| NL | 0.371*** | 0.010 | 0.017 |
| | (0.044) | (0.033) | (0.027) |
| BE | 0.589*** | -0.089** | -0.080*** |
| | (0.045) | (0.035) | (0.023) |
| FR | 0.590*** | -0.093*** | -0.085*** |
| | (0.045) | (0.036) | (0.024) |
| CH | 0.698*** | 0.037 | 0.049 |
| | (0.069) | (0.051) | (0.037) |
| AT | 0.129*** | 0.000 | 0.001 |
| | (0.049) | (0.039) | (0.039) |
| ES | 0.537*** | -0.012 | -0.004 |
| | (0.063) | (0.040) | (0.031) |
| IT | 0.578*** | -0.098** | -0.090*** |
| | (0.051) | (0.041) | (0.032) |
| invmills | | -0.024 | |
| | | (0.074) | |
| eligibleSR | 0.288*** | | |
| | (0.051) | | |
| log (1+sinceSR) | -0.016 | | |
| | (0.030) | | |
| fromNEtoE_SR | 0.183*** | | |
| | (0.043) | | |
| length_iv_module | -0.003*** | | |
| | (0.001) | | |
| length_iv_module$^2$/1000 | 0.008*** | | |
| | (0.002) | | |
| length_iv_missing | -0.587*** | | |
| | (0.052) | | |
| constant | 1.953* | 1.851 | 1.939 |
| | (1.066) | (1.511) | (1.541) |
| Observations | 17,313 | 6,125 | 6,125 |
| R-squared | | 0.053 | 0.052 |
| Pseudo R-squared | 0.035 | | |
| Adj R-squared | | 0.050 | 0.048 |

| | | | |
|---|---|---|---|
| Angrist and Pischke, first-stage F stat (Weak identification test) | | F(4, 660) | F(1, 660) |
| retired | | 30.909 | 99.839 |
| fromWtoR | | 22.264 | 80.654 |
| log(1+yearsinR) | | 24.245 | 87.390 |
| F-test joint significance (p-value) length variables | 0.000 | | |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

# Figures

**Figure 1: Average Memory Score by age**

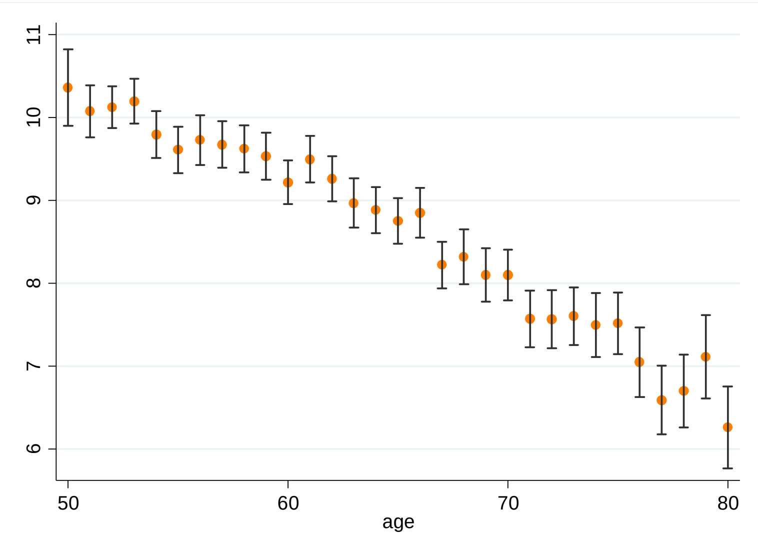
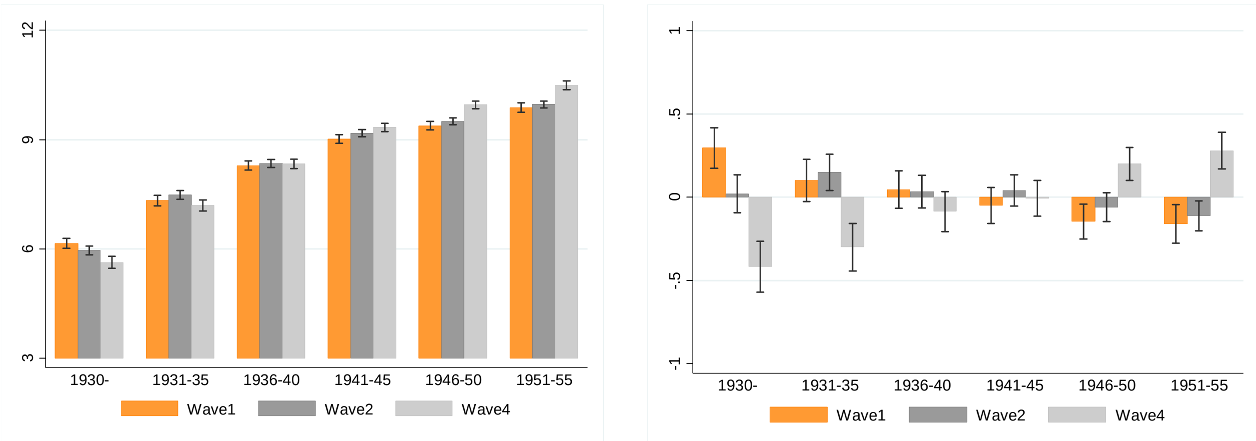



**Figure 2: Average Memory Score, by cohort and wave**




*Notes: On the left hand panel we show the average memory score by cohort and wave; on the right hand panel instead we show the mean of residuals by cohort and wave obtained by regressing the memory score on country and cohort dummies, time, gender, education and retirement status.*

**Figure 3: Percentage of High Decrease in Cognitive Abilities, by age in wave 4**

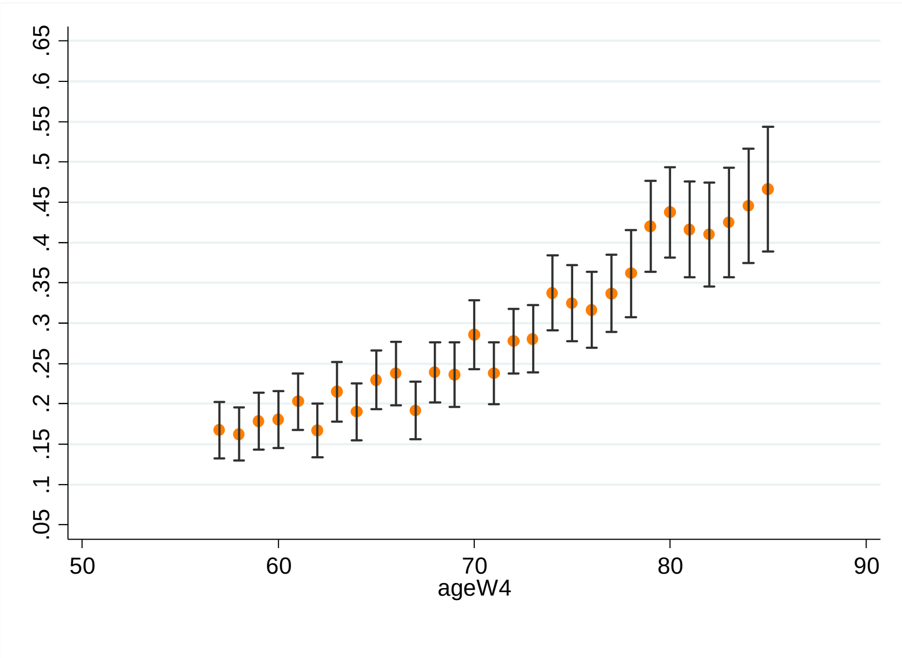
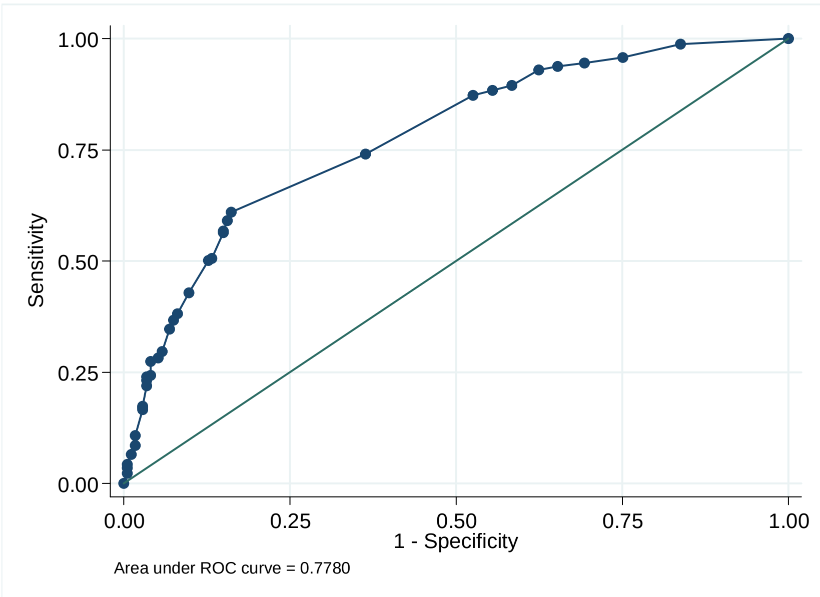



**Figure 4: Receiver Operating Characteristic (ROC) curve**



# Appendix A

Table 4, third and fourth column, shows estimates when controlling for a larger set of controls in addition to age, low cognition, gender, education, health, physical and other daily activities, and two dummies capturing individual who entered the sample in wave 2 and those having performed the test only twice. We describe below the additional controls.

To capture early-life conditions we define a zero-one dummy for the presence of less than 25 books at the parental home at age ten, and another dummy for living in rural areas at the age of ten (these play an important role in determining the returns to education according to Brunello et al. 2012). We include also a self-assessed measure of cognitive abilities when ten, exploiting a question asked about the relative position in mathematics: *mathskills* indicates that the individual declares to be better than the average of his/her schoolmates.[19]

As in Coe and Zamarro (2011), we consider the role of jobs' characteristics in shaping cognitive patterns. We include as controls *public*, *selfemployed* and *jobexperience40*. *Public* is a dummy variable that takes value 1 if the individual works or worked in the public sector (for retired individuals this regards the last job), similarly *selfemployed* indicates whether the individual works or worked as self-employed *Jobexperience40* captures individuals that worked for more than 40 years, and who entered the labour market very early, probably in low-skill positions.

Physical health conditions are captured by two variables: *nogstest*, and *eurodcat*. The first variable is a dummy that takes value 1 if the individual does not perform the hand-grip strength test, a situation that is usually considered a good predictor for future health problems among older adults, especially mortality and disability (Bohannon, 2008). *Eurodcat* is a dummy that equals 1 if the individual has at least one symptom of depression.[20]

We include also a measure of respondent cooperation, usually considered in survey participation analysis, *missingincome*, that takes value 1 if the respondent does not answer the question about household income.[21]

An engaged life style can be also maintained through social contacts that we proxy controlling for whether individuals have daily contacts with their children, *dailycontactchild*.

To account for changes in participation behaviour, we have *missingincomeW4*, that takes value 1 if the individual reports a valid income value in wave 1 or 2 but does not answer the question in wave 4. *NomissingincomeW4* instead captures the reverse situation.

In labelling changes between waves, we use the following notation: the suffix *_bl* denotes baseline observations, *drop* identifies cases in which the individual worsens his or her status compared to baseline, whereas *increase* cases in which he or she improves it. For hand-grip test for instance: *dropgstest* means that the respondent performed the test in wave 1 or 2 but did not perform it in wave 4, whereas *increasegstest* denotes an individual who did not perform the test in baseline (wave 1 or 2) but did it in wave 4.

We also consider some indicators that reflect the context in which the cognitive test was performed. For instance, the variable *dropalonecftest* identifies individuals who were alone with the interviewer when they

[19] The possible answers are Much better, Better, About the same, Worse, Much worse and Did not go to school; *mathskills* includes the first two options.

[20] Depression related questions in SHARE ask about depression, pessimism, suicidality, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment and tearfulness.

[21] "Don't know" are not considered regular values, even if there could be some information about income in terms of brackets.

took the memory test in wave 1 or 2, but who were not alone in wave 4 (i.e., there was someone else present during their test); *increasealonecftest* denotes the reverse situation. Also, the dummies *dropcontextcftest* and *increasecontextcftest* capture respectively situations in which there were no distractions in wave 1 or 2 but some in wave 4, and vice versa. There is an ample survey methodology literature pointing to the importance of taking such factors into consideration.

**Table A.1 Summary statistics**

| Variable | Description | % | Mean | SD |
|---|---|---|---|---|
| Memoryscore | Sum of words recalled in 1st and 2nd trial | | 9.140 | 3.228 |
| cognitive decline | 20% drop in memory score | 24.1 | | |
| Retired | being retired | 54.7 | | |
| fromWtoR | transition from work to retirement | 19.3 | | |
| log (1+yearsinR) | log of years spent in retirement | | 1.768 | 0.862 |
| eligibleER | being eligible for early retirement | 61.8 | | |
| fromNEtoE_ER | transition from being non eligible to being eligible for early retirement | 22.5 | | |
| log (1+sinceER) | log of years since eligibility for early retirement | | 2.068 | 0.875 |
| eligibleSR | being eligible for statutory retirement | 43.6 | | |
| fromNEtoE_SR | transition from being non eligible to being eligible for statutory retirement | 25.3 | | |
| log (1+sinceSR) | log of years since eligibility for statutory retirement | | 1.358 | 1.048 |
| retired (alt.) | being retired (alternative definition: additionally did not do any paid work in the four weeks before the baseline interview) | 51.4 | | |
| fromWtoR (alt.) | transition from work to retirement (alternative definition) | 19.3 | | |
| log (1+yearsinR) (alt.) | log of years spent in retirement (alternative definition) | | 1.692 | 0.932 |
| log age | log of age | | 4.131 | 0.131 |
| low cognition | baseline memory score is lower than the median value by wave and country | 50.8 | | |
| Female | being female | 42.1 | | |
| wave 2 | refreshment sample indicator | 25.5 | | |
| less repetitions | having performed the memory test twice (rather than three times) | 31.7 | | |
| jobexperience40 | worked for more than 40 years | 51.5 | | |
| eqincomeQ2 | 2nd Equivalent household income quartile | 24.3 | | |
| eqincomeQ3 | 3rd Equivalent household income quartile | 27.1 | | |
| eqincomeQ4 | 4th Equivalent household income quartile | 30.2 | | |
| incomemissing_bl | answered to household income question | 60.6 | | |
| missingincomeW4 | answered to household income question in baseline but not in wave 4 | 14.7 | | |
| nomissingincomeW4 | answered to household income question in wave 4 but not in baseline | 36.3 | | |
| Fewbooks | less than 25 books at home when 10 | 36.1 | | |
| math_skills | relative position in mathematics | 40.0 | | |
| Ruralarea | living in rural area when 10 | 41.9 | | |
| highschool | International Standard Classification of Education: 3-4 | 34.7 | | |
| college | International Standard Classification of | 28.1 | | |

| | | | | |
|---|---|---|---|---|
| | Education: 5-6 | | | |
| Public | works or worked in the public sector | 31.0 | | |
| Selfemployed | works or worked as self-employed | 12.3 | | |
| Partner | having a partner | 79.3 | | |
| nogstest_bl | no grip strenght test in baseline | 2.7 | | |
| Dropgstest | did the grip strength test in baseline not in wave 4 | 3.0 | | |
| Increasegstest | did the grip strength test in wave 4 not in baseline | 2.0 | | |
| alonecftest_bl | did the memory test alone in baseline | 84.7 | | |
| Dropalonecftest | did the memory test alone in baseline and was not alone in wave 4 | 8.7 | | |
| increasealonecftest | did the memory test alone in wave 4 but was not alone in baseline | 11.4 | | |
| contextcftest_bl | some contextual factors during the memory test in baseline | 7.7 | | |
| increasecontextcftest | some contextual factors during the memory test in baseline but not in wave 4 | 7.4 | | |
| dropcontextcftest | some contextual factors during the memory test in wave 4 but not in baseline | 2.9 | | |
| poorhealth_bl | being in poor health | 2.9 | | |
| Dropinhealth | health worsens between waves | 4.9 | | |
| Increaseinhealth | health improves between waves | 3.0 | | |
| eurodcat_bl | having at least one symptom of depression in baseline | 17.2 | | |
| increaseineurodcat | having at least one symptom of depression in baseline and no symptoms in wave 4 | 9.0 | | |
| Dropineurodcat | having at least one symptom of depression in wave 4 and no symptoms in baseline | 11.8 | | |
| physicalact_bl | practicing physical activity in baseline | 58.3 | | |
| increasephysicalact | started practicing physical activity in wave 4 | 13.6 | | |
| Dropphysicalact | stopped practicing physical activity in wave 4 | 20.5 | | |
| dailyact_bl | practicing any activity in baseline | 11.0 | | |
| Increasedailyact | started practicing any activity in wave 4 | 67.7 | | |
| Dropdailyact | stopped practicing activities in wave 4 | 1.9 | | |
| dailycontactchild_bl | having daily contacts with children | 42.5 | | |
| increase dailycontactchild | started having daily contacts with children in wave 4 | 9.5 | | |
| drop dailycontactchild | stopped having daily contacts with children in wave 4 | 16.8 | | |
| SE | Sweden | 12.5 | | |
| DK | Denmark | 13.9 | | |
| NL | The Netherlands | 9.5 | | |
| BE | Belgium | 13.5 | | |
| FR | France | 13.1 | | |
| CH | Switzerland | 6.7 | | |
| AT | Austria | 4.1 | | |
| ES | Spain | 6.6 | | |
| IT | Italy | 9.6 | | |
| DE | Germany | 10.5 | | |
| lenght_iv_m | Seconds | | 94.160 | 47.055 |
| length_iv_m_missing | No information about lenght of the IV module | 5.0 | | |

**Table A.2 Cognitive decline - TSLS estimates – (without any activities-related variable)**

| | (1)<br>Model 3 | (1)<br>Model 4 |
|---|---|---|
| retired | -0.104*** | -0.097*** |
| | (0.036) | (0.037) |
| fromWtoR | -0.027 | -0.018 |
| | (0.053) | (0.054) |
| log (1+yearsinR) | 0.057* | 0.057* |
| | (0.031) | (0.031) |
| log age | 0.598** | 0.538** |
| | (0.236) | (0.237) |
| low cognition | -0.168*** | -0.172*** |
| | (0.010) | (0.010) |
| Female | -0.029*** | -0.030*** |
| | (0.009) | (0.009) |
| wave 2 | -0.010 | -0.007 |
| | (0.012) | (0.021) |
| Highschool | -0.047*** | -0.042*** |
| | (0.011) | (0.011) |
| College | -0.097*** | -0.087*** |
| | (0.013) | (0.013) |
| less repetitions | | 0.001 |
| | | (0.019) |
| eqincomeQ2 | | -0.032** |
| | | (0.016) |
| eqincomeQ3 | | -0.044*** |
| | | (0.015) |
| eqincomeQ4 | | -0.042*** |
| | | (0.015) |
| poorhealth_bl | | 0.009 |
| | | (0.034) |
| Drophealth | | 0.082*** |
| | | (0.026) |
| Increasehealth | | -0.004 |
| | | (0.035) |
| SE | -0.020 | -0.016 |
| | (0.020) | (0.021) |
| DK | 0.006 | 0.010 |
| | (0.021) | (0.021) |
| NL | -0.008 | -0.001 |
| | (0.022) | (0.022) |
| BE | -0.074*** | -0.067*** |
| | (0.019) | (0.019) |
| FR | -0.080*** | -0.079*** |
| | (0.019) | (0.019) |
| CH | -0.048* | -0.041 |
| | (0.026) | (0.026) |
| AT | 0.056* | 0.058* |
| | (0.033) | (0.033) |
| ES | -0.009 | -0.004 |
| | (0.025) | (0.025) |
| IT | -0.060*** | -0.055** |
| | (0.023) | (0.023) |
| Constant | -2.049** | -1.785* |
| | (0.924) | (0.931) |
| Observations | 8,221 | 8,221 |
| R-squared | 0.079 | 0.082 |
| Adj R-squared | 0.077 | 0.080 |
| Sargan-Hansen (p-value) | 0.027 | 0.028 |
| F-test joint significance (p-value) retirement variables | 0.000 | 0.000 |
| Angrist and Pischke, first-stage F stat (Weak identification test) | | |
| Retired | 155.309 | 153.259 |
| fromWtoR | 51.880 | 51.628 |
| log(1+yearsinR) | 52.229 | 51.889 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table A.3 First stage estimates

| | retired | fromWtoR | log(1+yearsinR) |
|---|---|---|---|
| eligibleER | 0.275*** | 0.066 | 0.104** |
| | (0.035) | (0.047) | (0.050) |
| fromNEtoE_ER | 0.112*** | 0.125*** | 0.080** |
| | (0.016) | (0.033) | (0.031) |
| log (1+sinceER) | 0.212*** | -0.238*** | 0.449*** |
| | (0.026) | (0.029) | (0.053) |
| eligibleSR | 0.261*** | 0.057 | -0.422*** |
| | (0.035) | (0.043) | (0.063) |
| fromNEtoE_SR | 0.106*** | 0.165*** | -0.045 |
| | (0.023) | (0.035) | (0.028) |
| log (1+sinceSR) | -0.029 | -0.057*** | 0.471*** |
| | (0.018) | (0.018) | (0.033) |
| log age | -0.261* | 1.655*** | 0.868*** |
| | (0.158) | (0.191) | (0.316) |
| low cognition | 0.009 | -0.009 | 0.032*** |
| | (0.006) | (0.008) | (0.012) |
| female | -0.026*** | 0.024* | -0.032** |
| | (0.010) | (0.013) | (0.016) |
| wave 2 | 0.048*** | -0.092*** | 0.026 |
| | (0.017) | (0.021) | (0.028) |
| less repetitions | -0.027** | 0.027 | 0.011 |
| | (0.013) | (0.019) | (0.022) |
| eqincomeQ2 | -0.014 | 0.035*** | -0.001 |
| | (0.010) | (0.010) | (0.019) |
| eqincomeQ3 | -0.043*** | 0.070*** | -0.020 |
| | (0.011) | (0.012) | (0.021) |
| eqincomeQ4 | -0.074*** | 0.082*** | -0.081*** |
| | (0.012) | (0.013) | (0.021) |
| highschool | -0.013 | 0.004 | -0.009 |
| | (0.008) | (0.010) | (0.015) |
| college | -0.047*** | -0.004 | -0.053*** |
| | (0.010) | (0.012) | (0.018) |
| poorhealth_bl | 0.081*** | -0.054*** | 0.145*** |
| | (0.025) | (0.020) | (0.055) |
| drophealth | 0.014 | 0.012 | 0.056* |
| | (0.013) | (0.014) | (0.030) |
| increasehealth | 0.018 | -0.017 | 0.081 |
| | (0.024) | (0.023) | (0.051) |
| physicalact_bl | -0.033*** | 0.020* | -0.066*** |
| | (0.010) | (0.011) | (0.017) |
| increasephysicalact | -0.030*** | 0.039*** | -0.056** |
| | (0.012) | (0.013) | (0.022) |
| dropphysicalact | 0.008 | 0.013 | 0.013 |
| | (0.008) | (0.011) | (0.015) |
| dailyact_bl | 0.058*** | -0.046*** | 0.028 |
| | (0.013) | (0.016) | (0.024) |
| increasedailyact | -0.007 | 0.020** | -0.017 |
| | (0.008) | (0.009) | (0.017) |
| dropdailyact | -0.027 | 0.014 | -0.055 |
| | (0.024) | (0.029) | (0.043) |
| SE | -0.094*** | -0.003 | -0.206*** |
| | (0.019) | (0.026) | (0.034) |
| DK | -0.107*** | 0.009 | -0.324*** |
| | (0.018) | (0.025) | (0.035) |
| NL | -0.070*** | 0.017 | -0.221*** |
| | (0.019) | (0.027) | (0.030) |
| BE | -0.003 | 0.034 | -0.036 |
| | (0.021) | (0.027) | (0.030) |
| FR | -0.202*** | 0.153*** | -0.302*** |
| | (0.026) | (0.033) | (0.041) |
| CH | -0.075*** | -0.047* | -0.186*** |
| | (0.020) | (0.024) | (0.032) |
| AT | 0.047 | 0.042 | 0.179*** |
| | (0.034) | (0.041) | (0.052) |
| ES | -0.056** | 0.001 | -0.110*** |
| | (0.022) | (0.028) | (0.036) |
| IT | -0.086*** | 0.120*** | -0.090** |
| | (0.023) | (0.027) | (0.040) |
| constant | 1.202* | -6.552*** | -3.176** |
| | (0.630) | (0.758) | (1.254) |
| Observations | 8,221 | 8,221 | 8,221 |

| | | | |
|---|---|---|---|
| R-squared | 0.681 | 0.264 | 0.801 |
| F test of excluded instruments (p-value) | 0.000 | 0.000 | 0.000 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table A.4 First stage estimates - Early retirement

| Dep. Var. | retired | fromWtoR | log(1+yearsinR) |
|---|---|---|---|
| eligibleER | 0.391*** | 0.278*** | -0.541*** |
| | (0.023) | (0.028) | (0.049) |
| fromNEtoE_ER | 0.162*** | 0.184*** | -0.194*** |
| | (0.016) | (0.031) | (0.038) |
| log (1+sinceER) | 0.262*** | -0.336*** | 0.499*** |
| | (0.020) | (0.021) | (0.040) |
| log age | -0.376*** | 1.645*** | 3.872*** |
| | (0.143) | (0.173) | (0.326) |
| low cognition | 0.008 | -0.010 | 0.043*** |
| | (0.006) | (0.008) | (0.012) |
| female | -0.024** | 0.020 | 0.005 |
| | (0.010) | (0.013) | (0.021) |
| wave 2 | 0.043*** | -0.108*** | -0.002 |
| | (0.017) | (0.022) | (0.030) |
| less repetitions | -0.029** | 0.026 | 0.022 |
| | (0.013) | (0.019) | (0.025) |
| eqincomeQ2 | -0.012 | 0.036*** | -0.004 |
| | (0.010) | (0.011) | (0.020) |
| eqincomeQ3 | -0.041*** | 0.075*** | -0.022 |
| | (0.011) | (0.013) | (0.022) |
| eqincomeQ4 | -0.073*** | 0.085*** | -0.083*** |
| | (0.012) | (0.013) | (0.022) |
| highschool | -0.014* | 0.004 | -0.002 |
| | (0.008) | (0.010) | (0.016) |
| college | -0.047*** | -0.005 | -0.053*** |
| | (0.010) | (0.013) | (0.018) |
| poorhealth_bl | 0.084*** | -0.059*** | 0.136** |
| | (0.026) | (0.021) | (0.057) |
| drophealth | 0.012 | 0.011 | 0.068** |
| | (0.013) | (0.014) | (0.031) |
| increasehealth | 0.016 | -0.018 | 0.087* |
| | (0.024) | (0.024) | (0.052) |
| physicalact_bl | -0.030*** | 0.017 | -0.067*** |
| | (0.010) | (0.011) | (0.018) |
| increasephysicalact | -0.029** | 0.038*** | -0.055** |
| | (0.012) | (0.013) | (0.023) |
| dropphysicalact | 0.007 | 0.012 | 0.020 |
| | (0.008) | (0.011) | (0.015) |
| dailyact_bl | 0.060*** | -0.044*** | 0.021 |
| | (0.013) | (0.016) | (0.026) |
| increasedailyact | -0.005 | 0.025*** | -0.024 |
| | (0.008) | (0.009) | (0.017) |
| dropdailyact | -0.026 | 0.013 | -0.067 |
| | (0.025) | (0.028) | (0.047) |
| SE | -0.096*** | -0.008 | -0.183*** |
| | (0.021) | (0.028) | (0.045) |
| DK | -0.127*** | 0.007 | -0.279*** |
| | (0.018) | (0.023) | (0.041) |
| NL | -0.086*** | 0.002 | -0.137*** |
| | (0.019) | (0.025) | (0.036) |
| BE | -0.020 | 0.011 | 0.062* |
| | (0.021) | (0.024) | (0.036) |
| FR | -0.232*** | 0.142*** | 0.006 |
| | (0.024) | (0.030) | (0.044) |
| CH | -0.054*** | -0.056** | -0.202*** |
| | (0.020) | (0.023) | (0.035) |
| AT | 0.039 | 0.045 | 0.300*** |
| | (0.034) | (0.037) | (0.049) |
| ES | -0.061*** | -0.011 | -0.038 |
| | (0.023) | (0.028) | (0.043) |
| IT | -0.113*** | 0.127*** | 0.066 |
| | (0.021) | (0.025) | (0.042) |
| constant | 1.660*** | -6.500*** | -15.133*** |
| | (0.566) | (0.685) | (1.291) |
| Observations | 8,221 | 8,221 | 8,221 |
| R-squared | 0.674 | 0.245 | 0.780 |
| F test of excluded instruments (p-value) | 0.000 | 0.002 | 0.000 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

**Table A.5 First stage estimates - Statutory retirement**

| Dep. Var. | retired | fromWtoR | log(1+yearsinR) |
|---|---|---|---|
| eligibleSR | 0.574*** | -0.195*** | 0.018 |
| | (0.028) | (0.037) | (0.047) |
| fromNEtoE_SR | 0.204*** | 0.124*** | 0.031 |
| | (0.028) | (0.037) | (0.035) |
| log (1+sinceSR) | -0.060*** | -0.123*** | 0.519*** |
| | (0.017) | (0.019) | (0.027) |
| log age | 1.421*** | 0.942*** | 3.106*** |
| | (0.157) | (0.205) | (0.241) |
| low cognition | 0.007 | -0.004 | 0.028** |
| | (0.006) | (0.008) | (0.012) |
| Female | -0.024** | 0.029* | -0.035** |
| | (0.011) | (0.015) | (0.018) |
| wave 2 | 0.047*** | -0.109*** | 0.022 |
| | (0.018) | (0.022) | (0.028) |
| less repetitions | -0.028* | 0.031 | 0.011 |
| | (0.014) | (0.019) | (0.023) |
| eqincomeQ2 | -0.014 | 0.037*** | -0.002 |
| | (0.010) | (0.011) | (0.019) |
| eqincomeQ3 | -0.049*** | 0.077*** | -0.028 |
| | (0.011) | (0.012) | (0.021) |
| eqincomeQ4 | -0.087*** | 0.096*** | -0.101*** |
| | (0.013) | (0.013) | (0.021) |
| highschool | -0.010 | 0.005 | -0.007 |
| | (0.009) | (0.011) | (0.016) |
| College | -0.049*** | 0.000 | -0.057*** |
| | (0.010) | (0.013) | (0.018) |
| poorhealth_bl | 0.077*** | -0.048** | 0.135** |
| | (0.026) | (0.020) | (0.056) |
| drophealth | 0.010 | 0.012 | 0.050 |
| | (0.014) | (0.015) | (0.031) |
| increasehealth | 0.018 | -0.011 | 0.081 |
| | (0.024) | (0.023) | (0.051) |
| physicalact_bl | -0.032*** | 0.024** | -0.070*** |
| | (0.010) | (0.011) | (0.017) |
| increasephysicalact | -0.023* | 0.039*** | -0.052** |
| | (0.012) | (0.014) | (0.023) |
| dropphysicalact | 0.012 | 0.012 | 0.017 |
| | (0.008) | (0.012) | (0.015) |
| dailyact_bl | 0.062*** | -0.052*** | 0.033 |
| | (0.014) | (0.016) | (0.026) |
| increasedailyact | -0.007 | 0.018* | -0.015 |
| | (0.009) | (0.009) | (0.018) |
| dropdailyact | -0.034 | 0.020 | -0.060 |
| | (0.024) | (0.029) | (0.043) |
| SE | -0.096*** | 0.022 | -0.220*** |
| | (0.025) | (0.032) | (0.040) |
| DK | -0.065*** | -0.005 | -0.269*** |
| | (0.023) | (0.031) | (0.040) |
| NL | -0.033 | 0.021 | -0.185*** |
| | (0.024) | (0.035) | (0.037) |
| BE | 0.030 | 0.040 | -0.009 |
| | (0.023) | (0.034) | (0.038) |
| FR | -0.047* | 0.110*** | -0.110*** |
| | (0.025) | (0.032) | (0.037) |
| CH | -0.143*** | 0.020 | -0.302*** |
| | (0.027) | (0.032) | (0.040) |
| AT | 0.113*** | 0.022 | 0.269*** |
| | (0.034) | (0.044) | (0.059) |
| ES | -0.043* | 0.016 | -0.110*** |
| | (0.024) | (0.032) | (0.040) |
| IT | 0.057** | 0.043 | 0.114** |
| | (0.028) | (0.033) | (0.058) |
| Constant | -5.472*** | -3.661*** | -12.071*** |
| | (0.622) | (0.815) | (0.961) |
| Observations | 8,221 | 8,221 | 8,221 |
| R-squared | 0.651 | 0.207 | 0.785 |
| F test of excluded instruments (p-value) | 0.000 | 0.000 | 0.000 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table A.6. Cognitive decline - TSLS estimates

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Dropping: | | | | | | | | | |
| VARIABLES | All countries | AT | DE | SE | NL | ES | IT | FR | DK | CH | BE |
| retired | -0.089** | -0.094** | -0.087** | -0.114*** | -0.092** | -0.085** | -0.051 | -0.086** | -0.093** | -0.103*** | -0.068* |
| | (0.037) | (0.037) | (0.040) | (0.038) | (0.037) | (0.037) | (0.046) | (0.041) | (0.038) | (0.037) | (0.038) |
| fromWtoR | -0.012 | -0.014 | 0.003 | -0.045 | 0.005 | -0.003 | 0.036 | -0.022 | -0.026 | -0.043 | 0.004 |
| | (0.055) | (0.055) | (0.053) | (0.062) | (0.057) | (0.056) | (0.067) | (0.058) | (0.058) | (0.057) | (0.057) |
| log(1+yearsinR) | 0.056* | 0.061* | 0.069** | 0.046 | 0.067** | 0.068** | 0.078** | 0.049 | 0.037 | 0.040 | 0.050 |
| | (0.031) | (0.031) | (0.031) | (0.035) | (0.031) | (0.032) | (0.038) | (0.032) | (0.032) | (0.031) | (0.032) |
| log age | 0.510** | 0.506** | 0.410* | 0.608** | 0.470** | 0.419* | 0.290 | 0.544** | 0.650*** | 0.627*** | 0.487** |
| | (0.238) | (0.241) | (0.249) | (0.256) | (0.236) | (0.244) | (0.309) | (0.254) | (0.243) | (0.242) | (0.246) |
| Observations | 8,221 | 7,885 | 7,368 | 7,196 | 7,436 | 7,676 | 7,429 | 7,140 | 7,080 | 7,671 | 7,108 |
| Sargan-Hansen (p-value) | 0.019 | 0.019 | 0.013 | 0.103 | 0.030 | 0.021 | 0.017 | 0.035 | 0.076 | 0.047 | 0.008 |
| Angrist and Pischke, first-stage F stat (Weak identification test) | | | | | | | | | | | |
| retired | 153.831 | 138.345 | 103.196 | 167.349 | 181.572 | 152.659 | 81.266 | 133.381 | 192.630 | 139.948 | 176.567 |
| fromWtoR | 51.484 | 50.065 | 46.619 | 40.518 | 47.978 | 48.516 | 42.998 | 47.034 | 48.224 | 45.321 | 49.021 |
| log(1+yearsinR) | 51.982 | 49.094 | 44.042 | 42.022 | 50.484 | 48.178 | 43.567 | 47.684 | 53.191 | 46.424 | 46.548 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table A.7. Cognitive decline - TSLS estimates – Early retirement

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Dropping: | | | | | | | | | |
| VARIABLES | All countries | AT | DE | SE | NL | ES | IT | FR | DK | CH | BE |
| retired | -0.182*** | -0.204*** | -0.283** | -0.161*** | -0.167*** | -0.166*** | -0.340* | -0.162*** | -0.148*** | -0.187*** | -0.171*** |
| | (0.056) | (0.064) | (0.119) | (0.045) | (0.050) | (0.052) | (0.199) | (0.052) | (0.047) | (0.057) | (0.062) |
| fromWtoR | -0.310* | -0.348* | -0.588 | -0.213 | -0.258 | -0.248 | -0.668 | -0.217* | -0.240 | -0.327* | -0.380 |
| | (0.186) | (0.200) | (0.371) | (0.148) | (0.190) | (0.171) | (0.531) | (0.129) | (0.161) | (0.192) | (0.246) |
| log(1+yearsinR) | -0.122 | -0.144 | -0.313 | -0.052 | -0.085 | -0.072 | -0.347 | -0.064 | -0.089 | -0.126 | -0.181 |
| | (0.119) | (0.129) | (0.250) | (0.096) | (0.120) | (0.109) | (0.336) | (0.082) | (0.104) | (0.123) | (0.158) |
| log age | 1.781** | 1.989** | 3.137* | 1.295** | 1.542* | 1.439* | 3.503 | 1.388** | 1.518** | 1.806** | 2.099** |
| | (0.814) | (0.903) | (1.750) | (0.633) | (0.789) | (0.738) | (2.473) | (0.574) | (0.686) | (0.837) | (1.055) |
| Observations | 8,221 | 7,885 | 7,368 | 7,196 | 7,436 | 7,676 | 7,429 | 7,140 | 7,080 | 7,671 | 7,108 |
| Angrist and Pischke, first-stage F stat (Weak identification test) | | | | | | | | | | | |
| retired | 27.292 | 20.106 | 6.102 | 81.477 | 35.687 | 31.698 | 4.140 | 64.335 | 63.413 | 25.282 | 19.589 |
| fromWtoR | 14.190 | 12.606 | 5.326 | 20.170 | 13.041 | 15.083 | 3.512 | 26.760 | 16.443 | 12.955 | 9.234 |
| log(1+yearsinR) | 13.426 | 11.991 | 5.300 | 18.930 | 12.189 | 14.164 | 3.373 | 24.823 | 15.029 | 12.234 | 8.801 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Table A.8. Cognitive decline - TSLS estimates - Statutory retirement

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Dropping: | | | | | | | | | |
| VARIABLES | All countries | AT | DE | SE | NL | ES | IT | FR | DK | CH | BE |
| retired | 0.005 | 0.002 | 0.002 | -0.022 | 0.001 | 0.015 | 0.026 | 0.032 | -0.019 | -0.019 | 0.040 |
| | (0.051) | (0.052) | (0.053) | (0.055) | (0.051) | (0.053) | (0.054) | (0.061) | (0.051) | (0.052) | (0.053) |
| fromWtoR | 0.116* | 0.108 | 0.122* | 0.097 | 0.121* | 0.128* | 0.147** | 0.132* | 0.088 | 0.083 | 0.129* |
| | (0.069) | (0.069) | (0.065) | (0.085) | (0.071) | (0.071) | (0.074) | (0.080) | (0.074) | (0.073) | (0.068) |
| log(1+yearsinR) | 0.131*** | 0.135*** | 0.135*** | 0.129*** | 0.140*** | 0.146*** | 0.138*** | 0.141*** | 0.100** | 0.112*** | 0.131*** |
| | (0.040) | (0.041) | (0.038) | (0.049) | (0.041) | (0.042) | (0.043) | (0.047) | (0.042) | (0.043) | (0.042) |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| log age | -0.175 | -0.181 | -0.196 | -0.116 | -0.199 | -0.298 | -0.253 | -0.301 | 0.090 | -0.001 | -0.285 |
| | (0.340) | (0.347) | (0.335) | (0.391) | (0.337) | (0.356) | (0.363) | (0.405) | (0.347) | (0.358) | (0.352) |
| Observations | 8,221 | 7,885 | 7,368 | 7,196 | 7,436 | 7,676 | 7,429 | 7,140 | 7,080 | 7,671 | 7,108 |
| Angrist and Pischke, first-stage F stat (Weak identification test) | | | | | | | | | | | |
| retired | 127.127 | 117.465 | 116.450 | 94.124 | 141.356 | 117.012 | 106.781 | 90.613 | 131.615 | 105.373 | 140.644 |
| fromWtoR | 87.476 | 84.140 | 89.300 | 60.858 | 88.264 | 81.333 | 77.249 | 71.052 | 78.107 | 75.218 | 87.582 |
| log(1+yearsinR) | 104.632 | 100.340 | 106.220 | 71.256 | 109.730 | 96.505 | 91.976 | 79.300 | 98.934 | 88.528 | 109.565 |

Notes: Standard errors are robust to clustering at the country, gender and cohort level. The Angrist-Pischke F statistics are computed as suggested by Sanderson and Windmejier (2015). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

## Appendix B

The initial sources of information about early and normal retirement eligibility criteria are Gruber and Wise (1999, 2010) and Wise (2012). Other country specific auxiliary data sources are reported below. Figure B5 and B6 show the histograms of the retirement age by country, for males and females respectively. The vertical lines indicate the eligibility ages for early (blue) and statutory (red) retirement.

Austria (see Staubli and Zweimuller, 2011)
ER: 60 for men and 55 for women until 2001. From 2001 until 2004, early retirement depends on year of birth. For men it is 61 until 1942 and 62 from 1943 onwards. For women it is 56 for those born in 1947, 57 for those born from 1948 to 1951, 58 for those born from 1952. From 2005, it is 62 for men and women.
SR: 65 for men and 60 for women.

Belgium (see Jousten et al., 2010)
ER: No early retirement until 1966, 60 afterwards for men, for women 55 until 1986 and 60 from 1987.
SR: 65 for men, for women 60 until 1996, 61 from 1997 to 1999, 62 from 2000 to 2002, 63 from 2003 to 2005, 64 from 2006 to 2008, 65 from 2009.

Denmark (see Bingley et al., 2010)
ER: 60 for both men and women throughout the years, except from 1992 to 1993, when the ER was lowered to 55, and from 1994 to 1995, when it was 50.
SR: 67 until 2003, 65 from 2004, for both men and women.

France (see Hamblin, 2013)
ER: No early retirement until 1963. 60 from 1963 to 1980, 55 from 1981 onwards.
SR: 65 until 1982 and 60 from 1983 to 2010, from 2011 60 for those born till 1952, 61 for those born between 1953 and 1954 and 62 for those born since 1955.

Germany (see Berkel and Boersch-Supan, 2004, and Mazzonna and Peracchi., 2014)
ER: For men, no early retirement until 1972, 60 from 1973 until 2003, 63 from 2004 onwards. For women, no early retirement in 1961, 60 from 1962 until 2003, 62 from 2004 until 2005, 63 from 2006.
SR: 65 for all.

Italy
We follow Angelini et al., 2009, and refer to the details therein.

Netherlands (see Euwals et al., 2010)
ER: No early retirement until 1974. 60 from 1975 onwards, for both men and women.
SR: 65 for both men and women.

Spain (see Blanco, 2000)
ER: 64 until 1982, 60 from 1983 to 1993, 61 from 1994 onwards, for both men and women.
SR: 65 for both men and women.

Sweden (see Mazzonna and Peracchi, 2014)
ER: No early retirement until 1962, 60 from 1963 to 1997, 61 from 1998 onwards.
SR: 67 for both men and women until 1994, 65 from 1995 onward.

Switzerland (see Dorn and Sousa-Poza, 2003)
ER: No early retirement until 1996 for men and until 2000 for women. Then, 64 for men from 1997 until 2000 and 63 from 2001, for women 62 from 2001.
SR: 65 for men, for women 63 until 1963, 62 from 1964 until 2000, 63 from 2001 to 2004, 64 from 2005.

## Additional references for retirement ages


Angelini, V., Brugiavini, A., & Weber, G. (2009). Ageing and unused capacity in Europe: is there an early retirement trap?. Economic Policy, 24(59), 463-508.

Berkel, B. & Boersch-Supan, A. (2004). Pension Reforms in Germany: The Impact on Retirement Decisions. MEA Discussion Paper 62-2004.

Bingley, P., Datta Gupta, N. and Pedersen, P.J. (2010). Social Security, Retirement and Employment of the Young in Denmark. In: Gruber, J. and Wise, D., Social Security Programs And Retirement Around the World. The Relationship to Youth Employment. Chicago: University of Chicago Press.

Blanco, A. (2000) The Decision of Early Retirement in Spain. FEDEA Working Paper n. 76.

Dorn, D. & Alfonso Sousa-Poza, 2003. Why is the Employment Rate of Older Swiss so High? An Analysis of the Social Security System. The Geneva Papers on Risk and Insurance, 28 (4) 652-672.

Euwals, R., van Vuuren, D., & Wolthoff, R. (2010). Early retirement behaviour in the Netherlands: Evidence from a policy reform. De Economist, 158(3), 209-236

Gruber, J. and D. A. Wise (1999). Social Security and Retirement Around the World. University of Chicago Press.

Gruber J., and Wise D.A. (2010), Social Security Programs and Retirement around the World: The Relationship to Youth Employment. University of Chicago Press, Chicago.

Hamblin, K. A. (2013) Active Ageing in the European Union. Policy Convergence and Divergence. London: Palgrave Macmillan.

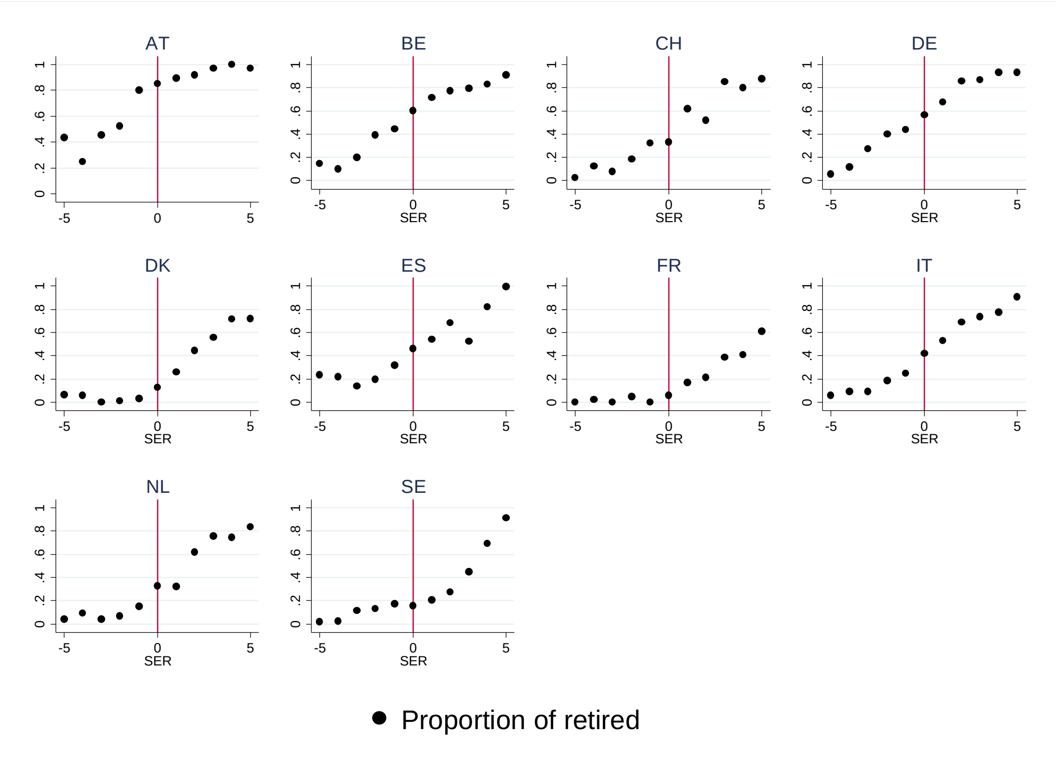
Jousten, A., Lefèbvre, M., Perelman, S. and Pestieau, P. (2010). The Effects of Early Retirement on Youth Unemployment: The Case of Belgium. In: Gruber, J. and Wise, D., Social Security Programs And Retirement Around the World. The Relationship to Youth Employment. Chicago: University of Chicago Press.

Mazzonna, F. & Peracchi, F. (2014). Unhealthy retirement? EIEF working paper 09/14.

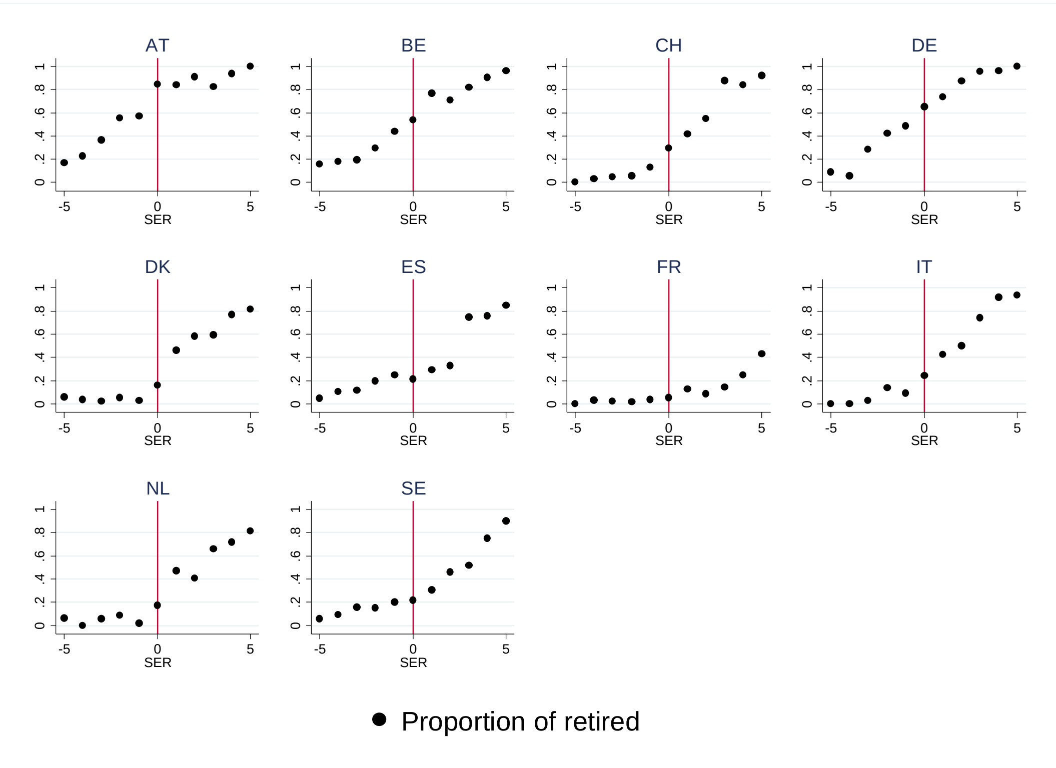
Staubli, S. & Zweilmuller, J. (2011). Does Raising the Retirement Age Increase Employment of Older Workers? IZA Discussion Paper 5863.

Wise D.A. (2012), Social Security Programs and Retirement around the World: Historical Trends in Mortality and Health, Employment, and Disability Insurance Participation and Reforms. University of Chicago Press, Chicago.

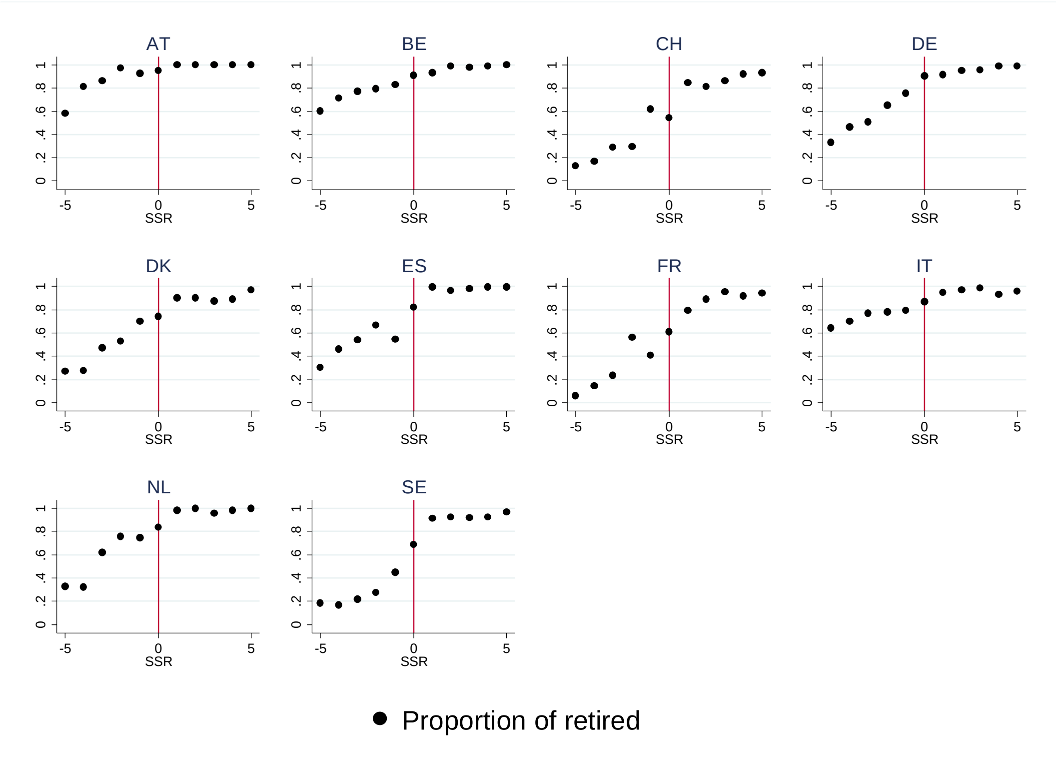
### Figure B.1: Males - Early retirement




### Figure B.2: Females - Early retirement



## Figure B.3: Males - Statutory retirement




## Figure B.4: Females - Statutory retirement

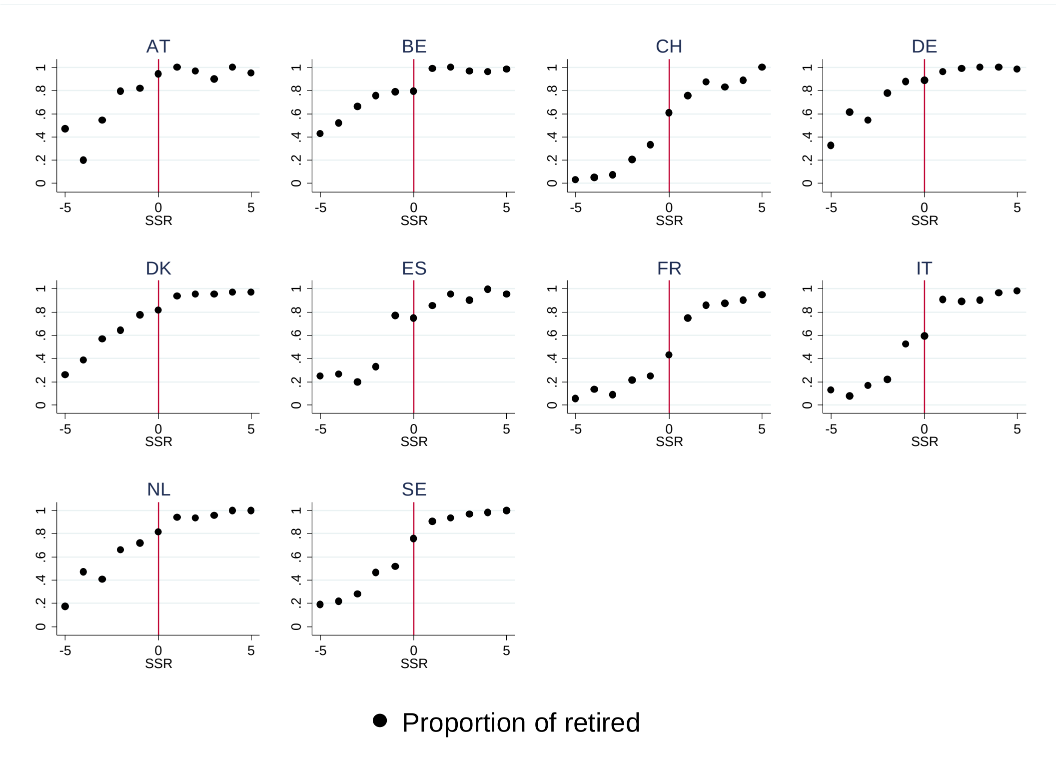


**Figure B.5: Early and statutory retirement eligibility ages (Males)**

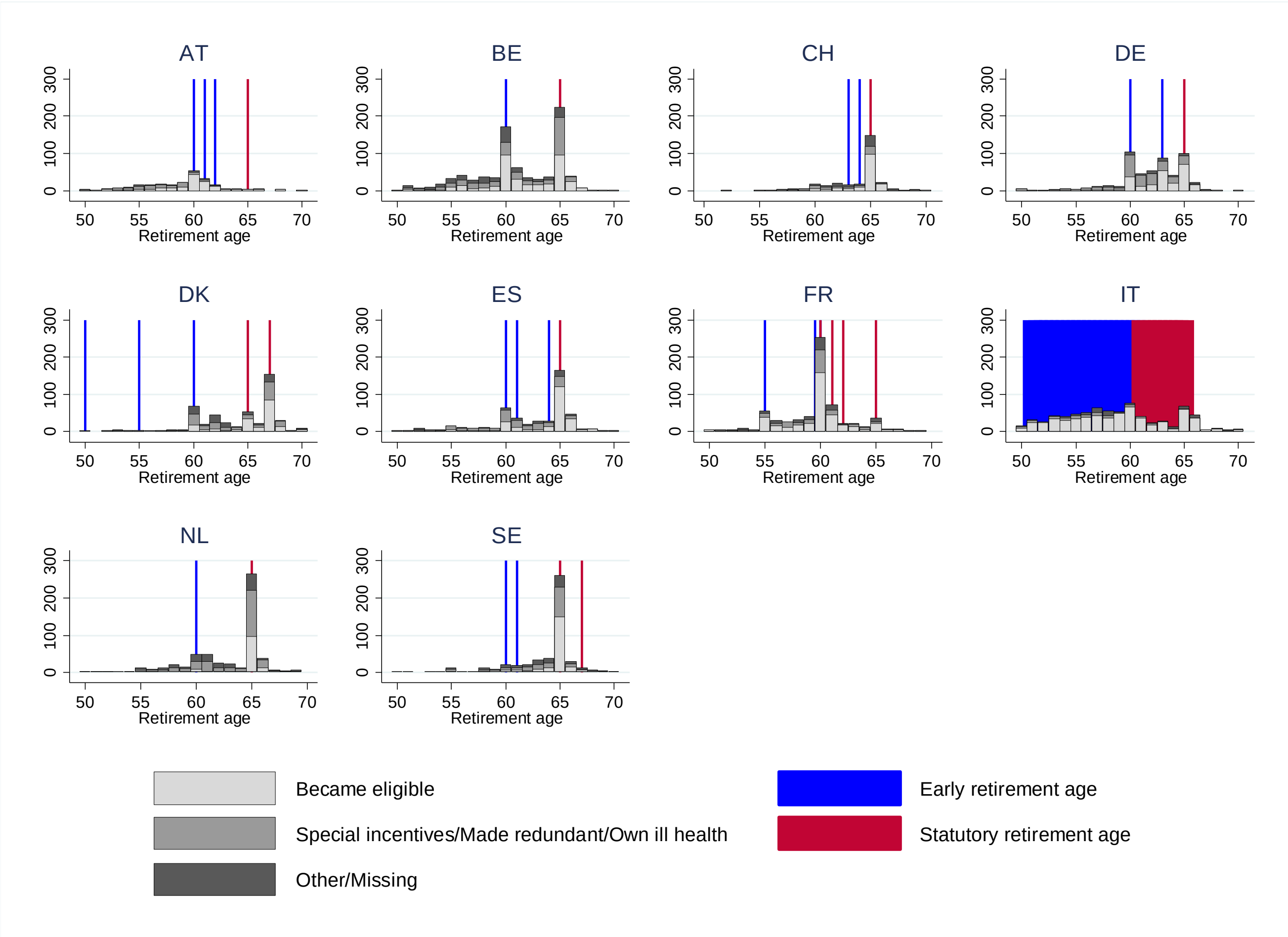


**Figure B.6: Early and statutory retirement eligibility ages (Females)**