UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche ed Aziendali "Marco Fanno"

EARLY RETIREMENT AND COGNITIVE DECLINE. A LONGITUDINAL ANALYSIS USING SHARE DATA

MARTINA CELIDONI
University of Padova

CHIARA DAL BIANCO
University of Venezia

GUGLIELMO WEBER
University of Padova

December 2013

"MARCO FANNO" WORKING PAPER N.174

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
Cannareggio 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

19 December 2013

Abstract

We use 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. 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. The evidence we produce confirms the 'mental retirement' hypothesis and suggests its relevance for the onset of dementia.

Keywords: Ageing, cognition, retirement, instrumental variable estimation

JEL: I12, I1, J26

1 Introduction

Cognitive abilities later in life have been traditionally and largely 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)

The progressive cognitive decline which is likely to interfere with independent functioning is called dementia in psychiatric medicine (American Psychiatric Association, 2000); it has recently received particular attention also in the economic literature due to the generated costs for public and private health expenditure. According to a recent study about 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. As Mazzonna and Peracchi (2012) notice, understanding the decline in cognitive abilities is important for economists also because cognition is relevant for decision making.

According to Salthouse (2006), individual heterogeneity in cognition levels and the rate of age-related change in cognitive functioning are likely to be associated with individuals' lifestyle such as the engagement in mentally stimulating activities. Maguire *et al.* (2000) show how stimulating environment and activities are able to improve, or maintain, cognitive functioning. Particular attention has been devoted to retirement (Adam *et al.*, 2007; Mazzonna and Peracchi, 2012; Rohwedder and Willis, 2010; Bingley and Martinello, 2013; Coe and Zamarro, 2011; Bonsang *et al.*, 2012) since this transition could represent a remarkable change in individuals' life-style and involvement activities that are likely to affect cognition at old ages.

A well-established hypothesis is that retirement has a negative impact on health, mostly mental health (Minkler, 1981), since it can be a stressful event and lead to a break with support networks and friends, with emotional and mental negative impacts. According to Rohwedder and Willis (2010), retirees are engaged in less mental exercise than workers: the latter are exposed to environments that are considered more cognitively challenging and stimulating compared to the non-work condition (the so-called *unengaged lifestyle hypothesis*). On the other hand, some authors, such as Ekerdt *et al.* (1983), Gall *et al.* (1997), Mein *et al.* (2003), Mojon-Azzi *et al.* (2007) and Westerlund *et al.* (2010), suggest that retirement could be beneficial to the extent that it eliminates work-related stress and preserves the health of retirees. Drentea (2002) especially supports the hypothesis that work is alienating and retirement liberating, because retirees experience less anxiety and distress.

Understanding the consequences of retirement on health becomes even more important if we consider that, recently, most developed countries have passed reforms aimed at increasing retirement age to ensure the financial sustainability of social security systems. A positive side effect could be the reduction of long-term care expenditure if later retirement delays the onset of dementia (Dave *et al.*, 2008; Bonsang *et al.*, 2012).

As pointed out in the literature, 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, or a selection problem could arise. As Coe and Zamarro (2011, p. 78) note, 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. For instance, Charles (2004) looks at longitudinal data, drawn from the Health and Retirement Study (HRS), and finds that depressed

*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 and the 4th SHARE user conference held in Liège, November 2013. We are grateful for comments and suggestions made by participants.

people tend to select into retirement.

Endogeneity bias is typically addressed by exploiting changes in public pension eligibility rules as a source of exogenous variation.

In this paper we adopt the same strategy to account for the endogenous nature of retirement: we use eligibility ages for early retirement and 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 of public pension eligibility reflects gender, time of retirement and country of residence, and this ensures that the instruments based on it are informative.

We improve upon most recent papers in several dimensions:

- we follow a longitudinal approach using SHARE data, that is we compare word recalled by the same individual across waves (as in Bonsang *et al.*, 2012);
- we adopt a sharper measure of cognitive decline than the standard change in cognition score, based on a 20% drop in the number of words recalled across waves (see Dal Bianco *et al.*, 2013);
- we show that the measure we use predicts well the onset of dementia in a commonly used US data set, the Aging, Demographics and Memory Study (ADAMS).

The cross-sectional variability in the data, largely exploited so far in the literature, could be uninformative about the potential cumulative effect of years spent in retirement on cognitive decline, if people born in different periods of time have enjoyed different standards of living (the current young old are typically much better off than the young old thirty years ago). These differences are known as cohort effects. To control for cohort effects one could pool together several cross sections, and condition on year of birth, but this type of analysis requires very large samples (with repeated cross sections, year-of-birth averages are assumed to converge to the true population means, as explained in Deaton, 1985). An arguably better solution is to use datasets where the same individuals are followed over a period of time (panel data).

The sharper measure of cognitive decline based on a 20% drop in words recalled is arguably 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. The authors argue that high decreases in words recalled are more informative about actual declines, as opposed to straight changes in the score.

We are able to show in this paper that the 20% decline measure is highly predictive of the onset of dementia: using a small sample of 432 individuals aged 70 or more who took part in ADAMS and were later medically assessed for dementia, we show that our test correctly classifies 74% of individuals according to their later dementia status.

Our econometric analysis shows that retirement status *per-se* has not 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) results. Our evidence therefore supports the cumulative negative effect hypothesis. Early retirement increases the probability of experiencing cognitive decline at older ages - and this reinforces the view that it can be detrimental to the well being of individuals (as argued for other reasons by Angelini *et al.* (2009)).

We find that variables capturing 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. We estimate a protective role of physical activity but also activities in general if performed daily. Job characteristics play a role in terms of job experience: individuals with high job experience (more than 40 years), are more likely to experience a drop in cognition (for a given age, they are more likely to have had low-skilled jobs). Finally, test administration features are important (whether the test is done in presence of someone else or there were distracting noises during the test) - they explain part of the heterogeneity that we see in the data.

The paper is organized as follows. Section 2 reviews the literature about the effect of retirement on cognition based on memory score levels and presents the issues in considering changes in cognition. Section 3 presents the data and section 4 describes the empirical strategy. Section 5 comments the results, section 6 concludes.

2 Literature review

It is largely documented in the literature that ageing is associated with a decline in the ability to perform several cognitive tasks (Dixon *et al.*, 2004), called *fluid abilities*, that include also episodic memory.[1]

This decline in cognitive abilities, also in episodic memory, though is heterogeneous across the population, with some individuals being able to more efficiently use their cognitive resources.

Salthouse (2006) stressed the role of individuals' lifestyle in shaping heterogeneity in cognition levels and rate of age-related change in cognitive functioning. If this is the case, retirement, that represents a remarkable change in life-style and involvement activities, plays an important role for cognition at older ages.

One of the first paper that analyses the causal impact of retirement on levels of cognitive abilities is Rohwedder and Willis (2010) who emphasize the *use it or lose it* hypothesis, i.e. an undemanding environment is likely to enhance the process of cognitive decline. The authors use data drawn from HRS, SHARE and ELSA in 2004 to investigate this effect and exploit the cross-country variation in pension policies to instrument retirement and assess causality. They find a significant and quantitatively important negative effect of retirement on cognitive functioning, i.e. a drop close to 40% in average cognitive score. The measure of cognition used in their analysis is the total number of words remembered in questions of immediate and delayed recall as in Adam *et al.* (2007) and several other papers. The endogenous variable *retired* equals one if a person reports she is not currently working for pay. They include therefore not only people that are retired but also out-of-the-labour-force individuals, unemployed or disabled, relying on a very broad definition of retirement.

A drawback of their analysis, as pointed out by Mazzonna and Peracchi (2012), is to consider retirement as a binary treatment, it means in fact identifying a one-time shift in cognitive abilities levels, without allowing for a potential cumulative effect of years spent in retirement. Another limitation of their analysis is the lack of important controls such as gender, education and country of residence as well as early life conditions. Early life conditions, in particular, are recognised to be important in determining the development of both cognitive and socio-emotional skills (Heckman and Cunha, 2007).

Using the same broad definition of retired, Coe and Zamarro (2011) use SHARE data and focus their attention on men who worked at least some time in their lives. They estimate the effect of being retired on several health outcomes, ranging from self-reported health to verbal fluency. The memory score in particular, defined as in Rohwedder and Willis (2010) on a 20-point scale, is found to be unaffected by retirement when increasing the number of controls, even after instrumenting it.

[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).

Another paper dealing with the effect of retirement on cognitive abilities using SHARE cross-sectional data is Mazzonna and Peracchi (2012) who provide also a theoretical framework, borrowed from Grossman (1972), to describe the link between cognitive abilities, ageing and retirement. The underlying idea is that individuals can partly control their level of cognitive capital by investing in cognitive repair activities, i.e. all types of cognitive-promoting behaviour, to mitigate the exogenous age-related deterioration. Differently from previous studies, they do not consider simply the status *not currently working for pay* but they account for the distance from retirement to identify the potential cumulative effect. Looking mainly at the first wave of SHARE, they find a negative and statistically significant effect of years in retirement, after taking into account the endogeneity of the retirement decision, using as instruments the positive parts of the difference between the actual and the legislated ages eligibility for early and normal retirement. They perform also a number of robustness checks, including the refreshment sample of the second wave, adding education and cohort controls as well as a variable that captures intra-household learning effects. Learning effects are found to be positive, significant and higher for high-school graduates.

More recently, Coe *et al.* (2012), using the pooled cross sections of all HRS waves between 1996 and 2008, model the effect of retirement duration on several measures of cognitive abilities including immediate and delayed recall, instrumenting retirement duration with offers of early retirement windows, that are legally required to be nondiscriminatory and therefore unrelated to cognitive functioning. Results do not suggest a negative cumulative effect of years spent in retirement, but even a beneficial effect for blue collars.

The only paper we are aware of that exploits the longitudinal dimension of the data is Bonsang *et al.* (2012). This paper uses HRS data to overcome the limitations of the cross-sectional analysis. The authors report a negative and statistically significant effect of retirement duration on words recalled, but only when the logarithm of duration is used as a regressor (not the level).

Longitudinal analyses of cognitive decline are well known to suffer from re-testing effects (Ferrer *et al.*, 2004), which cause an upward bias in cognitive ability measurement due to learning from tests performed in previous waves. To the extent that re-testing effects introduce non-classical measurement error in the dependent variable, they may cause bias in the estimates reported in Bonsang *et al.* (2012).

In this paper we provide evidence about the effect of years spent in retirement along the lines of Bonsang *et al.* (2012) for a sample of European 50+ individuals, but account for re-testing effects. We follow standard practice and measure cognitive abilities as the sum of words recalled in the first and second trial with a score ranging from 0 to 20 in each wave (as Rohwedder and Willis, 2010; Bonsang *et al.*, 2012), but use an alternative measure of cognitive decline, i.e. 20% drop in words recalled between waves, based on Dal Bianco *et al.* (2013) who argue that it is more informative about the actual decline when re-testing effects are important.

We argue that our measure is useful on the basis of a validation exercise on US data. We use ADAMS and show that our measure of cognitive decline is predictive of a pathological impairment related to dementia.

3 Data

In our empirical analysis we use data drawn from the Survey of Health, Ageing and Retirement in Europe (SHARE)[2] which collects information on health, socio-economic status and social and

[2]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

family networks. The SHARE target population are individuals aged fifty or over who speak the official language of their 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.

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 reinterviewed 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). For the most part we pool males and females. However, we also provide estimates separately by gender to understand whether the more interrupted careers of females affect our results.

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

Coe and Zamarro (2011) instead do not include among the retired those individuals that report being retired, simply because they left their *career job.*

If we compare in our sample the latter definition with the one we adopt, we observe that they are very similar: the 94% of individuals in wave 1 or 2 that declare to be retired did not any work for pay in the previous four weeks. In order to understand whether a different definition of retired can affect our results, we perform our analysis using also Coe and Zamarro (2011)'s version of *retired*: as shown later, results do not change.

Another key variable in our analysis is years spent in retirement which is computed on the basis of the year when the interviewee declares to retire from work asked in SHARELIFE. When this information is missing or the respondent retired between wave 3 and 4, we use a comparable question from wave 4 questionnaire. This definition of years spent in retirement can be considered conservative among waves, since the wording of questions is the same asking precisely retirement from work.

Table 1 here

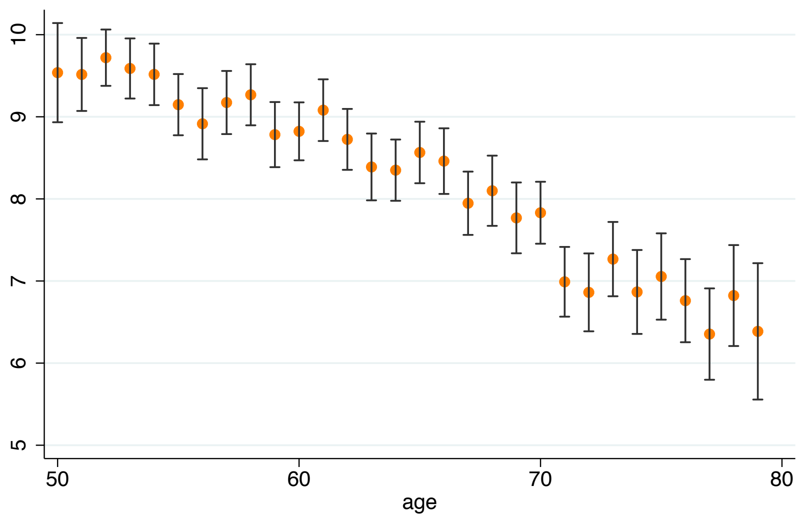
In our analysis, we excluded proxy interviews because in those cases individuals do not perform cognitive tests; we also do not consider interviewees with missing values in tests' scores in at least one of the two measurement occasions (i.e. baseline and wave4). The final sample is

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.

a balanced panel with two time periods for 8262 individuals (6176 observed between wave 1 and 4, 2086 refresher of wave 2).[3]

Data about cognitive abilities are collected in each regular wave of SHARE[4], especially a series of brief tests are included in the CAPI questionnaire. Among the test proposed, only verbal fluency and verbal learning tests are performed in wave 1, 2 and 4.[5] It is worth noting that, since the cognitive decline is a multidimensional phenomenon, each test usually measures a different aspect of the cognitive concept. We are particularly interested in memory scores to provide results that are comparable with the existing literature, we therefore concentrate upon the memory verbal test which is a modified version of the Rey's Auditory Verbal Learning Test-RAVLT. In the ten word-list-learning test the respondent is asked to 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. Unlike the simplify SHARE version, the original RAVLT consists of five consecutive trials each followed by an immediate recall and one delayed recall (Trial 6), which enables to compute several indices from the data collected (Estevez-Gonzalez *et al.*, 2003). 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 recall with a score ranging from 0 to 20 in each wave. This test is preferred to verbal fluency 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 this memory score, we will compute our main outcome of interest, i.e. high decrease, as explained in the next subsection.

Figure 1: Average Memory Score by age, Males



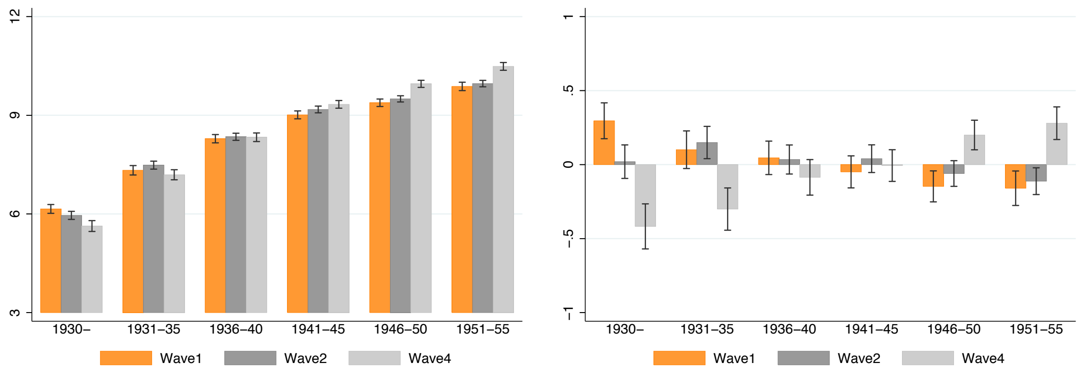
The standard negative association between cognitive abilities and age, that most papers on

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

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

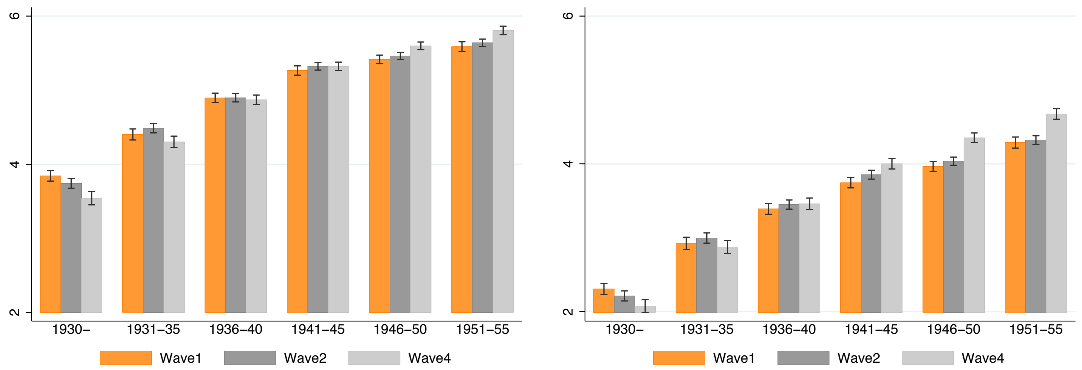
[5]The verbal learning test has the same technical features in all three waves, the only exception is that words used in the fourth wave of SHARE are different with respect to those used in the previous waves for details see Malter and Borsch-Supan (2013).

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: Average Score of Immediate and Delayed Recall, by cohort and wave



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 it 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, 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 word recalled in the first and second trial by cohort and wave. Following cohorts across waves we notice that cognitive decline is not systematic for all but is heavily concentrated among 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 older individuals tend to fall over time as one might expect, whereas the number of words recalled improves over time for younger individuals. Figure 3 represents a more detailed breakdown by immediate and delayed recall; we notice how especially younger cohorts tend to improve their performances in the delayed recall (Figure 3 on the right) over time suggesting a potentially relevant upward bias due to re-testing.

As emphasized in the literature, in fact, longitudinal analysis suffers from learning or re-testing 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. Rabbitt *et al.* (2001) using a 17-year longitudinal study found that fluid abilities decrease over age, with larger declines among older individuals, but at the same time performance improves across measurement occasions. Especially, the positive retest effect persists up to the fourth (and last) occasion, although the size of it decreases across repetitions. Therefore if we are interested in the development of cognitive abilities over time, to correctly identify declines in cognition we need to account for re-testing effects.

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 but a potential problem could arise when the age increase and the retest occasion are highly correlated. This approach helps mostly when retest interval presents variation and there is ample spread in age, but this is not our case since variability of re-test intervals is almost zero. Ferrer *et al.* (2004) stresses the fact that, for more refined strategies, separating changes in cognition due to age from practice effects in short panels, e.g. two-occasion data, requires restrictions to be imposed in the model leading to the impossibility of fully describing the phenomenon.

Table 2 here

To minimize the potential effects of the upward bias due to learning we follow an alternative strategy adopted in Dal Bianco *et al.* (2013). We consider as outcome of interest a high rather than any decrease in cognition between two measurement occasions, where high decrease means a drop higher than 20% of the initial memory score. 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 25-th percentile 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%.

The strategy of analysing high decreases in cognitive abilities could be considered conservative in presence of re-testing effects: we should identify correctly the most vulnerable individuals that would have had a drop in cognition also without learning effects, but we might miss-specified mild decreases since the true drop in cognition could have been mitigated by the re-test effects. Focusing on high decreases rather than any decrease helps in reducing measurement issues, since the latter 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[6], 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 compared the diagnosed cognitive status with the predicted probability estimated with a probit model where we use our measure of high decline, interval length, gender, test features (if there was any distraction factor during the test) and low memory score in baseline, i.e. whether the initial number of words recalled is lower than the median value computed in that specific wave.[7] Table 3 summarises our results: if we are particularly interested in correctly identifying individuals with any sign of dementia, we see that a high decline in cognition is a good measure as in the 73% of cases it corresponds to a clinical assessment of the pathology.

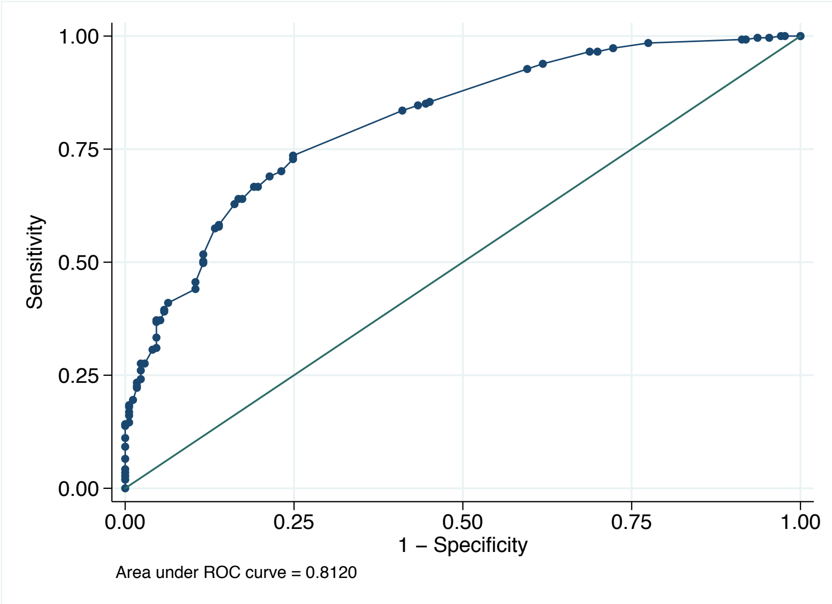
To validate our measure, we provide also results coming from a widely used approach in the field of diseases diagnosis, the Receiver Operating Characteristic (ROC) curve. This curve measures the extent to which a given signal - predictions based on high decreases in cognition in our framework - can detect 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 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

[6]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.

[7]In order to compare predicted and actual classification we used a cutoff of 0.57.

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 0.81 (95% Confidence interval: 0.77-0.85), value considered good in the literature.

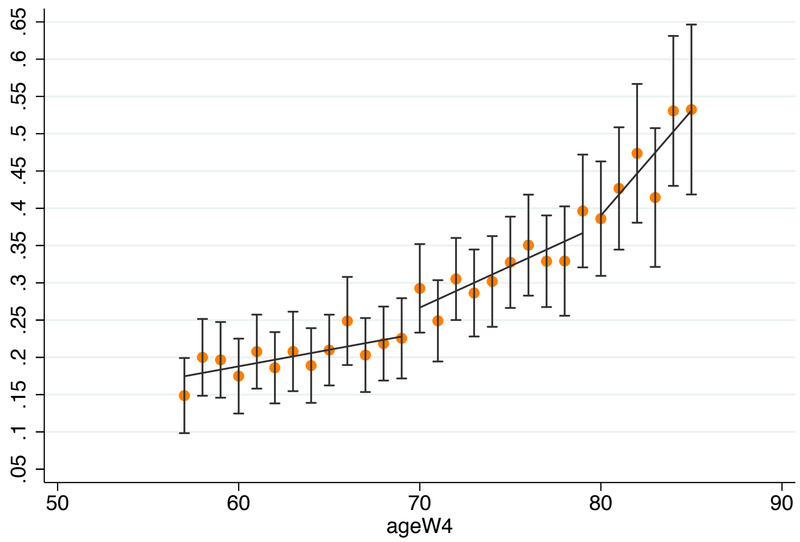
Figure 4: Receiver Operating Characteristic (ROC) curve



After validating our measure of cognitive decline, we now focus on another issue: disentangling the age effect from years spent in retirement. The former refers more to biological changes and age-related cerebral modifications in the brain, the latter instead should capture life-style and activities. There is some evidence that the magnitude of age-related decline accelerates at older ages (Salthouse, 2009), especially, according to a longitudinal analysis in Aartsen *et al.* (2002), cognitive decline might start after midlife, but most often occurs at higher ages (70 or higher). The same result was previously found in Schaie (1989): *most abilities tend to peak in early midlife, plateau until the late fifties or sixties, and then show decline, initially at a slow pace, but accelerating as the late seventies are reached*. Therefore, even if there is still no consensus about when the rate of decline in cognitive abilities begins, there is evidence that after 60, more precisely around 70, the magnitude of age-related decline accelerates.

This can be explained by the *brain reserve capacity* approach. The concept of brain reserve has been proposed in the literature to account for the disjunction between the degree of brain damage or pathology and the related clinical manifestations (Stern, 2009). According to this approach, brains can sustain a number of insults before a clinical deficit emerges, especially the larger the brain the higher the number of insults it sustains before manifesting a pathology. Brain reserve has been codified in the threshold model (Satz, 2009), which recognises not only that there are differences in brain reserve, but also the existence of a critical threshold beyond which specific clinical or functional deficits emerge. This approach is called *passive* because it assumes a fixed cutoff below which functional impairment will occur for everyone; it is complementary to active models, such as cognitive reserve models, which focus on the processes that allow

Figure 5: Percentage of High Decrease in Cognitive Abilities, by age in wave 4



individuals to sustain damage and maintain functions.

In our econometric analysis we control for age effects both in the standard way, that is by introducing a second order age polynomial, and by modeling age effects using a linear spline. The latter approach takes into account the evidence about acceleration in cognitive decline particulary around specific ages (such as 70 and 80). We specify the linear spline by defining the following dummies: individuals younger than 70 in wave 4, those whose age in wave 4 is in between 70 and 79, and individuals that are older than 80.

In Figure 5, we describe, graphically, with fitted regression lines, how the percentages of individuals with high decrease in cognitive abilities change according to the age groups identified. Compared to individuals younger than 70, there is a larger proportion of respondents with high decrease in cognition in the age group 70-79. After 80, cognitive deterioration accelerates and this results in a marked increase in the percentage of high drops in memory scores among very old individuals. We performed a Chow (1960)'s test and we reject the hypothesis of equality among the estimated coefficients for the three age bands.[8]

After presenting all the key variables in this analysis and the related issues, we now turn to the description of the controls (or covariates) included in the model.

3.1 Covariates

We include in our cross-sectional specifications time invariant controls related to education expressed according to the International Standard Classification of Education (ISCED): *highschool* corresponds to ISCED codes 3 and 4 whereas *college* to ISCED codes 5 and 6. We also include some indicators of early-life conditions, such as a zero-one dummy that captures 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

[8]Results are available upon request.

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[9]. As in Coe and Zamarro (2011), we consider the role of jobs' characteristics in shaping cognitive patterns, we especially include controls such as *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. In addition, we include also *jobexperience40* to capture 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 three variables: *nogstest*, *poorhealth*, and *eurodcat*. The first variable is a dummy that takes value one 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). *Poorhealth* indicates that the individual self-declares to be in poor health, whereas *eurodcat* is a dummy that equals one if the individual has at least one symptom of depression.[10]

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. 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.[11] In order to capture the role of activities, we have also *physicalact*, a dummy for whether the individual performs a vigorous physical activity at least weekly, and *dailyact*, that takes value 1 if the individual performs daily, during the last month, activities such as voluntary work, training course, participation in religious or political organisations. 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*. Finally we have also country dummies and a control for the wave 2 refreshment sample.

When we focus on a high decrease in cognitive abilities, we further include controls indicating changes between waves for potentially relevant time varying covariates. 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 _bf 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. *Drophealth* identifies individuals that declared 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. We also control for changes in activities among physical exercise, daily activities and contacts with children: a *drop* means that the individual used to perform the activity and stopped it in wave 4, while *increase*

[9]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.

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

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

means that the respondent started performing the activity in question in wave 4. We include in the model also a dummy, *lowcognition*, that takes value one if baseline memory score was lower than the median value by wave and country, as a 20% decrease is more likely to lead to dementia if the initial value was already low.

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 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.

4 Empirical strategy

The aim of the empirical analysis is to test the cumulative role of years spent in retirement 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:

$$y{*}_i = (score_{i,t} - score_{i,t-1})/score_{i,t-1}. \tag{1}$$

We further define our sharper measure of high cognitive decline, y_i, as a dummy variable that takes value 1 if $y{*}_i$ is lower than -0.2 and zero otherwise.[12] For the most part, we adopt the following linear specification for y_i:

$$y_i = \beta_1 retired_{i,t-1} + \beta_2 yearsinR_{i,t-1} + \beta_3 fromWtoR_{i,t-1} + X_i^T \beta_4 + \epsilon_i \tag{2}$$

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 $t-1$, $retired_{i,t-1}$, years spent in retirement, $yearsinR_{i,t-1}$, and whether we observe the transition from working to retirement between waves, $fromWtoR_i$. We include in the model also a vector of covariates, X_i^T as described above.

The advantage of the 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 IV strategy (Rohwedder and Willis, 2010), we exploit not only the cross-country variability in eligibility ages,

[12] 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.

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 a lot to the European heterogeneity of pension entitlements.

We use the following instruments for $retired_{i,t-1}$, $yearsinR_{i,t-1}$ and $fromWtoR_i$: two dummy variables that take value 1 if the individual is eligible for early and normal retirement, two variables indicating the number of years since eligibility ages for the two type of pensions and two dummies that equal 1 if we observe the transitions from not being eligible to being eligible between the two 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 and are arguably exogenous.

Differently from Coe and Zamarro (2011) and Rohwedder and Willis (2010) that consider only the binary treatment of retirement, we 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, it could be that the effect of retirement is not 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, as argued by Bonsang *et al.* (2012, p.492). 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

In this section we first present estimation results that replicate what has already been done in the literature using HRS or SHARE data. We then look at the specific evidence when the dependent variable is instead the cognitive decline measure described above, that we have shown to predict well the onset of dementia. Finally, we show the results of robustness analysis in a number of directions.

5.1 Replicating previous findings

We first replicate the findings of the existing literature that models memory score levels at a given point in time. In Table 4, we report the coefficients estimated when the outcome variable is the memory score and age effects are modeled either as a set of ten-year wide dummy variables (columns 1 and 2) or as a polynomial (columns 3 and 4). Looking at the OLS estimates, first and third column of Table 4, we notice that among retirement related variables, only years spent in retirement is significant and with a negative sign. This is consistent with a cumulative negative effect on words recalled.

Table 4 here

We compare now OLS parameters with TSLS estimates, second and fourth column[13]. We notice that the coefficient associated to retired is again not significant, whereas the estimated effect of years spent in retirement is negative and significant at the 1% level in both specifications. Table 4 reports also Angrist-Pischke (AP) and Sargan statistics. The former, in particular, tests for weak-identification in presence of more than one endogenous regressor, *retired* and *yearsinR* in this case. The latter, instead, test for the validity of the over-identifying restrictions that are implicitly imposed when the number of instruments exceeds the number of endogenous variables on the right hand side of the equation. According to the test results we see that our instruments are not weak and are valid, i.e. the over-identifying restrictions are not rejected.

Comparing the OLS and TSLS estimates on the variable of interest, *yearsinR*, we observe that the TSLS estimate is larger in absolute value than the OLS estimate. One way to interpret this finding is that OLS suffers from attenuation bias, like in the classical measurement error case. Retirement status is unlikely to be measured with error, as it is based on a self-report, but the engagement associated to having a job is not necessarily captured in a error-free measure by retirement status. In this sense, retirement is an error-ridden measure of the variable of interest. To clarify this point, consider individuals who are not already retired, but who decrease the level of mental exercise on the job because of the nearing prospect of retirement. Rohwedder and Willis (2010, p.128) argue that there is a *on-the-job* retirement effect depending on the remaining working life: *for workers late in their careers, the value of continuing to build work-related human capital is very sensitive to the length of the remaining working life.* At the same time, due to the *honeymoon phase*, individuals right after retirement can engage in different activities that were set aside because of their work. The combination of the two effects might result in a weaker association between retirement and cognitive abilities than the one captured by TSLS. This attenuation bias is in line with what Bonsang *et al.* (2012) find using HRS data: the partial correlation is lower that the causal effect.

Looking at other determinants, as in Coe and Zamarro (2011) and Mazzonna and Peracchi (2012), we find that higher education is associated with higher memory scores on average. Regarding household equivalent income, considered here in quartiles, we find that the higher the income quartile the higher the memory score, in line with the standard health-wealth gradient. Early-life conditions also explain part of the heterogeneity in memory score levels: especially, few books at home during childhood or living in rural areas at the age of ten reduces the number of words recalled, whereas a better than average self-reported math ability increases cognition. Poor health, i.e. reporting being in poor health and having at least one symptom of depression, is associated with a lower number of words recalled.

Differently from previous studies, we consider also variables related to how the test has been carried out, since they could affect results substantially: having performed the test alone, without the presence of the partner or other relatives, is associated with a higher memory score. Other variables that are statistically significant are: job experience that has negative sign (as it may capture individuals that entered the job market at early ages - probably in low-skill jobs); having a partner has a positive effect on words recalled. Some country dummies are also significant, capturing other differences in words recalled compared to Germany.

Table 5 here

Table 6 here

The following two tables present estimation results that exploit the panel dimension of the data, in line with the analysis carried out on HRS data in Bonsang *et al.* (2012). Table 5 shows

[13] In all our analyses, we test for the endogeneity of retirement-related variables by means of standard Hausman tests and strongly reject the null hypothesis of exogeneity.

fixed effects (FE) estimates for years spent in retirement on memory scores, controlling for time and age respectively in dummies or continuous. Estimation results in our case do not seem to depend on the functional form used for age: years spent in retirement decreases the number of words recalled in both specifications (in this sense the evidence is more robust than the one in Bonsang *et al.* (2012)). Looking at Table 6, where FE-IV estimates are shown, we see that there is a statistically significant negative effect of years spent in retirement on cognition over and above the pure age effect.

5.2 New evidence on cognitive decline

Table 7 here

We now focus on estimation results for the model where the dependent variable is defined to be the cognitive decline measure that we introduced in this paper.

Table 7 reports OLS and TSLS estimates when the high decrease in cognitive abilities indicator is the outcome variable. We propose here again two specifications, one with age dummies (first and second column) and the other one with an age polynomial (third and fourth column). Focusing on OLS estimates (first column), we can see that years spent in retirement is significant with a positive sign, whereas retired and the transition from working to retired have no effect. The positive sing on the key variable of interest (years in retirement) is consistent with the negative sign reported above, where the outcome variable was the change in the test score, because a decline is associated to a fall in the memory test score. 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 being retired and the transition from working into retirement again have no significant effect, whereas years spent in retirement remains highly significant with a positive sign, supporting the negative cumulative hypothesis. The same results hold when controlling for a polynomial in age. 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. It must be kept in mind that in our specification there are several age-related controls, e.g. *poorhealth, nogstest, physicalact, jobexperience40* and their changes between waves, that are likely to capture part of what is generally considered an age effect. This can explain why, after controlling for several factors, the decline in words recalled due to age, seems rather weak.

In the specification we can see again that OLS suffers from attenuation bias compared to TSLS. The likely reason is that retirement is an error-ridden measurement of the disengagement associated to the end of a working career, for the reasons explained above. One could alternatively argue that the TSLS estimates differ from OLS estimates because they capture the effect of retirement on cognition for those groups of the population that are driven in retirement by the pension eligibility criteria that we use as instrumental variables. This LATE (Local Average Treatment Effect) interpretation is less than compelling, as our pension eligibility variables group together individuals who retire as soon as possible (early retirement eligibility) and individuals who instead retire as late as possible (normal retirement eligibility). It is hard to imagine why individuals who retire somewhere in between these two extremes should be systematically different, and in particular why for these individuals the time spent in retirement should have no effect or even a positive effect on cognition.

It is interesting also to look at other variables included in the model 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 exception of Mazzonna and Peracchi (2012), or variables associated with activities, often proposed as determinants of cognitive aging patterns (Salthouse, 2006). Table 7 shows 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 did the test alone 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. As far as changes in activities are concerned, we find that physical activity seems to have a protective role for cognitive abilities (according the old-time say "*mens sana in corpore sano*", a healthy mind in a health body). Also an increase or a drop in such activities have a significant effect on the probability of a high decrease, but these represent only associations that might not correspond to the true causal effect. 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. These effects of course could be due to reverse causality (we stress that all key estimation results are unaffected by dropping this group of variables).

At the bottom of Table 7 we report the Angrist-Pischke and Sargan statistics. On the basis of the test results, we conclude that the instruments used are not weak and the over-identifying restrictions are not rejected.

Table 8 here

The estimates shown in Table 7 are based on a linear probability model. This model normally produces a reasonable approximation to non-linear models that take into account the discrete, 0-1, nature of the dependent variables, such as probit or logit. However, the linear probability model cannot be used for prediction, because it can predict out of range. On the other hand, it is easy to estimate by instrumental variables (IV) the linear probability model; it is less easy to obtain IV estimates for a probit model if some endogenous regressor is a discrete variable, as in our case. Fortunately for us, these discrete endogenous regressors have an insignificant effect on the outcome variable and we can drop them from the specification retaining consistency.

To properly quantify the effect of one year more spent in retirement on the probability to experience a high decrease in cognitive abilities, we provide also IV probit estimates (Table 8), considering in the specification the only significant retirement related variable, *yearsinR* that we treat as continuous. Looking at Table 8 where marginal effects are reported when we alternatively control for age effects by means of age dummies and an age polynomial, we see that *yearsinR* is highly significant. We find that an additional year spent in retirement increases the probability of experiencing a drop in cognition by 1.24% (95% Conf. Interval: 0.82-1.66%) or 1.61% (95% Conf. Interval: 0.52-2.71%) depending on how age is specified. These marginal effects are of course computed at the sample averages of all continuous explanatory variables - and at value zero of all 0-1 dummies.

We also observe that the probability of experiencing a drop in memory score increases with age and accelerates starting especially at age 70-79, as the medical literature suggests (Fratiglioni *et al.*, 2008).

5.3 Robustness analysis

We present here the outcome of some of the many robustness analyses that we performed.

Table 9 here

Table 10 here

First, to make sure that we are really capturing a retirement duration effect rather than an age effect, we restrict our sample to those individuals aged between 50 and 80 in baseline. In Table 9 we report OLS and TSLS whereas in Table 10 we provide IV probit estimates for this restricted sample. We see that results are robust to the exclusion of very old individuals, suggesting that we are properly capturing the role of retirement duration on the probability of experiencing a drop in cognition rather than a pure age effect.

The IV probit estimates in this case show that an addition year in retirement causes an increase of 1.47% (95% Conf. Interval: 0.98-1.96%) and 1.27% (95% Conf. Interval: 0.06-2.48%) that are not very different from Table 8 coefficients.

Table 11 here

We also ran the same IV probit analysis separately by gender (see Table 11) to understand whether results are driven by the inclusion of a rather special sample of females and if there are gender specific differences in retirement duration effects. Looking at Table 11, we can see that one year more in retirement increases the chances of experiencing a high decrease in memory score by 1.02% for men (95% Conf. Interval: 0.45-1.59%) and 1.41% for women (95% Conf. Interval: 0.80-2.03%). A noticeable difference has to do with early life conditions: only *mathskills* is significant for both males and females, whereas *ruralarea* increases the probability of a cognitive drop only for men.

Table 12 here

Table 12 instead shows IV probit estimates separately by first and second trial of the memory test (or immediate and delayed recall) to understand if results are stable when we consider separately the two types of memory measured, i.e. short- and long-term memory. The effect of years spent in retirement is similar for both trials: one year more in retirement increases the probability of a high decrease by 0.89% in the first trial (95% Conf. Interval: 0.06-1.32%) - immediate recall - and by 0.99% in the second trial - delayed recall (95% Conf. Interval: 0.53-1.46%). Both point estimates are lower than those for the sum of first and second trials, shown in Table 8, but this may reflect the less fine measure that is used when the maximum score is 10 rather than 20. It is interesting to notice that the age effect seems to be weaker for the second trial, while early life conditions play a role in both.

Table 13 here

The next robustness check that we ran addresses the issue of how we define retirement. We rely on a self-reported retirement status, without any correction that relates to work activities carried out in the four weeks prior to the interview. Such corrections are instead normally implemented in the literature. We report in Table 13, first and second column, IV probit estimates when we consider as retired those individuals that are permanently out of the labour force, i.e. not only those who declared to be retired from work, but additionally also those who did not do any paid work in the four weeks before the interview in wave 1 and 2 and in the year before the interview in wave 4. With this alternative definition we do not include among the retired those individuals that report being retired, simply because they left their *career job*, even if they work full- or part-time (Coe and Zamarro, 2011). Comparing estimates in Table 8 with Table 13, we can see that results do not point to different conclusions: *yearsinR* continues to be highly significant with positive effect on the probability of observing a high decrease in cognition, with no noticeable differences in magnitude.

As a final robustness check we consider how results change when years in retirement enter the right hand side in logarithmic form as in Bonsang *et al.* (2012). Table 13, third and fourth column, reports the IV probit estimates corresponding to this specification: years in retirement keeps having a positive and significant effect, but we notice that age effects become insignificant. We conclude from this that our key finding is not driven by functional form assumptions.

6 Conclusions

In this paper we have used a new measure of cognitive decline that we have shown to be 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.

Using SHARE data, that cover ten different European countries, we have shown 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. The evidence we have produced confirms the 'mental retirement' hypothesis and suggests its relevance for the onset of dementia.

Our measure of cognitive decline 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 74% of individuals according to their later dementia status.

Considering high decrease in cognition as outcome, our econometric analysis shows that retirement status *per-se* has no effect on cognitive decline, but years spent in retirement has a significant, positive effect, after controlling for age and education. 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. An implication of our analysis is that early retirement increases the risk of cognitive decline at older ages.

We find that early life conditions affect the probability of experiencing cognitive decline: especially doing well in maths at age 10 lowers the probability, whereas living in rural area during childhood increases it. We also estimate a protective role of physical activity as well as of activities in general if performed daily. Even if the causal effect must be established, our findings suggest that there is an association between life-styles and the decline in cognitive abilities that is worth further investigation because of its potential policy implications.

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.

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.

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

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.

Brunello, G., Weber, G. and Weiss, C. T. (2012) Books are forever: Early life conditions, education and lifetime income, IZA Discussion Papers 6386, Institute for the Study of Labor (IZA).

Charles, K. K. (2004) Is retirement depressing?: Labour force inactivity and psychological well-being 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. Borsch-Supan, M. Brandt and G. Weber, De Gruyter, Berlin.

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

Deaton, A. (1985) Panel data from time series of cross-sections, *Journal of Econometrics*, **30**, 109–126.

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.

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.

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 Borsch-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.

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.

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 od adult intellectual development, *American Psychologist*, **49**, 304–313.

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

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

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, *BMJ*, **341**, c6149.

Tables

Table 1: Summary statistics by country

	Age in baseline Mean	Early ret. age - Males Mean	Early ret. age - Females Mean	Normal ret. age - Males Mean
Austria	63	61	56	65
Germany	63	63	63	65
Sweden	64	61	61	65
Netherland	62	60	60	65
Spain	65	61	61	65
Italy	64	60	55	63
France	63	60	60	61
Denmark	69	60	60	66
Switzerland	62	63	60	65
Belgium	64	60	60	65

	Normal ret. age - Females Mean	Years in retirement %	Transitions into retirement %	Retired
Austria	60	6.2	17	73
Germany	65	4.7	15	59
Sweden	65	4.3	21	52
Netherland	65	3.8	17	49
Spain	65	5.7	11	63
Italy	58	7.4	11	73
France	61	6.1	16	61
Denmark	66	6.9	11	84
Switzerland	63	2.9	12	44
Belgium	62	6.2	17	63

Table 2: Percentage decrease in memory score between waves

Males and Females							
Waves \ Percentile	5	10	**25**	50	75	90	95
W1 W4	-0.615	-0.429	**-0.214**	0.000	0.300	0.714	1.00
W2 W4	-0.600	-0.444	**-0.231**	0.000	0.250	0.625	1.00
Males and Females 80-							
W1 W4	-0.584	-0.417	**-0.200**	0.000	0.300	0.714	1.00
W2 W4	-0.571	-0.429	**-0.222**	0.000	0.273	0.625	1.00

Table 3: Predicted probability of cognitive status versus assessed cognitive status

	Predicted		
Assessed	Normal	Demented	All
Normal	75.14%	24.86%	100.0
Demented	27.20%	72.80%	100.0
Total	46.31%	53.69%	100.0
Correctly classified			74.00%

Table 4: Memory score - Males and Females

	Age dummies						Continuous Age					
	OLS			TSLS			OLS			TSLS		
	Coef.	Std.Err.		Coef.	Std.Err.		Coef.	Std.Err.		Coef.	Std.Err.	
retired	-0.0675	0.1066		0.3977	0.3296		0.0289	0.1075		0.3983	0.3177	
yearsinR	-0.0364	0.0084	***	-0.1270	0.0187	***	-0.0168	0.0094	*	-0.1438	0.0404	***
age60_69	-0.4796	0.1028	***	-0.3980	0.2026	**						
age70_79	-1.0092	0.1470	***	-0.2787	0.2662							
age80	-1.7388	0.2353	***	-0.2640	0.3722							
age							0.0183	0.0609		-0.2702	0.1144	**
age^2							-0.0007	0.0005		0.0021	0.0009	**
female	1.0484	0.0678	***	1.0437	0.0682	***	1.0524	0.0676	***	1.0390	0.0685	***
jobexperience40	-0.1836	0.0718	**	-0.3288	0.0765	***	-0.1059	0.0736		-0.3461	0.1143	***
wave_2	0.2853	0.0792	***	0.2841	0.0797	***	0.2898	0.0790	***	0.2854	0.0799	***
eqincomeQ2	0.2782	0.0916	***	0.3043	0.0925	***	0.2692	0.0915	***	0.3051	0.0929	***
eqincomeQ3	0.5042	0.0943	***	0.5346	0.0972	***	0.4914	0.0941	***	0.5297	0.0967	***
eqincomeQ4	0.6413	0.0992	***	0.6591	0.1043	***	0.6257	0.0991	***	0.6490	0.1038	***
missingincome	-0.0928	0.0632		-0.0813	0.0636		-0.0853	0.0632		-0.0774	0.0638	
fewbooks	-0.4863	0.0717	***	-0.4659	0.0727	***	-0.4829	0.0716	***	-0.4547	0.0738	***
mathskills	0.5196	0.0636	***	0.5218	0.0639	***	0.5193	0.0635	***	0.5215	0.0641	***
ruralarea	-0.2718	0.0645	***	-0.2708	0.0648	***	-0.2671	0.0644	***	-0.2742	0.0650	***
highschool	0.7171	0.0791	***	0.6958	0.0797	***	0.7078	0.0790	***	0.6919	0.0801	***
college	1.2137	0.0943	***	1.1591	0.0965	***	1.2214	0.0941	***	1.1481	0.1017	***
public	0.1309	0.0737	*	0.1549	0.0745	**	0.1286	0.0736	*	0.1629	0.0760	**
selfemployed	0.0117	0.0953		-0.0143	0.0972		0.0360	0.0952		-0.0284	0.1017	
partner	0.3465	0.0804	***	0.3146	0.0812	***	0.3361	0.0804	***	0.3226	0.0813	***
nogstest	-0.1663	0.1796		-0.1720	0.1805		-0.1471	0.1793		-0.1612	0.1812	
alonecftest	0.2589	0.0856	***	0.2705	0.0860	***	0.2638	0.0854	***	0.2712	0.0862	***
contextcftest	-0.2085	0.1165	*	-0.2226	0.1171	*	-0.2127	0.1163	*	-0.2196	0.1174	*
poorhealth	-0.5255	0.1719	***	-0.4523	0.1738	***	-0.5493	0.1717	***	-0.4454	0.1794	**
eurodcat	-0.5102	0.0819	***	-0.5098	0.0823	***	-0.5100	0.0818	***	-0.5109	0.0825	***
physicalact	0.1888	0.0637	***	0.1742	0.0644	***	0.1747	0.0637	***	0.1751	0.0644	***
dailyact	0.1721	0.0994	*	0.1502	0.1017		0.1757	0.0993	*	0.1550	0.1016	
dailycontactchild	-0.0517	0.0677		-0.0505	0.0689		-0.0751	0.0679		-0.0649	0.0687	
SE	0.3277	0.1318	**	0.2798	0.1341	**	0.3499	0.1316	***	0.2758	0.1396	**
DK	0.4584	0.1466	***	0.3150	0.1496	**	0.4992	0.1465	***	0.2906	0.1617	*
NL	0.0962	0.1400		0.0664	0.1410		0.0831	0.1397		0.0738	0.1415	
BE	-0.5749	0.1274	***	-0.5356	0.1286	***	-0.5960	0.1271	***	-0.5079	0.1325	***
FR	-0.8151	0.1282	***	-0.7741	0.1292	***	-0.8496	0.1281	***	-0.7489	0.1340	***
CH	0.3210	0.1470	**	0.2525	0.1505	*	0.3355	0.1467	**	0.2546	0.1569	
AT	0.1765	0.1789		0.2706	0.1849		0.1164	0.1790		0.3256	0.2048	
ES	-1.9017	0.1645	***	-1.9274	0.1654	***	-1.9203	0.1641	***	-1.9168	0.1656	***
IT	-1.1141	0.1373	***	-0.9790	0.1410	***	-1.1761	0.1379	***	-0.9215	0.1674	***
Constant	8.2496	0.2016	***	8.2586	0.2119	***	9.3874	1.9672	***	16.8490	3.5461	***
Observations	8262			8262			8262			8262		
Adj R-squared	0.2906			0.2804			0.2927			0.2769		
Sargan statistic (overidentification test of all instruments - p-value)												
				0.6411						0.5380		
Angrist and Pischke first-stage F statistics (Weak identification test)												
				F(3,8222)						F(3,8223)		
retired				303.58						342.19		
yearsinR				679.50						150.71		

*Notes: Significant levels as follows: p-value *** $\leq$ 0.01, ** $\leq$ 0.05, * $\leq$ 0.1.*

Table 5: Memory score - Males and Females - Fixed Effects

	Age dummies			Continuous Age		
	Coef.	Std.Err.		Coef.	Std.Err.	
yearsinR	-0.0963	0.0153	***	-0.1285	0.0131	***
age60_69	0.2473	0.1009	**			
age70_79	0.2535	0.1577				
age80	-0.6348	0.2337	***			
age				-0.3418	0.0871	***
wave_1	-0.7883	0.0943	***	-3.2141	0.5874	***
wave_2	-0.5424	0.0879	***	-2.0262	0.368	***
Constant	10.0279	0.198	***	34.3733	6.087	***
Observations	16524			16524		
R-squared	0.027			0.021		
Number of individuals	8262			8262		

*Notes: Significance levels as follows: p-value *** ≤ 0.01, ** ≤ 0.05, * ≤ 0.1.*

Table 6: Memory score - Males and Females - Fixed Effects - IV

	Age dummies			Continuous Age		
	Coef.	Std.Err.		Coef.	Std.Err.	
yearsinR	-0.1826	0.0213	***	-0.1967	0.0163	***
age60_69	0.0855	0.1048				
age70_79	0.2721	0.1580	*			
age80	-0.4350	0.2365	*			
age				-0.2947	0.0874	***
wave_1	-1.1814	0.1157	***	-3.2280	0.5883	***
wave_2	-0.7776	0.0967	***	-2.0360	0.3685	***
Observations	16524			16524		
R-squared	0.023			0.017		
Number of individuals	8262			8262		
Sargan statistic (p-value) (overidentification test of all instruments)						
	0.8681			0.2463		

*Notes: Significance levels as follows: p-value *** $\leq$ 0.01, ** $\leq$ 0.05, * $\leq$ 0.1.*

Table 7: High decrease - Males and Females

	Age dummies						Continuous Age					
	OLS			TSLS			OLS			TSLS		
	Coef.	Std.Err.		Coef.	Std.Err.		Coef.	Std.Err.		Coef.	Std.Err.	
retired	0.0012	0.0181		0.0102	0.0364		-0.0230	0.0201		-0.0438	0.0774	
fromWtoR	0.0031	0.0175		0.0021	0.0549		-0.0093	0.0177		-0.0160	0.0615	
yearsinR	0.0062	0.0012	***	0.0126	0.0023	***	0.0043	0.0014	***	0.0148	0.0061	**
lowcognition	-0.1708	0.0097	***	-0.1735	0.0098	***	-0.1718	0.0097	***	-0.1733	0.0098	***
female	-0.0179	0.0104	*	-0.0162	0.0104		-0.0180	0.0104	*	-0.0166	0.0104	
wave_2	-0.0032	0.0126		-0.0139	0.0157		0.0025	0.0127		-0.0145	0.0170	
age60_69	0.0409	0.0177	**	0.0269	0.0249							
age70_79	0.0927	0.0229	***	0.0374	0.0335							
age80	0.1651	0.0298	***	0.0527	0.0441							
age							0.0077	0.0105		0.0304	0.0221	
age^2							0.0000	0.0001		-0.0002	0.0001	
jobexperience40	0.0379	0.0109	***	0.0495	0.0117	***	0.0293	0.0111	***	0.0509	0.0173	***
eqincomeQ2	-0.0405	0.0139	***	-0.0408	0.0139	***	-0.0389	0.0139	***	-0.0423	0.0140	***
eqincomeQ3	-0.0391	0.0143	***	-0.0377	0.0146	***	-0.0370	0.0143	***	-0.0400	0.0146	***
eqincomeQ4	-0.0394	0.0151	***	-0.0361	0.0156	**	-0.0374	0.0151	**	-0.0390	0.0156	**
missingincome_bl	0.0469	0.0137	***	0.0470	0.0137	***	0.0472	0.0137	***	0.0466	0.0138	***
missingincomeW4	0.0523	0.0156	***	0.0526	0.0156	***	0.0536	0.0156	***	0.0527	0.0156	***
nomissingincomeW4	-0.0526	0.0121	***	-0.0532	0.0122	***	-0.0531	0.0121	***	-0.0535	0.0122	***
fewbooks	0.0174	0.0109		0.0148	0.0109		0.0175	0.0109		0.0151	0.0111	
mathskills	-0.0328	0.0097	***	-0.0330	0.0097	***	-0.0328	0.0097	***	-0.0329	0.0097	***
ruralarea	0.0209	0.0098	**	0.0212	0.0097	**	0.0213	0.0098	**	0.0218	0.0098	**
highschool	-0.0276	0.0120	**	-0.0255	0.0120	**	-0.0269	0.0120	**	-0.0255	0.0121	**
college	-0.0447	0.0144	***	-0.0391	0.0145	***	-0.0472	0.0144	***	-0.0409	0.0162	**
public	-0.0217	0.0112	*	-0.0247	0.0114	**	-0.0197	0.0112	*	-0.0235	0.0125	*
selfemployed	0.0011	0.0144		0.0052	0.0145		-0.0020	0.0145		0.0027	0.0161	
partner	0.0092	0.0127		0.0120	0.0128		0.0105	0.0127		0.0122	0.0128	
nogstest_bl	0.0178	0.0516		0.0203	0.0515		0.0142	0.0515		0.0198	0.0518	
dropgstest	0.0906	0.0254	***	0.0876	0.0254	***	0.0870	0.0254	***	0.0885	0.0255	***
increasegstest	-0.0052	0.0597		-0.0075	0.0596		-0.0022	0.0596		-0.0078	0.0598	
alonecftest_bl	-0.0661	0.0238	***	-0.0639	0.0238	***	-0.0660	0.0238	***	-0.0634	0.0238	***
dropalonecftest	0.0251	0.0167		0.0217	0.0168		0.0244	0.0167		0.0219	0.0168	
increasealonecftest	-0.0439	0.0262	*	-0.0419	0.0262		-0.0444	0.0262	*	-0.0413	0.0262	

contextcftest_bl	0.1542	0.0800	*	0.1507	0.0803	*	0.1603	0.0799	**	0.1488	0.0803	*
dropcontextcftest	0.1487	0.0274	***	0.1493	0.0274	***	0.1471	0.0274	***	0.1516	0.0275	***
increasecontextcftest	-0.1739	0.0816	**	-0.1700	0.0818	**	-0.1803	0.0815	**	-0.1677	0.0818	**
poorhealth_bl	-0.0044	0.0320		-0.0105	0.0321		-0.0024	0.0320		-0.0100	0.0329	
drophealth	0.0478	0.0213	**	0.0437	0.0214	**	0.0484	0.0213	**	0.0429	0.0217	**
increasehealth	-0.0002	0.0311		-0.0031	0.0310		0.0010	0.0310		-0.0026	0.0311	
eurodcat_bl	0.0287	0.0169	*	0.0263	0.0169		0.0289	0.0169	*	0.0261	0.0170	
dropeurodcat	0.0208	0.0144		0.0183	0.0144		0.0226	0.0144		0.0186	0.0147	
increaseeurodcat	-0.0299	0.0216		-0.0253	0.0216		-0.0293	0.0216		-0.0244	0.0218	
physicalact_bl	-0.0361	0.0124	***	-0.0313	0.0125	**	-0.0333	0.0124	***	-0.0325	0.0126	***
dropphysicalact	0.0448	0.0129	***	0.0431	0.0130	***	0.0437	0.0129	***	0.0439	0.0130	***
increasephysicalact	-0.0341	0.0155	**	-0.0296	0.0155	*	-0.0318	0.0155	**	-0.0302	0.0156	*
dailyact_bl	-0.0752	0.0192	***	-0.0762	0.0194	***	-0.0750	0.0192	***	-0.0740	0.0194	***
dropdailyact	0.0531	0.0355		0.0539	0.0354		0.0525	0.0354		0.0531	0.0355	
increasedailyact	-0.0587	0.0124	***	-0.0587	0.0124	***	-0.0592	0.0124	***	-0.0586	0.0124	***
dailycontactchild_bl	-0.0137	0.0124		-0.0130	0.0125		-0.0129	0.0124		-0.0142	0.0126	
dropdailycontactchild	0.0085	0.0150		0.0087	0.0150		0.0091	0.0150		0.0090	0.0150	
increasedailycontactchild	-0.0043	0.0162		-0.0072	0.0162		-0.0037	0.0162		-0.0082	0.0164	
SE	0.0042	0.0200		0.0083	0.0200		-0.0009	0.0200		0.0063	0.0219	
DK	0.0142	0.0222		0.0192	0.0225		0.0072	0.0223		0.0225	0.0236	
NL	0.0031	0.0212		0.0036	0.0212		0.0024	0.0212		0.0026	0.0214	
BE	-0.0704	0.0194	***	-0.0790	0.0195	***	-0.0650	0.0195	***	-0.0775	0.0226	***
FR	-0.0928	0.0195	***	-0.0990	0.0196	***	-0.0891	0.0195	***	-0.0982	0.0213	***
CH	-0.0135	0.0223		-0.0068	0.0225		-0.0176	0.0223		-0.0108	0.0249	
AT	0.0378	0.0271		0.0244	0.0275		0.0464	0.0272	*	0.0271	0.0334	
ES	-0.0467	0.0255	*	-0.0462	0.0255	*	-0.0466	0.0255	*	-0.0463	0.0255	*
IT	-0.0762	0.0216	***	-0.0902	0.0219	***	-0.0690	0.0217	***	-0.0903	0.0278	***
Constant	0.4038	0.0406	***	0.3995	0.0408	***	-0.0641	0.3735		-0.6824	0.8183	
Observations	8262			8262			8262			8262		
Adj R-squared	0.1004			0.0969			0.1017			0.0955		
Sargan statistic (overidentification test of all instruments - p-value)												
				0.6990						0.8218		
Angrist and Pischke first-stage F statistics (Weak identification test)												
				F(4,8201)						F(4,8202)		
retired				433.45						97.38		
fromWtoR				159.51						126.84		
yearsinR				755.56						110.84		

Table 8: High decrease - Males and Females - IV PROBIT

	Age dummies			Continuous Age		
	dy/dx	Std.Err.		dy/dx	Std.Err.	
yearsinR	0.0124	0.0022	***	0.0161	0.0056	***
low_cognition	-0.1865	0.0105	***	-0.1866	0.0105	***
female	-0.0157	0.0110		-0.0156	0.0110	
wave_2	-0.0141	0.0129		-0.0198	0.0159	
age60_69	0.0449	0.0190	**			
age70_79	0.0705	0.0263	***			
age80	0.0797	0.0465	*			
age				0.0319	0.0122	***
age^2				-0.0002	0.0001	**
jobexperience40	0.0496	0.0119	***	0.0556	0.0166	***
eqincomeQ2	-0.0386	0.0134	***	-0.0394	0.0135	***
eqincomeQ3	-0.0377	0.0140	***	-0.0379	0.0141	***
eqincomeQ4	-0.0347	0.0150	**	-0.0334	0.0150	**
missingincome_bl	0.0487	0.0139	***	0.0493	0.0139	***
missingincomeW4	0.0559	0.0175	***	0.0565	0.0175	***
nomissingincomeW4	-0.0555	0.0122	***	-0.0559	0.0122	***
fewbooks	0.0157	0.0114		0.0141	0.0115	
mathskills	-0.0340	0.0101	***	-0.0338	0.0101	***
ruralarea	0.0229	0.0103	**	0.0235	0.0103	**
highschool	-0.0265	0.0123	**	-0.0255	0.0124	**
college	-0.0438	0.0145	***	-0.0410	0.0150	***
public	-0.0260	0.0115	**	-0.0273	0.0117	**
selfemployed	0.0023	0.0149		0.0042	0.0152	
partner	0.0124	0.0132		0.0111	0.0132	
nogstest_bl	0.0138	0.0519		0.0173	0.0524	
dropgstest	0.0803	0.0288	***	0.0816	0.0289	***
increasegstest	-0.0060	0.0586		-0.0095	0.0582	
alonecftest_bl	-0.0650	0.0258	**	-0.0637	0.0258	**
dropalonecftest	0.0186	0.0176		0.0175	0.0176	
increasealonecftest	-0.0372	0.0242		-0.0363	0.0243	
contextcftest_bl	0.1675	0.0958	*	0.1635	0.0959	*
dropcontextcftest	0.1562	0.0334	***	0.1573	0.0334	***
increasecontextcftest	-0.1381	0.0499	***	-0.1356	0.0510	***
poorhealth_bl	-0.0048	0.0317		-0.0090	0.0317	
drophealth	0.0424	0.0228	*	0.0397	0.0229	*
increasehealth	-0.0009	0.0315		-0.0007	0.0315	
eurodcat_bl	0.0285	0.0183		0.0273	0.0183	
dropeurodcat	0.0175	0.0154		0.0161	0.0156	
increaseeurodcat	-0.0261	0.0214		-0.0242	0.0217	
physicalact_bl	-0.0329	0.0132	**	-0.0319	0.0132	**
dropphysicalact	0.0447	0.0144	***	0.0439	0.0143	***
increasephysicalact	-0.0317	0.0155	**	-0.0311	0.0155	**
dailyact_bl	-0.0739	0.0172	***	-0.0742	0.0172	***
dropdailyact	0.0617	0.0420		0.0634	0.0420	
increasedailyact	-0.0615	0.0133	***	-0.0614	0.0134	***
dailycontactchild_bl	-0.0170	0.0130		-0.0161	0.0130	
dropdailycontactchild	0.0105	0.0164		0.0105	0.0164	
increasedailycontactchild	-0.0102	0.0165		-0.0115	0.0166	
SE	0.0068	0.0212		0.0107	0.0219	
DK	0.0214	0.0237		0.0255	0.0252	
NL	0.0052	0.0225		0.0062	0.0225	
BE	-0.0766	0.0178	***	-0.0801	0.0187	***
FR	-0.0974	0.0168	***	-0.0996	0.0172	***
CH	-0.0078	0.0233		-0.0053	0.0237	
AT	0.0287	0.0291		0.0193	0.0310	
ES	-0.0463	0.0235	**	-0.0458	0.0236	*
IT	-0.0831	0.0191	***	-0.0905	0.0213	***
Observations	8262			8262		
Log likelihood	-25659.67			-25512.158		
Amemiya-Lee-Newey minimum chi-sq statistic (p-value)						
	0.8424			0.7379		

*Notes: Significance levels as follows: p-value *** ≤ 0.01, ** ≤ 0.05, * ≤ 0.1.*

Table 9: High decrease - Males and Females 80-

	Age dummies						Continuous Age					
	OLS			TSLS			OLS			TSLS		
	Coef.	Std.Err.		Coef.	Std.Err.		Coef.	Std.Err.		Coef.	Std.Err.	
retired	0.0036	0.0181		0.0022	0.0372		-0.0136	0.0202		-0.0169	0.0823	
fromWtoR	0.0035	0.0175		0.0036	0.0551		-0.0023	0.0179		0.0085	0.0673	
yearsinR	0.0061	0.0013	***	0.0148	0.0028	***	0.0036	0.0014	**	0.0118	0.0063	*
low_cognition	-0.1709	0.0098	***	-0.1738	0.0098	***	-0.1718	0.0098	***	-0.1732	0.0099	***
female	-0.0158	0.0105		-0.0141	0.0105		-0.0161	0.0105		-0.0150	0.0105	
wave_2	-0.0011	0.0127		-0.0134	0.0161		0.0055	0.0128		-0.0051	0.0176	
age60_69	0.0415	0.0176	**	0.0266	0.0248							
age70_79	0.0943	0.0229	***	0.0307	0.0337							
age80	0.1673	0.0299	***	0.0372	0.0455							
age							-0.0094	0.0129		0.0072	0.0289	
age^2							0.0001	0.0001		-0.0000	0.0002	
jobexperience40	0.0353	0.0110	***	0.0507	0.0121	***	0.0268	0.0113	**	0.0427	0.0176	**
eqincomeQ2	-0.0353	0.0141	**	-0.0360	0.0142	**	-0.0328	0.0141	**	-0.0356	0.0143	**
eqincomeQ3	-0.0319	0.0145	**	-0.0307	0.0148	**	-0.0287	0.0145	**	-0.0312	0.0147	**
eqincomeQ4	-0.0328	0.0152	**	-0.0290	0.0158	*	-0.0291	0.0152	*	-0.0305	0.0158	*
missingincome_bl	0.0447	0.0139	***	0.0450	0.0139	***	0.0452	0.0138	***	0.0452	0.0139	***
missingincomeW4	0.0465	0.0158	***	0.0467	0.0157	***	0.0483	0.0157	***	0.0473	0.0157	***
nomissingincomeW4	-0.0501	0.0123	***	-0.0515	0.0123	***	-0.0509	0.0122	***	-0.0518	0.0123	***
fewbooks	0.0156	0.0110		0.0124	0.0111		0.0160	0.0110		0.0139	0.0113	
mathskills	-0.0317	0.0098	***	-0.0320	0.0098	***	-0.0318	0.0098	***	-0.0318	0.0098	***
ruralarea	0.0214	0.0099	**	0.0220	0.0099	**	0.0216	0.0098	**	0.0221	0.0098	**
highschool	-0.0287	0.0121	**	-0.0264	0.0122	**	-0.0279	0.0121	**	-0.0270	0.0122	**
college	-0.0495	0.0145	***	-0.0431	0.0147	***	-0.0519	0.0145	***	-0.0463	0.0164	***
public	-0.0190	0.0113	*	-0.0227	0.0116	**	-0.0171	0.0113		-0.0209	0.0127	*
selfemployed	0.0035	0.0147		0.0078	0.0148		-0.0000	0.0147		0.0044	0.0164	
partner	0.0080	0.0129		0.0103	0.0130		0.0091	0.0129		0.0099	0.0130	
nogstest_bl	0.0185	0.0528		0.0240	0.0529		0.0164	0.0528		0.0203	0.0530	
dropgstest	0.0905	0.0269	***	0.0910	0.0269	***	0.0888	0.0268	***	0.0904	0.0268	***
increasegstest	-0.0013	0.0609		-0.0069	0.0608		0.0001	0.0608		-0.0037	0.0609	
alonecftest_bl	-0.0675	0.0243	***	-0.0656	0.0243	***	-0.0673	0.0242	***	-0.0663	0.0243	***
dropalonecftest	0.0209	0.0169		0.0164	0.0170		0.0198	0.0169		0.0178	0.0170	
increasealonecftest	-0.0419	0.0267		-0.0402	0.0267		-0.0426	0.0267		-0.0412	0.0267	

contextcftest_bl	0.1392	0.0811	*	0.1304	0.0816		0.1434	0.0810	*	0.1366	0.0813	*
dropcontextcftest	0.1484	0.0282	***	0.1499	0.0282	***	0.1468	0.0281	***	0.1493	0.0282	***
increasecontextcftest	-0.1612	0.0827	*	-0.1517	0.0831	*	-0.1657	0.0826	**	-0.1578	0.0829	*
poorhealth_bl	0.0016	0.0327		-0.0062	0.0328		0.0037	0.0327		-0.0028	0.0336	
drophealth	0.0581	0.0220	***	0.0523	0.0221	**	0.0578	0.0220	***	0.0537	0.0223	**
increasehealth	-0.0150	0.0318		-0.0171	0.0318		-0.0135	0.0317		-0.0153	0.0317	
eurodcat_bl	0.0265	0.0172		0.0244	0.0172		0.0269	0.0172		0.0253	0.0172	
dropeurodcat	0.0153	0.0147		0.0126	0.0147		0.0170	0.0147		0.0142	0.0149	
increaseeurodcat	-0.0256	0.0219		-0.0214	0.0219		-0.0254	0.0218		-0.0225	0.0220	
physicalact_bl	-0.0337	0.0125	***	-0.0290	0.0126	**	-0.0307	0.0125	**	-0.0298	0.0127	**
dropphysicalact	0.0431	0.0130	***	0.0410	0.0131	***	0.0413	0.0130	***	0.0411	0.0131	***
increasephysicalact	-0.0313	0.0156	**	-0.0268	0.0157	*	-0.0290	0.0156	*	-0.0276	0.0157	*
dailyact_bl	-0.0748	0.0194	***	-0.0752	0.0196	***	-0.0746	0.0194	***	-0.0737	0.0195	***
increasedailyact	-0.0565	0.0125	***	-0.0563	0.0125	***	-0.0566	0.0125	***	-0.0565	0.0125	***
dropdailyact	0.0555	0.0361		0.0583	0.0361		0.0560	0.0361		0.0570	0.0361	
dailycontactchild_bl	-0.0137	0.0125		-0.0126	0.0127		-0.0127	0.0125		-0.0130	0.0128	
dropdailycontactchild	0.0104	0.0151		0.0105	0.0151		0.0111	0.0151		0.0108	0.0151	
increasedailycontactchild	-0.0037	0.0165		-0.0077	0.0165		-0.0027	0.0165		-0.0064	0.0167	
SE	0.0034	0.0202		0.0081	0.0203		-0.0017	0.0202		0.0042	0.0221	
DK	0.0165	0.0228		0.0238	0.0231		0.0088	0.0228		0.0192	0.0241	
NL	0.0078	0.0214		0.0074	0.0214		0.0065	0.0214		0.0070	0.0215	
BE	-0.0747	0.0196	***	-0.0853	0.0198	***	-0.0698	0.0197	***	-0.0799	0.0228	***
FR	-0.0924	0.0197	***	-0.1008	0.0199	***	-0.0899	0.0197	***	-0.0972	0.0214	***
CH	-0.0141	0.0224		-0.0075	0.0226		-0.0186	0.0224		-0.0120	0.0249	
AT	0.0367	0.0272		0.0191	0.0277		0.0454	0.0272	*	0.0288	0.0337	
ES	-0.0342	0.0258		-0.0349	0.0258		-0.0354	0.0258		-0.0352	0.0257	
IT	-0.0774	0.0218	***	-0.0962	0.0223	***	-0.0695	0.0218	***	-0.0871	0.0280	***
Constant	0.3971	0.0411	***	0.3952	0.0413	***	0.5016	0.4513		0.0618	1.0314	
Observations	7994			7994			7994			7994		
Adj R-squared	0.0951			0.08961			0.0970			0.0933		
Sargan statistic (overidentification test of all instruments - p-value)												
				0.6950						0.8979		
Angrist and Pischke first-stage F statistics (Weak identification test)												
				F(4,7933)						F(4,7934)		
retired				400.40						82.33		
fromWtoR				152.11						101.26		
yearsinR				529.41						103.65		

Table 10: High decrease - Males and Females 80- - IV PROBIT

	Age dummies			Continuous Age		
	dy/dx	Std.Err.		dy/dx	Std.Err.	
yearsinR	0.0147	0.0025	***	0.0127	0.0062	**
low_cognition	-0.1843	0.0104	***	-0.1841	0.0104	***
female	-0.0135	0.0110		-0.0137	0.0110	
wave_2	-0.0152	0.0130		-0.0114	0.0164	
age60_69	0.0409	0.0189	**			
age70_79	0.0521	0.0279	*			
age80	0.0486	0.0481				
age				0.0181	0.0166	
age^2				-0.0001	0.0001	
jobexperience40	0.0513	0.0122	***	0.0466	0.0174	***
eqincomeQ2	-0.0341	0.0137	**	-0.0335	0.0138	**
eqincomeQ3	-0.0304	0.0143	**	-0.0297	0.0143	**
eqincomeQ4	-0.0270	0.0152	*	-0.0263	0.0152	*
missingincome_bl	0.0464	0.0140	***	0.0471	0.0139	***
missingincomeW4	0.0498	0.0175	***	0.0508	0.0176	***
nomissingincomeW4	-0.0535	0.0123	***	-0.0536	0.0123	***
fewbooks	0.0128	0.0115		0.0133	0.0117	
mathskills	-0.0328	0.0101	***	-0.0328	0.0101	***
ruralarea	0.0237	0.0104	**	0.0238	0.0104	**
highschool	-0.0269	0.0124	**	-0.0269	0.0124	**
college	-0.0465	0.0145	***	-0.0474	0.0150	***
public	-0.0243	0.0116	**	-0.0232	0.0118	*
selfemployed	0.0051	0.0151		0.0038	0.0154	
partner	0.0104	0.0134		0.0097	0.0134	
nogstest_bl	0.0179	0.0532		0.0159	0.0531	
dropgstest	0.0859	0.0306	***	0.0847	0.0306	***
increasegstest	-0.0062	0.0591		-0.0037	0.0595	
alonecftest_bl	-0.0667	0.0263	**	-0.0667	0.0263	**
dropalonecftest	0.0128	0.0177		0.0135	0.0177	
increasealonecftest	-0.0355	0.0245		-0.0357	0.0245	
contextcftest_bl	0.1447	0.0956		0.1490	0.0965	
dropcontextcftest	0.1572	0.0343	***	0.1561	0.0343	***
increasecontextcftest	-0.1265	0.0522	**	-0.1284	0.0519	**
poorhealth_bl	-0.0025	0.0324		-0.0017	0.0327	
drophealth	0.0505	0.0237	**	0.0515	0.0239	**
increasehealth	-0.0125	0.0312		-0.0109	0.0313	
eurodcat_bl	0.0266	0.0185		0.0268	0.0186	
dropeurodcat	0.0113	0.0155		0.0121	0.0157	
increaseeurodcat	-0.0226	0.0217		-0.0231	0.0218	
physicalact_bl	-0.0297	0.0133	**	-0.0295	0.0132	**
dropphysicalact	0.0414	0.0143	***	0.0412	0.0143	***
increasephysicalact	-0.0282	0.0156	*	-0.0284	0.0156	*
dailyact_bl	-0.0734	0.0172	***	-0.0736	0.0172	***
dropdailyact	0.0678	0.0433		0.0678	0.0433	
increasedailyact	-0.0585	0.0134	***	-0.0586	0.0134	***
dailycontactchild_bl	-0.0161	0.0131		-0.0159	0.0131	
dropdailycontactchild	0.0123	0.0165		0.0126	0.0165	
increasedailycontactchild	-0.0107	0.0167		-0.0097	0.0169	
SE	0.0075	0.0214		0.0067	0.0219	
DK	0.0257	0.0243		0.0223	0.0256	
NL	0.0088	0.0226		0.0095	0.0227	
BE	-0.0827	0.0175	***	-0.0796	0.0188	***
FR	-0.0987	0.0167	***	-0.0965	0.0172	***
CH	-0.0080	0.0232		-0.0092	0.0235	
AT	0.0201	0.0288		0.0253	0.0319	
ES	-0.0367	0.0241		-0.0361	0.0242	
IT	-0.0889	0.0190	***	-0.0842	0.0220	***
Observations	7994			7994		
Log likelihood	-24705.463			-24559.491		
Amemiya-Lee-Newey minimum chi-sq statistic (p-value)						
	0.5936			0.8666		

*Notes: Significance levels as follows: p-value *** $\leq$ 0.01, ** $\leq$ 0.05, * $\leq$ 0.1.*

Table 11: High decrease - Separately by gender - IV PROBIT

	Males			Females		
	dy/dx	Std.Err.		dy/dx	Std.Err.	
yearsinR	0.0102	0.0029	***	0.0141	0.0031	***
lowcognition	-0.1971	0.0142	***	-0.1745	0.0155	***
wave_2	-0.0240	0.0170		0.0040	0.0202	
age60_69	0.0372	0.0271		0.0560	0.0269	**
age70_79	0.0806	0.0367	**	0.0676	0.0381	*
age80	0.1037	0.0646		0.0812	0.0688	
jobexperience40	0.0479	0.0159	***	0.0494	0.0189	***
eqincomeQ2	-0.0590	0.0177	***	-0.0130	0.0205	
eqincomeQ3	-0.0610	0.0184	***	-0.0174	0.0216	
eqincomeQ4	-0.0529	0.0197	***	-0.0225	0.0234	
missingincome_bl	0.0472	0.0187	**	0.0471	0.0211	**
missingincomeW4	0.0558	0.0236	**	0.0463	0.0261	*
nomissingincomeW4	-0.0343	0.0168	**	-0.0825	0.0180	***
fewbooks	0.0243	0.0154		0.0080	0.0172	
mathskills	-0.0404	0.0135	***	-0.0259	0.0152	*
ruralarea	0.0327	0.0141	**	0.0070	0.0152	
highschool	-0.0475	0.0169	***	-0.0094	0.0182	
college	-0.0551	0.0202	***	-0.0370	0.0212	*
public	-0.0334	0.0160	**	-0.0163	0.0170	
selfemployed	0.0086	0.0190		-0.0113	0.0241	
partner	-0.0144	0.0213		0.0289	0.0178	
nogstest_bl	0.0358	0.0741		-0.0319	0.0689	
dropgstest	0.1003	0.0421	**	0.0656	0.0400	
increasegstest	-0.0252	0.0795		0.0416	0.0937	
alonecftest_bl	-0.0807	0.0307	***	-0.0253	0.0486	
dropalonecftest	0.0283	0.0227		0.0004	0.0283	
increasealonecftest	-0.0760	0.0269	***	0.0484	0.0544	
contextcftest_bl	0.0962	0.1206		0.2822	0.1534	*
dropcontextcftest	0.1307	0.0447	***	0.1829	0.0506	***
increasecontextcftest	-0.1078	0.0792		-0.1727	0.0553	***
poorhealth_bl	0.0242	0.0462		-0.0383	0.0431	
drophealth	0.0507	0.0307	*	0.0337	0.0344	
increasehealth	-0.0410	0.0388		0.0456	0.0523	
eurodcat_bl	0.0137	0.0298		0.0329	0.0229	
dropeurodcat	0.0175	0.0221		0.0168	0.0216	
increaseeurodcat	0.0078	0.0375		-0.0408	0.0261	
physicalact_bl	-0.0450	0.0179	**	-0.0139	0.0199	
dropphysicalact	0.0550	0.0194	***	0.0323	0.0216	
increasephysicalact	-0.0059	0.0220		-0.0612	0.0215	***
dailyact_bl	-0.0657	0.0243	***	-0.0823	0.0246	***
dropdailyact	0.0579	0.0565		0.0609	0.0633	
increasedailyact	-0.0605	0.0173	***	-0.0600	0.0212	***
dailycontactchild_bl	-0.0097	0.0178		-0.0168	0.0193	
dropdailycontactchild	0.0196	0.0216		-0.0064	0.0253	
increasedailycontactchild	-0.0156	0.0233		0.0061	0.0239	
SE	-0.0267	0.0283		0.0488	0.0325	
DK	0.0300	0.0339		0.0224	0.0341	
NL	-0.0272	0.0276		0.0443	0.0388	
BE	-0.1111	0.0221	***	-0.0360	0.0292	
FR	-0.1198	0.0218	***	-0.0674	0.0265	**
CH	-0.0492	0.0291	*	0.0506	0.0384	
AT	0.0212	0.0399		0.0476	0.044	
ES	-0.0863	0.0271	***	0.0116	0.048	
IT	-0.1214	0.0236	***	-0.0338	0.0324	
Observations	4674			3588		
Log likelihood	-14357.525			-11139.957		
Amemiya-Lee-Newey minimum chi-sq statistic (p-value)						
	0.7988			0.7317		

*Notes: Significance levels as follows: p-value *** $\leq$ 0.01, ** $\leq$ 0.05, * $\leq$ 0.1.*

Table 12: High decrease - Separately by trial - IV PROBIT

	First trial			Second trial		
	dy/dx	Std.Err.		dy/dx	Std.Err.	
yearsinR	0.0089	0.0022	***	0.0099	0.0024	***
low_cognition	-0.1300	0.0105	***	-0.1962	0.0111	***
female	-0.0197	0.0111	*	-0.0206	0.0117	*
wave_2	-0.0081	0.0132		-0.0102	0.0140	
age60_69	0.0513	0.0193	***	0.0352	0.0193	*
age70_79	0.1003	0.0269	***	0.0702	0.0275	**
age80	0.1259	0.0487	**	0.0703	0.0484	
jobexperience40	0.0481	0.0120	***	0.0269	0.0127	**
eqincomeQ2	-0.0073	0.0142		-0.0454	0.0147	***
eqincomeQ3	-0.0022	0.0148		-0.0494	0.0152	***
eqincomeQ4	-0.0113	0.0157		-0.0466	0.0162	***
missingincome_bl	0.0337	0.0143	**	0.0575	0.0150	***
missingincomeW4	0.0462	0.0174	***	0.0600	0.0184	***
nomissingincomeW4	-0.0343	0.0126	***	-0.0624	0.0132	***
fewbooks	0.0256	0.0116	**	0.0266	0.0123	**
mathskills	-0.0204	0.0102	**	-0.0399	0.0108	***
ruralarea	0.0183	0.0104	*	0.0130	0.0110	
highschool	-0.0192	0.0126		-0.0219	0.0134	
college	-0.0424	0.0148	***	-0.0496	0.0157	***
public	-0.0123	0.0118		-0.0189	0.0125	
selfemployed	0.0081	0.0152		0.0118	0.0162	
partner	0.0078	0.0134		0.0121	0.0143	
nogstest_bl	0.0250	0.0541		-0.0655	0.0506	
dropgstest	0.1012	0.0296	***	-0.0068	0.0278	
increasegstest	-0.0152	0.0587		0.1037	0.0742	
alonecftest_bl	-0.0569	0.0259	**	-0.0385	0.0270	
dropalonecftest	0.0162	0.0178		0.0343	0.0191	*
increasealonecftest	-0.0379	0.0249		-0.0225	0.0280	
contextcftest_bl	-0.0109	0.0829		0.2309	0.0972	**
dropcontextcftest	0.1351	0.0328	***	0.1557	0.0338	***
increasecontextcftest	-0.0153	0.0839		-0.1860	0.0538	***
poorhealth_bl	-0.0200	0.0314		0.0181	0.0361	
drophealth	0.0807	0.0240	***	0.0283	0.0242	
increasehealth	0.0242	0.0334		-0.0192	0.0335	
eurodcat_bl	0.0311	0.0186	*	0.0024	0.0191	
dropeurodcat	-0.0086	0.0151		0.0100	0.0164	
increaseeurodcat	-0.0109	0.0223		-0.0279	0.0237	
physicalact_bl	-0.0036	0.0133		-0.0204	0.0141	
dropphysicalact	0.0194	0.0140		0.0407	0.0150	***
increasephysicalact	-0.0220	0.0161		-0.0307	0.0170	*
dailyact_bl	-0.0637	0.0180	***	-0.0592	0.0201	***
dropdailyact	0.0329	0.0402		0.0611	0.0435	
increasedailyact	-0.0518	0.0135	***	-0.0424	0.0141	***
dailycontactchild_bl	-0.0067	0.0132		-0.0054	0.0140	
dropdailycontactchild	0.0207	0.0167		-0.0093	0.0170	
increasedailycontactchild	-0.0144	0.0167		-0.0170	0.0178	
SE	0.0156	0.0215		0.0597	0.0241	**
DK	0.0207	0.0237		0.0804	0.0270	***
NL	-0.0106	0.0220		0.0608	0.0257	**
BE	-0.0750	0.0182	***	-0.0372	0.0215	*
FR	-0.1021	0.0171	***	-0.0551	0.0210	***
CH	-0.0007	0.0236		0.0337	0.0265	
AT	0.0298	0.0294		0.0717	0.0328	**
ES	-0.0943	0.0212	***	-0.0218	0.0278	
IT	-0.0853	0.0195	***	-0.0365	0.0237	
Observations	8262			8262		
Log likelihood	-25835.982			-26178.078		
Amemiya-Lee-Newey minimum chi-sq statistic (p-value)	0.3594			0.4670		

*Notes: Significance levels as follows: p-value *** $\leq$ 0.01, ** $\leq$ 0.05, * $\leq$ 0.1.*

Table 13: High decrease - Alternative definition of R and non linear effects - IV PROBIT

	Alternative definition of R			Non linear effects		
	dy/dx	Std.Err.		dy/dx	Std.Err.	
yearsinR	0.0123	0.0022	***			
log(yearsinR+1)				0.0774	0.0140	***
lowcognition	-0.1866	0.0105	***	-0.1864	0.0105	***
female	-0.0167	0.0109		-0.0149	0.0110	
wave_2	-0.0133	0.0129		-0.0162	0.0130	
age60_69	0.0457	0.0190	**	0.0296	0.0197	
age70_79	0.0750	0.0259	***	0.0232	0.0323	
age80	0.0858	0.0459	*	0.0721	0.0487	
jobexperience40	0.0493	0.0119	***	0.0457	0.0117	***
eqincomeQ2	-0.0387	0.0134	***	-0.0363	0.0134	***
eqincomeQ3	-0.0370	0.0140	***	-0.0322	0.0142	**
eqincomeQ4	-0.0335	0.015	**	-0.0255	0.0153	*
missingincome_bl	0.0482	0.0139	***	0.0504	0.0139	***
missingincomeW4	0.0553	0.0175	***	0.0572	0.0175	***
nomissingincomeW4	-0.0546	0.0122	***	-0.0547	0.0122	***
fewbooks	0.0159	0.0114		0.0145	0.0114	
mathskills	-0.0335	0.0101	***	-0.0346	0.0101	***
ruralarea	0.0235	0.0103	**	0.0220	0.0103	**
highschool	-0.0266	0.0123	**	-0.0268	0.0123	**
college	-0.0431	0.0146	***	-0.0419	0.0146	***
public	-0.0264	0.0115	**	-0.0273	0.0116	**
selfemployed	0.0034	0.0149		0.0044	0.0150	
partner	0.0109	0.0132		0.0085	0.0132	
nogstest_bl	0.0171	0.0523		0.0206	0.0526	
dropgstest	0.0792	0.0288	***	0.0853	0.0290	***
increasegstest	-0.0063	0.0586		-0.0127	0.0576	
alonecftest_bl	-0.0649	0.0258	**	-0.0664	0.0258	**
dropalonecftest	0.0187	0.0176		0.0178	0.0176	
increasealonecftest	-0.0377	0.0242		-0.0385	0.0241	
contextcftest_bl	0.1644	0.0956	*	0.1706	0.0960	*
dropcontextcftest	0.1570	0.0335	***	0.1510	0.0333	***
increasecontextcftest	-0.1368	0.0503	***	-0.1401	0.0493	***
poorhealth_bl	-0.0054	0.0317		-0.0067	0.0315	
drophealth	0.0421	0.0228	*	0.0437	0.0228	*
increasehealth	-0.0009	0.0315		0.0021	0.0315	
eurodcat_bl	0.0281	0.0183		0.0327	0.0184	*
dropeurodcat	-0.0266	0.0214		-0.0316	0.0211	
increaseeurodcat	0.0175	0.0154		0.0172	0.0154	
physicalact_bl	-0.0317	0.0132	**	-0.0331	0.0132	**
dropphysicalact	0.0446	0.0144	***	0.0428	0.0143	***
increasephysicalact	-0.0315	0.0155	**	-0.0339	0.0154	**
dailyact_bl	-0.0732	0.0173	***	-0.0767	0.0171	***
dropdailyact	0.0633	0.0421		0.0658	0.0423	
increasedailyact	-0.0614	0.0133	***	-0.0611	0.0133	***
dailycontactchild_bl	-0.0172	0.0130		-0.0128	0.0130	
dropdailycontactchild	0.0108	0.0164		0.0107	0.0164	
increasedailycontactchild	-0.0104	0.0165		-0.0076	0.0166	
SE	0.0079	0.0213		0.0106	0.0214	
DK	0.0209	0.0236		0.0175	0.0235	
NL	0.0047	0.0225		0.0066	0.0226	
BE	-0.0773	0.0177	***	-0.0761	0.0178	***
FR	-0.0983	0.0168	***	-0.0981	0.0168	***
CH	-0.0072	0.0233		-0.0067	0.0234	
AT	0.0304	0.0291		0.0200	0.0289	
ES	-0.0476	0.0234	**	-0.0475	0.0234	**
IT	-0.0826	0.0191	***	-0.0847	0.0191	***
Observations	8262			8262		
Log likelihood	-26054.846			-10748.721		
Amemiya-Lee-Newey minimum chi-sq statistic (p-value)						
	0.7685			0.1909		

*Notes: Significance levels as follows: p-value *** ≤ 0.01, ** ≤ 0.05, * ≤ 0.1.*

Table 14: Summary statistics

	Mean	Std.Err.		%
memory score	8.9	3.24	contextcftest	0.07
yearsinR	5.51	6.85	increasecontextcftest	0.07
sinceER	5.69	6.62	dropcontexcftest	0.03
sinceSR	3.69	5.33	poorhealth	0.03
	%		drophealth	0.05
high decrease	0.25		increasehealth	0.03
lowcognition	0.53		eurodcat	0.18
female	0.43		dropeurodcat	0.09
wave_2	0.25		increaseeurodcat	0.12
age60_69	0.37		physicalact	0.56
age70_79	0.35		dropphysicalact	0.13
age80	0.15		increasephysicalact	0.21
jobexperience40	0.38		dailyact	0.11
eqincomeQ2	0.25		increasedailyact	0.66
eqincomeQ3	0.27		dropdailyact	0.02
eqincomeQ4	0.28		dailycontactchild	0.43
missingincome	0.61		increasedailycontactchild	0.10
missingincomeW4	0.14		dropdailycontactchild	0.15
nomissingincomeW4	0.37		SE	0.12
fewbooks	0.40		DK	0.08
mathskills	0.40		NL	0.09
ruralarea	0.43		BE	0.15
highschool	0.34		FR	0.13
college	0.25		CH	0.08
public	0.30		AT	0.04
selfemployed	0.13		ES	0.07
partner	0.78		IT	0.14
nogstest	0.03		retired	0.62
dropgstest	0.03		fromWtoR	0.15
increasegstest	0.02		eligibleER	0.64
alonecftest	0.84		fromNEtoE_ER	0.21
dropalonecftest	0.09		eligibleSR	0.52
increasealonecftest	0.12		fromNEtoE_SR	0.22