The association between health problems and retirement decisions

Lau, M.K.F. (352283)
Erasmus University Rotterdam

Supervisor: Lumsdaine, R.L.
Co-reader: Kole, H.J.W.G.

August 4, 2017

Abstract

This paper studies the association between health problems and retirement decisions by distinguishing between physical- and mental health and also distinguishing between full- and partial retirement. To account for the endogeneity issue between health and retirement, a two-stage IV approach is taken. In the first-stage, self-perceived health and retirement variables are instrumented. In the second-stage, predicted values of these variables are used to model the association between health and retirement. Consistent with previous studies, the results show that full retirement is associated with worse physical health and partial retirement is associated with better physical health. In terms of magnitude, the marginal effects of full retirement on self-perceived health are larger (difference of 0.1 to 3.4 percentage points in probability) than those of partial retirement in all health categories. Contrary to previous studies, the initial results show that full retirement improves cognitive abilities. In terms of magnitude, partial retirement has larger marginal effects (difference of 0.002 to 0.050 in coefficient) on fluency and numeracy scores than full retirement but smaller effects (difference of 0.020 to 0.068 in coefficient) on memory scores. Fortunately, after including an interaction term between age and retirement, this research finds evidence (contrary to the initial results) that full retirement and partial retirement are associated with declining cognitive measures for people of different ages.

Keywords: Self-perceived health; Cognition; Endogeneity; Instrumental variables

Contents

1 Introduction

Population ageing has put the sustainability of pension and public heath systems under pressure in many countries. To deal with this issue, many countries consider the approach of raising the retirement ages. This makes understanding what effect an individual's labor force exit has on their health conditions crucial to the successful implementation of the policy reforms. For some people, retirement may be a positive. After decades of daily work that is taxing for their bodies and minds, they get to relax. For others, retirement could be negative period filled with health problems and limitations.

This paper aims to describe the association between health problems and exit from the labor market by considering both physical and mental health problems. This paper aims to answer the main research question:

- How do the retirement decisions of labor market participants of age 50 and more affect their physical and mental health?

Furthermore, this paper is distinguishing between partial and full retirement decisions. This paper plans to answer the main question by considering the two following sub questions:

- How are the partial and full retirement decisions associated with the physical health problems?
- How are the partial and full retirement decisions associated with the mental health problems?

The analysis is conducted using data from the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE is a multidisciplinary and cross-national longitudinal database of micro data on health, socio-economic status and social and family networks of approximately 123,000 individuals aged 50 or older (more than 293,000 interviews). As of 2016, SHARE has covered 27 European countries and Israel. The SHARE dataset contains a plethora of variables. This paper focuses on the information of physical and mental health conditions as well as the retirement status of individuals aged 50 or older.

1.1 Literature review

Many previous studies have analyzed the relationship between health condition and retirement and have found strong association between health and individual retirement behavior (See Alavinia and Burdorf (2008), Gannon (2009), Coe and Zamarro (2011) and Pagán (2012)). Results from these studies indicate that health condition does seem to influence retirement decisions and that health condition and retirement are possibly endogenously related. Bad health may cause individuals to consider retiring, but physically demanding work may also cause health to deteriorate. In addition to the endogeneity issue, another question that arises is: "what kind of health measures should be used?" According to Kerkhofs, Lindeboom and Theeuwes (1999), health has an effect on the retirement decision but the size of the effect depends crucially on the health measure used. They conclude that subjective health measures overstate the effect of health on retirement and that endogeneity of health suppresses the health effect. The subjective measures of self-perceived health seem to be the commonest measure of health used in related studies (See Kerkhofs, Lindeboom and Theeuwes (1999) and McGarry (2004)). Hagan, Jones and Rice (2009) deal with the potential problems of reverse causality and measurement error in subjective self-perceived health measures by regressing the self-perceived health measures on a set of more objective health indicators and using the predicted values from these regressions as a measure of a latent health stock. In addition to physical health measures, mental health measures (cognitive abilities) also seem to play a role in retirement decisions (See Mazzonna and Peracchi (2012), Coe, von Gaudecker, Lindeboom and Maurer (2012), Rohwedder and Willis (2010)). According to these studies, retirement is associated with an increase in cognitive decline. As there are with physical health measures, the potential problems of reverse causality also exist with cognition and retirement. Declining cognition may cause individuals to consider retiring, but retirement may also cause cognitive skills to decline. Coe and her colleagues deal with this issue by instrumenting for retirement behavior using unexpected early retirement window offers. These offers are a special kind of incentives, like a cash bonus or improved pension benefits offered to older employers to encourage them to leave a firm at a particular time. Coe and colleagues argue that these offerings serve as a valid instrument for retirement behavior, because these offerings clearly affect retirement behavior but they do not affect an individual's cognitive function since firms cannot limit the eligibility of these offerings to specific individuals. The firms can only select broad groups of workers to be eligible for early retirement windows and the firms' power to limit eligibility is bounded due to anti-discrimination laws[1]. Another reason why the early retirement window offers are a valid instrumental variable for retirement decisions is the fact that the timing of these offers is unanticipated in the sense that the workers have no control over the timing of these offerings.

As can be seen above, there have been many studies[2] conducted on the impact of health conditions on retirement decisions. This study chooses to analyze the impact of retirement on health conditions. Furthermore, much of the previous research does not put their focus on two areas. First, the literature pays little attention to analyzing the effects of health at the intensive margin. Among workers suffering from comparable health problems, some may want to stop working all together while others may just want to reduce their amount of working hours. In this research a distinction will be made between transitions to partial and full retirement due to health issues. Second, previous studies rarely make a distinction between different types of health conditions. Mental health issues may affect job performance differently than physical issues. As SHARE also provides information on the cognitive abilities of their respondents, this study will focus on analysis of the effects that labor market exit through partial and full retirement have on physical health as well as mental health. As the previous studies have done, this paper also takes the endogeneity of the physical health conditions and retirement into account in the analysis. In order to deal with the potential problems of reverse causality and measurement error in subjective self-perceived health measures, this paper adopts the method of Hagan, Jones and Rice (2009). As SHARE provides self-perceived health measures as well as other more objective health measures such as body mass index and grip strength measures, it is viable to adopt this approach. As for the potential problems of reverse causality between

[1] One example is the Americans with Disabilities Act (ADA). ADA prohibits discrimination in employment on the basis of disabilities and requires that employers reasonably accommodate individuals with disabilities who can otherwise perform a job.

[2] Alavinia and Burdorf (2008), Gannon (2009), Coe and Zamarro (2011), Pagán (2012), Kerkhofs, Lindeboom and Theeuwes (1999), McGarry (2004), Hagan, Jones and Rice (2009), Mazzonna and Peracchi (2012), Coe, von Gaudecker, Lindeboom and Maurer (2012) and Rohwedder and Willis (2010).

cognitive functioning and retirement behavior, while the approach of Coe, von Gaudecker, Lindeboom and Maurer (2012) is interesting, it is not viable with the SHARE dataset since information on offerings of early retirement windows is not available. Instead, this paper attempts to instrument for the retirement decisions by considering variables such as the number of children, the number of grandchildren, the marital status and receiving pension benefits as instrumental variables. These variables may be viable instruments because they may be related to retirement decisions but do not directly influence cognition. The timing of these variables are also unanticipated in the sense that people are unlikely to have complete control over the timing of having a partner and the birth of a (grand)child.

After the introduction of the research in this section, the following section of this paper explains the data and the construction of the data. In the third section, the dependent variable as well as the explanatory variables are explained in detail. In the fourth part, the methodology is described. Furthermore, the summary statistics of the sample are analyzed. The results of the models discussed in the methodology section are given in the following part. Finally, this paper is concluded with a summary and a discussion.

2 Data

As mentioned before, the data that is considered is obtained from the Survey of Health, Ageing and Retirement in Europe[3]. Questions about different aspects of life are asked of approximately 123,000 individuals aged 50 or older in 27 European countries and Israel. There have been six waves of SHARE conducted, of which the first wave was collected in 2004 and the sixth wave was gathered in 2015. Wave 1, wave 2, wave 4, wave 5 and wave 6 are similar questionnaires. The questionnaire starts with a coverscreen interview on the household level, which is answered by one of the household members. It gathers basic demographic information about each individual currently living in the household. The main survey consists of multiple modules, including the employment and pensions, physical health, mental health and cognitive function modules. Those modules contain the information that is of interest for this study. Wave 3 (SHARELIFE) differs from the other waves of SHARE as it is a retrospective survey that focuses on people's life histories. Since the information of interest for this study is not included in wave 3, data from wave 3 is not used in this research.

2.1 Initial sample

In order to include as much information as well as to have the largest sample possible, this paper treats respondents who have participated in multiple waves of SHARE as different respondents. For example, a respondent who appears in wave 1, wave 2, wave 4, wave 5 and wave 6 has five different observations in the sample and has the information corresponding to each respective wave. In this case, the age of this particular respondent is for example, 50, 52, 55, 58 and 60 in the respective waves. In this way, the respondent who appears in

[3] *This paper uses data from SHARE Waves 1, 2, 3 (SHARELIFE), 4, 5 and 6 (DOIs: 10.6103/SHARE.w1.600, 10.6103/SHARE.w2.600, 10.6103/SHARE.w3.600, 10.6103/SHARE.w4.600, 10.6103/SHARE.w5.600, 10.6103/SHARE.w6.600), see Börsch-Supan et al. (2013) for methodological details. The SHARE data collection has been primarily funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812) and FP7 (SHARE-PREP: N° 211909, SHARE-LEAP: N° 227822, SHARE M4: N° 261982). Additional funding from the German Ministry of Education and Research, 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, OGHA_04-064) and from various national funding sources is gratefully acknowledged (see www.share-project.org).*

five different waves is treated as five different respondents, each with different information included in the sample. Using this method, the initial sample of wave 1, wave 2, wave 4, wave 5 and wave 6 consists of 260,244 person-wave observations with 120,047 unique respondents. In wave 2 20,916 of the 30,434 respondents from wave 1 are carried over, while 16,258 are new to wave 2. The carryovers in the following waves can be seen in table 1.

Table 1: The number of respondents per wave in the initial sample

	Wave 1	**Wave 2**	**Wave 4**	**Wave 5**	**Wave 6**	**Total**
Number of respondents	30,434	37,174	58,184	66,221	68,231	260,244
Carryover from the previous wave		20,916	18,781	39,009	47,523	

2.2 Missing values

As is common with data obtained from a large longitudinal survey, some of the variables contain missing observations. According to the SHARE missing codes, the observations can be missing due to:

- The respondent answering a question with "Don't know"
- The respondent refusing to answer a question
- The respondent answering a question with an implausible value
- The answer to a question by the respondent being deemed as not codable or not yet coded by SHARE
- A question not being applicable to the respondent

In order to keep the sample as large as possible and to obtain reliable results, this paper tries to deal with the problem of missing observations by replacing some of the unavailable observations with sensible assumptions for each variable with missing values. In addition, SHARE also provides imputations that are performed with the multiple imputation methodology for these variables (See Christelis (2011) and De Luca, Celidoni and Trevisan (2015)).

With multiple imputations, five values are generated for each missing value. The five implicat datasets are generated independently of each other. In some cases where sensible replacements cannot be assumed for the missing values, this paper chooses to use the average values over the five imputed values to replace the missing values. According to the SHARE Release Guide 6.0.0 (page 46), since the five values are independently generated from each other, there is no specific reason to prefer one particular implicat to the others. The Release Guide also states that neglecting the uncertainty of the imputed data (by selecting for instance only one of the five available replications) may lead to misleadingly precise estimates. The next section provides the list of variables and their details.

3 Variables

This section details the list of variables that are of interest from the SHARE dataset, discussing the dependent variables and the multitude of control variables.

3.1 Dependent variables - the health variables

The physical health as well as mental health variables are used as the dependent variables in the models to assess the association between health problems and retirement decisions. For physical health, the variable *Self-perceived health* is used because it is thought to capture more physical rather than mental health (See Shields and Shooshtari (2001)). To account for the endogeneity issue, the instrumental variables: *Limitation with activities*, *Chronic diseases*, *BMI* and *Maxgrip* are used for *Self-perceived health.* For mental health, the variables: *Euro-D*, *Verbal Fluency*, *Numeracy* and *Recall memory* are used.

Physical health variables

Self-perceived health (US version)

This variable is based on how each respondent values his or her own health. In wave 1, there are two versions of self-perceived health. The first version uses the scaling from the World Health Organization (WHO version): (1) very good, (2) good, (3) fair, (4) bad and (5) very bad. The second version (US version) uses the scaling: (1) excellent, (2) very good, (3) good, (4) fair and (5) poor. Due to the fact that from wave 2 onward only the US version of self-perceived health exists, it would be ideal to also use this version for respondents from wave 1. However in wave 1, the respondents are asked to assess their health twice with the two types of answer-categories and which type they receive first is randomized. Of the 30,434 respondents in wave 1, 15,192 (49.92 %) of them received the US version of the question first, while the rest of the respondents in wave 1 received the WHO version first. Lumsdaine and Exterkate (2013) find that the question-order may play a role in the responses. Following Lumsdaine and Exterkate (2013), a respondent is defined as word concordant if he or she answered exactly the same words to both the US version and the WHO version of the self-perceived health question (for example, "Good" as a response to both versions). Next, a respondent is defined as numerical concordant if he or she answered both versions using the same numerical value (for example, he or she chose number three in both versions, with three corresponding to “good” in the US version and “fair” in the WHO version). Lastly, a respondent is defined as discordant if he or she gave entirely different (not word or numerical concordant) answers to both versions of the question. The following approach is taken to account for the two versions of self-perceived health: For the respondents who were asked the US version first, their answers to this question are given as their value for this *self-perceived health* variable. For those who were given the WHO version of the question first, those who are word or numerical concordant (about 44 % of wave 1) are given their first response (mapped to the US scale) as their value for *self-perceived health*. The discordant people who were asked the WHO version first (about 6 % of wave 1), their first response (mapped to the US scale) are used. The robustness of this decision will be tested in the robustness section by also analyzing the results if their second response are used instead. Overall, *self-perceived health* contains 981 (0.38 % of the initial sample) missing values across all waves.

Instrumental variables for physical health

Limitation with activities

This variable gives the number of limitations with activities of daily living each respondent has. The respondents are asked whether they have any difficulty doing several of the everyday activities. These activities include:

- Dressing, including putting on shoes and socks
- Walking across a room
- Bathing or showering
- Eating, such as cutting up your food
- Getting in or out of bed
- Using the toilet, including getting up or down.

This variable has 976 (0.38%) missing values.

Chronic diseases

This variable counts the number of chronic diseases each respondent has. Some examples of chronic diseases are: heart issues (heart attack or congestive heart failure), diabetes, cancer, hip fracture and others. This variable has 1,080 (0.41 %) missing values.

Body Mass Index (BMI)

This variable gives the body mass index (BMI) of each respondent. Using the weight and height of each respondent, the BMI is calculated as follow: $BMI = \frac{(weight\ in\ kilograms)}{(height\ in\ meters)^2}$. There are 9,042 (3.47%) missing observations due to either a missing value for the weight or height.

Maxgrip

In each wave of SHARE, the respondents are asked whether they are willing to take part in a gripping exercise in order to assess the strength of their hands. If they are willing, then the handgrip strength of both of their hands are measured twice with the aid of a dynamometer. The variable *Maxgrip* gives the maximum value of the four measures. For this variable, there are 17,733 (9.23%) missing values. The large number of missing values is due to the respondents' unwillingness to take part in the exercise or in some cases the respondents are unable to complete the test.

Mental health variables

Depressed

The EURO-D depression scale is constructed by summing the 0 or 1 (no or yes) scores for 12 depression related questions in the survey as a composite index of twelve items: depressed mood, pessimism, suicidality, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment and tearfulness. The scale ranges from 0 (not depressed) to 12 (very depressed). This variable has 9,711 (3.73%) missing values, of which 3,302 (1.27%)

are missing due to a missing value in one or more of the 12 depression related questions and the rest are missing due to missing values in all twelve of the questions. The 3,302 missing values are replaced by the sum of the remaining available questions and re-scaled accordingly. For example if only one question is missing, and the sum over the 11 questions is 11, the re-scaled EURO-D score is 10 ($\frac{11}{12} \times 11$). Next, a binary variable *Depressed* is created that has value one if a respondent's EURO-D score is higher than three and zero otherwise. According to Denny (2011), an individual with a Euro-D score greater than three is considered to be at risk of depression.

Verbal fluency

In the cognitive function module of each wave of the SHARE survey, the respondents are asked to name as many animals as they can within a minute. Their scores for *verbal fluency* are equal to the number of animals that they have named. This variable has 9,422 (3.62%) missing values.

Numeracy

In addition to fluency, the respondents are also asked four numeracy related questions (numeracy1). In these questions the respondents are asked to perform some calculations. On the basis of these four questions, Dewey and Prince (2005) construct a numeracy indicator, which has value 1, 2, 3, 4 or 5. The exact wordings of these four questions are:

1. If the chance of getting a disease is 10 per cent, how many people out of one thousand would be expected to get the disease?

2. In a sale, a shop is selling all items at half price. Before the sale a sofa costs 300 euro. How much will it cost in the sale?

3. A second hand car dealer is selling a car for 6,000 euro. This is two-thirds of what it costs new. How much did the car cost new?

4. Let's say you have 2,000 euro in a saving account. The account earns ten per cent interest each year. How much would you have in the account at the end two years?

The method that Dewey and Prince have applied works as follow: If a person answered question (1) correctly then he/she was asked question (3) and if he/she answers correctly again he/she was asked question (4). If a respondent only answered question (1) correctly then he/she is given a score of 3. If he/she answered question (3) correctly but not question (4), then a score of 4 is assigned to him/her. If he/she answered question (4) correctly, he/she would get a score of 5. On the other hand if he/she answers question (1) incorrectly then he/she is directed to question (2). If the respondent answers question (2) correctly, it results in a score of 2. If the incorrect answer is given to question (2), it results in a score of 1. This *numeracy1* variable has 121,527 (46.7%) missing values. The high amount of missing values of *numeracy1* (mostly in waves 4,5 and 6) is due to the fact that starting in wave 4 only the new respondents receive the questions above. Respondents who participated in an earlier wave, only receive a new set of five numeracy questions (numeracy2). The new respondents also receive this new set of questions. These new questions are related to subtraction rather than related to percentage in the first numeracy test. Since *numeracy2* is newly introduced in wave 4, the respondents from wave 1 and wave 2 who do not appear again in future waves will have no values for this variable. In order to make sure that most respondents have a score for numeracy test, the score of the second numeracy test is used as a proxy for those with missing values for the first numeracy test. This choice comes from the assumption that numeracy scores from both tests are most likely related and for simplicity a linear relationship is assumed. First *numeracy1* is regressed on *numeracy2*, then using the coefficients of the regression it is possible to replace 116,309 of the 121,527 missing values of the first numeracy score with the predicted values of *numeracy1*.

Recall memory

For recall memory, the respondents are asked to carefully listen to a list of 10 words that the interviewer would read to them. Then they are immediately asked to recall the list of words. The number of words that they can recall is their score of the *immediate recall memory*. The interviewer would proceed with the interview and after a while he or she would ask the respondents to recall the list of words once again. The number of words that they can recall this time is their score of *delayed recall memory*. The variable *Immediate recall memory* has 9,612 (3.69%) missing values and the variable *Delayed recall memory* contains 10,812 (4.15%) missing observations.

3.2 Explanatory variables

The explanatory variables that this study uses include the retirement variables (*Partially and Fully*) with the instrumental variables: *number of children*, *number of grandchildren*, *Pension* and *Marital status* and the demographic variables: *Age*, *Male*, *Education*, *Income*, *Wealth*, *country dummies* and *wave dummies*.

The retirement variables

Since the focus of this study is on analyzing the association between health problems and retirement decisions of older adults, this paper distinguishes between partial retirement decisions and full retirement decisions. Both retirement decisions are modeled as explanatory variables in regression models to assess the impact that retirement decisions have on physical health as well as mental health. The variable for the *partial retirement decision* is a binary variable, which takes the value 1 if the respondent is currently partially retired and the value 0 otherwise. Likewise, the variable for *full retirement decision* is also a binary variable, which takes the value 1 if the respondent is currently fully retired and the value 0 otherwise. In the employment and pensions module of SHARE, the respondents are asked to describe their current job situation. The choices to this questions are: (1) Retired (retired from own work, including semi-retired, partially retired, early retired, pre-retired), (2) Employed or self-employed (including working for family business), (3) Unemployed, (4) Permanently sick or disabled, (5) Homemaker or (6) other. Since SHARE does not distinguish between fully retired and partially retired, additional information is needed to construct the binary variables. To this end, the working hours per week of the respective respondent is considered. In the employment and pensions module of SHARE, respondents who self reported that they are currently employed and respondents who have done any paid work, are asked how many hours a week they usually work in their main job and second job. Using this information, a variable for working hours of the respondents can be created as the sum of the number of hours over the main and second jobs. A respondent is considered to be partially retired if his or her total number of working hours is less than 20 hours per week. There are no missing values for these variables.

The number of children

This variable gives the number of children of the respondents including all natural children, fostered, adopted and stepchildren. There are no missing values for this variable.

The number of grandchildren

This variable gives the number of grandchildren of the respondents including grandchildren of spouse or partner from previous relationships. There are no missing values for this variable.

Pension

This binary variable has value 1 if the respondent currently receives any pension benefits and 0 otherwise. There are no missing values for this variable.

Marital status

Respondents are asked about their marital status in each wave of SHARE. The six possible answers to this question are: (1) Married and living together with spouse, (2) registered partnership, (3) married, living separated from spouse, (4) never married, (5) divorced, (6) widowed. The respondents are also asked whether they currently live together with their partner. A dummy variable *marital status* is created using this information. This dummy variable takes value 1 if the respondent has a partner or spouse at the time of the interview, which corresponds with the answers (1) to (3) of the question asked about their marital status. On the other hand, the variable has value 0 if the respondent does not have a partner, which corresponds with the answers (4) to (6). There are also no missing values for this variable.

Demographic variables

Age

This variable gives the age of each respondent at the time of the interview. This is determined by using the year of birth of each respondent. This variable has 27 (0.01%) missing values.

Male

This dummy variable *male* takes value 1 if the respondent is male and value 0 if the respondent is female. There are no missing values for this variable.

Education

SHARE contains information about highest obtained educational degree of each respondent. In order to take different education systems across the different countries into account, the 1997 International Standard Classification (provided by SHARE) is used. This variable classifies each respondent's education level into: (0) pre-primary education , (1) primary education , (2) lower secondary education, (3) upper secondary education, (4) post-secondary education, (5) first tertiary education or (6) second tertiary education. This variable has 4,646 (1.79%) missing observations.

Log(Income)

In each wave SHARE has its own generated variables, among which is the *annual household income*. Some modifications are made to this variable. First of all, due to the fact that the household income of wave 1 is given in gross terms and the household incomes of the other waves are given in terms of net household income, the tax rates of the different countries at the time of the interview (obtained from the OECD database) are used to correct for the household income of wave 1. Afterwards, this *income* variable is adjusted for relative purchasing power parity by dividing the incomes with the purchasing power parity-rates of the interview year (provided by SHARE). Finally, in order to obtain income per person, information about the number of household members contributing to the household income is considered. The following is assumed regarding the other household members: the respondents' children younger than the age of 16 and his/her parents older than the age of 70 are unlikely to contribute to the household income. Following Avendano, Aro and Mackenbach (2005), the income is divided by the square root of an adjusted household size (with the number of children younger than age 16 and parents older than age 70 subtracted from the regular household size) to obtain the equivalent disposable income per standard person. The variable *income* has no missing values.

Wealth

The SHARE survey covers a wide range of financial and real assets, from which one can calculate *wealth*. Financial assets include seven categories: bank and other transaction accounts, government and corporate bonds, stocks, mutual funds, individual retirement accounts, contractual savings for housing, and life insurance policies. For each category the respondents are asked whether they hold any assets in the respective category. If so, they are asked to give a value for their total holdings in the respective category. This paper defines total financial wealth as the sum of the seven categories of financial assets listed above. This variable has 87,485 (33.62%) missing observations. This large number of missing values is due to missing value in any one of the seven financial wealth related questions. These unavailable observations are replaced by SHARE imputations.

Country dummies

In order to control for country effects, dummy variables for the country from which the respondents are from are included as explanatory variables. There are respondents from 21 countries in the initial sample.

Table 2: List of the countries in SHARE wave 1-6

No.	**Country**	**Wave 1**	**Wave 2**	**Wave 4**	**Wave 5**	**Wave 6**
1	Austria	x	x	x	x	x
2	Germany	x	x	x	x	x
3	Sweden	x	x	x	x	x
4	Netherlands	x	x	x	x	
5	Spain	x	x	x	x	x
6	Italy	x	x	x	x	x
7	France	x	x	x	x	x
8	Denmark	x	x	x	x	x
9	Greece	x	x			x
10	Switzerland	x	x	x	x	x
11	Belgium	x	x	x	x	x
12	Israel	x	x			x
13	Czech Republic		x	x	x	x
14	Poland		x	x		x
15	Ireland		x			
16	Luxembourg				x	x
17	Hungary			x		
18	Portugal			x		x
19	Slovenia			x	x	x
20	Estonia			x	x	x
21	Croatia					x

Starting in wave 1, there are 12 countries in the sample. Wave 2 has added respondents from Czech Republic, Poland and Ireland. Wave 4 has introduced Hungary, Portugal, Slovenia and Estonia. Wave 5 has added Luxembourg, while wave 6 has added Croatia.

Wave dummies

In order to absorb wave specific effects, *wave dummies* are also added as explanatory variables.

3.3 Analysis sample

Starting from the initial sample, some adjustments are made to deal with the missing observations mentioned above. There are two types of respondents in the sample: one-time respondents (those who only appear once) and longitudinal respondents (those who appear in multiple waves). Since it is desired to retain as many longitudinal respondents as possible for the analysis, the adjustments are first applied to the one-time respondents (see table 3). First, respondents younger than age 50 are dropped from the initial sample, since SHARE is only representative for older respondents (age 50 or older). Next, the respondents with missing observations in the dependent variables are dropped. Since the missing observations may overlap between the variables, the following observation dropping procedure is used: first, the respondents with the most missing observations in a certain physical health variable are dropped, then the respondents with the next highest number of missing values in the health variable are dropped. After dropping respondents with unavailable observations in the physical health variables, the same procedure is applied to the mental health variables and then the independent (demographic) variables. After these first adjustments, the sample consists of 251,344 person-wave observations with 111,147 unique respondents. Of these 251,344 observations, there are still 28,427 person-wave observations with 19,248 unique people containing missing observations across the different variables. Of the 19,248 unique people, 2952 of them are missing a middle wave (For example someone who appears in wave 1, 2 and 4, has missing value for a certain variable in wave 2). About 60% of the 19,248 unique respondents have missing values in physical health (mainly in grip strength measures), about 30% in cognition (mainly in recall memory) and 10% in demographics (mainly in education). On grip strength, using measures from other waves, it is possible to interpolate for 2438 of them. Linear interpolation is used since the SHARE imputations do not contain imputed values for these missing values in grip strength. The remaining respondents with missing values in grip strength measures are omitted from the sample. For the respondents with missing values in the remaining variables (mainly in recall memory and education), SHARE imputations are used as replacements. After these adjustments, the final analysis sample contains 235,454 person-wave observations, which is 90.47% of the initial sample and it has a total of 109,388 (91.12% of the initial sample) unique respondents over the five waves of SHARE.

Table 3: Analysis sample

	Wave 1	**Wave 2**	**Wave 4**	**Wave 5**	**Wave 6**
Initial sample	30,434	37,174	58,184	66,221	68,231
One-time resp.					
Dropping resp. with age <50	30,165	36,937	56,795	66,004	67,764
Dropping resp. with missing obs. in PH var.	29,024	35,835	56,025	64,930	66,351
Dropping resp. with missing obs. in MH var.	28,929	35,779	55,907	64,819	66,212
Dropping resp. with missing obs. in Dem. Var.	28,869	35,740	55,883	64,772	66,080
Longitudinal resp.					
Dropping resp. with missing obs. in maxgrip	27,341	33,639	52,671	60,781	61,113
Carryover from previous wave		18,835	17,492	35,424	42,729
Cumulative sample	27,341	60,980	113,651	174,432	**235,545**

4 Methodology

4.1 First-stage models

As mentioned before, this study will focus on analysis of the effects that labor market exits through partial and full retirement have on physical health and mental health. However, there is one issue with analyzing the contemporaneous effect of retirement behavior on health condition. That is, that retirement may deteriorate or improve health, which means that retirement and health conditions are potentially endogenous in the analysis of the effect of retirement behavior on health conditions. A two-stage instrumental variables approach is taken to tackle this issue of endogeneity. To account for the endogeneity of the physical health condition, the approach in Hagan, Jones and Rice (2009) is used. First, the self-perceived health measures are regressed on a set of health indicators such as chronic diseases, activity limitations and other demographics. Next, the retirement behaviors are instrumented in two equations following Coe et all.(2012). These nine equations form the first-stage models:

$$SPH_i^* = x_i'\alpha + \nu_i, \quad \nu_i \sim N(0, \sigma^2) \tag{1}$$

$$SPH_i = \begin{cases} 1 \;\; if & SPH_i^* \leq \mu_1, \\ 2 \;\; if \;\; \mu_1 < & SPH_i^* \leq \mu_2, \\ 3 \;\; if \;\; \mu_2 < & SPH_i^* \leq \mu_3, \\ 4 \;\; if \;\; \mu_3 < & SPH_i^* \leq \mu_4, \\ 5 \;\; if & SPH_i^* > \mu_4 \end{cases} \tag{2}$$

$$P_{ij} = P(SPH_i = j|x_i) = \begin{cases} \Phi(\frac{\mu_1 - x_i'\alpha}{\sigma}) & for \ \ j = 1 \\ \Phi(\frac{\mu_2 - x_i'\alpha}{\sigma}) - \Phi(\frac{\mu_1 - x_i'\alpha}{\sigma}) & for \ \ j = 2 \\ \Phi(\frac{\mu_3 - x_i'\alpha}{\sigma}) - \Phi(\frac{\mu_2 - x_i'\alpha}{\sigma}) & for \ \ j = 3 \\ \Phi(\frac{\mu_4 - x_i'\alpha}{\sigma}) - \Phi(\frac{\mu_3 - x_i'\alpha}{\sigma}) & for \ \ j = 4 \\ 1 - \Phi(\frac{\mu_4 - x_i'\alpha}{\sigma}) & for \ \ j = 5 \end{cases} \tag{3}$$

$$PR_i^* = x_i'\beta + \epsilon_i, \ \ \epsilon_i \sim N(0, \sigma^2) \tag{4}$$

$$PR_i = \begin{cases} 1 \ \ if \ \ PR_i^* > 0 \\ 0 \quad otherwise \end{cases} \tag{5}$$

$$P(PR_i = 1|x_i) = \Phi(\frac{x_i'\beta}{\sigma}) \tag{6}$$

$$FR_i^* = x_i'\gamma + \eta_i, \ \ \eta_i \sim N(0, \sigma^2) \tag{7}$$

$$FR_i = \begin{cases} 1 \ \ if \ \ FR_i^* > 0 \\ 0 \quad otherwise \end{cases} \tag{8}$$

$$P(FR_i = 1|x_i) = \Phi(\frac{x_i'\gamma}{\sigma}) \tag{9}$$

That is, in equation (1) the self-perceived health (SPH) measure is determined by explanatory variables such as other health indicators (the instrumental variables) and demographic variables using an ordered Probit model. This model is the model of choice due to the fact that the *self-perceived health* variable is observed in the data as a categorical variable with five possible values: 'excellent', 'very good', 'good', 'fair' and 'poor'. The ordered Probit model uses a latent variable approach, where the dependent variable has the potential values (more than two) with some natural ordering. In this case, the latent SPH_i^* is observed in discrete form through the censoring mechanism given in equation

(2), where $\mu_1, ..., \mu_4$ are the cut points (thresholds). The probability that each individual i chooses alternative j is given in equation (3), where $\Phi(.)$ is the cumulative standard normal distribution. Furthermore, the ordered Probit model assumes normal distribution for the error term ν and makes use of maximum likelihood estimation to estimate the coefficients $\boldsymbol{\alpha}$.

An approach similar to instrumenting the self-perceived health is taken to instrument for the retirement behaviors to account for the endogeneity between retirement and the mental health condition. In two regression equations (equation (4) and (7)) partial and full retirement decisions are determined by instrumental variables such as the number of children, the number of grandchildren, the marital status and pension among other control variables using Probit models, where PR and FR are the dummy variables for partial and full retirement decisions. The Probit model like the ordered Probit model uses a latent variable approach, but the dependent variable can only take two possible values (retired or not retired). The underlying mechanisms of the two Probit models are given in equation (5) and (8). The probability that individual i is partially or fully retired is given in equation (6) and (9), respectively. Furthermore, the Probit models also assume normal distributions for the error terms (ϵ in equation (4) and η in equation (7)) and make use of maximum likelihood estimation to estimate the coefficients (β in equation (4) and γ in equation (7)).

4.2 Second-stage models

In the second stage, physical health and mental health measures are interacted with retirement behaviors among other control variables. To account for the aforementioned reverse causality issue, values of self-perceived health and the retirement behaviors predicted from the respective first-stage models must be used in place of the observed values from the sample. To obtain predicted values of self-perceived health from the ordered Probit model, the cut points ($\mu_1, ..., \mu_4$) obtained from the regression are used as the thresholds (as in equation (2)). To obtain predicted values for the retirement decisions, the predicted probabilities ($P(PR_i = 1)$ and $P(FR_i = 1)$) obtained from the Probit regressions are used.

A threshold of a predicted probability of larger than 0.5 is used as an indicator of predicted full retirement. For partial retirement, a threshold of 0.05 is used. These thresholds are chosen due to the fact that the fractions of fully and partially retired people in the sample are roughly 50% and 5%, respectively. The robustness of the choice of these thresholds will be checked in the robustness section. The second-stage models can be illustrated in several equations:

$$SPH_i^* = x_i'\delta + \phi_i, \quad \phi_i \sim N(0, \sigma^2) \tag{10}$$

$$SPH_i = \begin{cases} 1 \ \ if \quad\quad\quad SPH_i^* \leq \mu_1, \\ 2 \ \ if \ \ \mu_1 < SPH_i^* \leq \mu_2, \\ 3 \ \ if \ \ \mu_2 < SPH_i^* \leq \mu_3, \\ 4 \ \ if \ \ \mu_3 < SPH_i^* \leq \mu_4, \\ 5 \ \ if \quad\quad\quad SPH_i^* > \mu_4 \end{cases} \tag{11}$$

$$P_{ij} = P(SPH_i = j|x_i) = \begin{cases} \Phi(\frac{\mu_1 - x_i'\delta}{\sigma}) & for \ \ j = 1 \\ \Phi(\frac{\mu_2 - x_i'\delta}{\sigma}) - \Phi(\frac{\mu_1 - x_i'\delta}{\sigma}) & for \ \ j = 2 \\ \Phi(\frac{\mu_3 - x_i'\delta}{\sigma}) - \Phi(\frac{\mu_2 - x_i'\delta}{\sigma}) & for \ \ j = 3 \\ \Phi(\frac{\mu_4 - x_i'\delta}{\sigma}) - \Phi(\frac{\mu_3 - x_i'\delta}{\sigma}) & for \ \ j = 4 \\ 1 - \Phi(\frac{\mu_4 - x_i'\delta}{\sigma}) & for \ \ j = 5 \end{cases} \tag{12}$$

$$Depr_i^* = x_i'\theta + \kappa_i, \quad \kappa_i \sim N(0, \sigma^2) \tag{13}$$

$$Depr_i = \begin{cases} 1 \ \ if \ \ Depr_i^* > 0 \\ 0 \quad otherwise \end{cases} \tag{14}$$

$$P(Depr_i = 1|x_i) = \Phi(\frac{x_i'\theta}{\sigma}) \tag{15}$$

$$Num_i^* = x_i'\lambda + \xi_i, \;\; \xi_i \sim N(0, \sigma^2) \tag{16}$$

$$Num_i = \begin{cases} 1 \;\; if & Num_i^* \leq \mu_1, \\ 2 \;\; if & \mu_1 < Num_i^* \leq \mu_2, \\ 3 \;\; if & \mu_2 < Num_i^* \leq \mu_3, \\ 4 \;\; if & \mu_3 < Num_i^* \leq \mu_4, \\ 5 \;\; if & Num_i^* > \mu_4 \end{cases} \tag{17}$$

$$P_{ij} = P(Num_i = j|x_i) = \begin{cases} \Phi(\frac{\mu_1 - x_i'\lambda}{\sigma}) & for \;\; j = 1 \\ \Phi(\frac{\mu_2 - x_i'\lambda}{\sigma}) - \Phi(\frac{\mu_1 - x_i'\lambda}{\sigma}) & for \;\; j = 2 \\ \Phi(\frac{\mu_3 - x_i'\lambda}{\sigma}) - \Phi(\frac{\mu_2 - x_i'\lambda}{\sigma}) & for \;\; j = 3 \\ \Phi(\frac{\mu_4 - x_i'\lambda}{\sigma}) - \Phi(\frac{\mu_3 - x_i'\lambda}{\sigma}) & for \;\; j = 4 \\ 1 - \Phi(\frac{\mu_4 - x_i'\lambda}{\sigma}) & for \;\; j = 5 \end{cases} \tag{18}$$

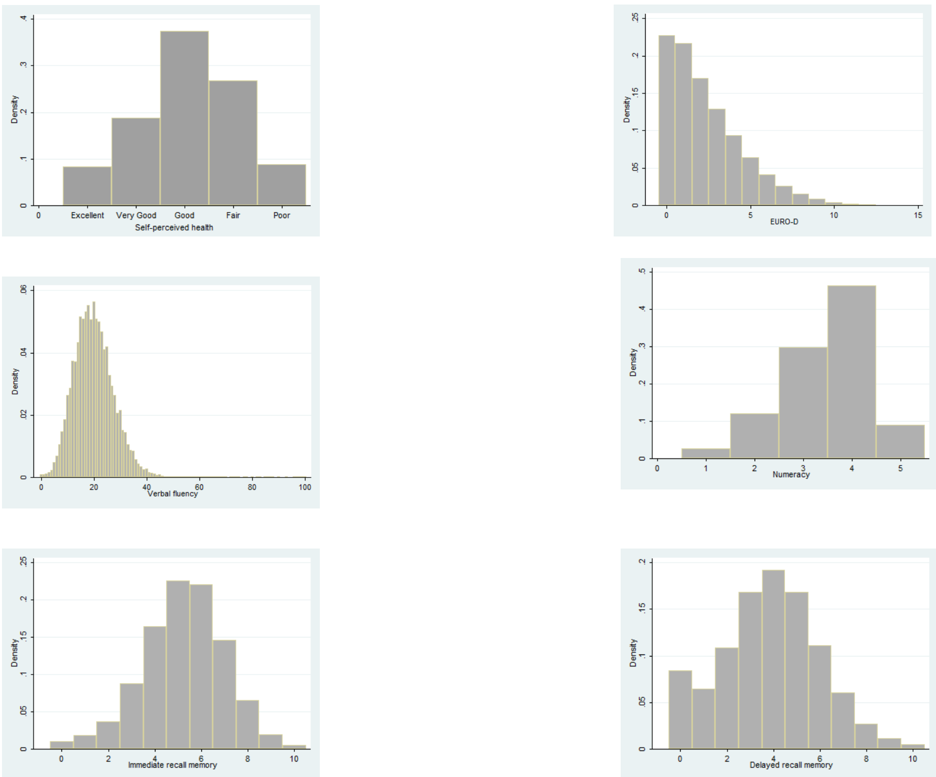
$$VF_i = x_i'\tau + \pi_i \tag{19}$$

$$IRM_i = x_i'\zeta + \psi_i \tag{20}$$

$$DRM_i = x_i'\omega + \chi_i \tag{21}$$

That is, the self-perceived health (predicted from the first-stage) and mental health conditions are used as dependent variables to interact with explanatory variables including retirement behaviors (predicted from the first-stage) among other control variables. For self-perceived health (equation (10)-(12)), an ordered probit model is used. For mental health, first *Depressed* (Depr) is modeled in a Probit model (equation (13)-(15)) to interact with the retirement behaviors. Due to the ordinal nature of the numeracy measures, numeracy scores (Num) are modeled in an ordered Probit model (equation (16)-(18)). Following Schneeweis, Skirbekk and Winter-Ebmer (2014), verbal fluency scores (VF), immediate recall memory (IRM) and delayed recall memory (DRM) are treated as continuous variables and ordinary least squares (equation (19), (20) and (21)) is used to interact these mental health measures with retirement behaviors. This is due to the fact that both memory scores and verbal fluency follow approximately normal distributions around their means (See figure 1 below). Using equations[4] (1)-(21), the two sub questions can be answered.

Figure 1: Frequency graphs of the physical and mental health variables



[4]This paper does not use sampling weights to estimate the models. It is mentioned in the SHARE release guide 6.0.0 that they strongly discourage users to rely on sampling design weights for standard analyses of the SHARE data due to the fact that SHARE data are survey data, problems of unit nonresponse causes the sampling design weights to be invalid (the estimator would be biased). Solon, Haider and Wooldridge (2015) discussed the different advantages and disadvantages of using sampling weights. They mentioned that when using weights to correct for endogenous sampling, if the model is correctly specified, the error term is not related to sampling criterion. Sampling weights may not be needed in that case. They mentioned that if one wants to correct for over- or under-sampling of respondents from certain states, one could include state dummies instead of using sampling weights. In our case, country and wave dummies are included to account for over- or under-sampling for certain countries or waves.

5 Summary statistics

Before the models are estimated, the summary statistics and the distribution of the variables in the sample are analyzed.

5.1 Explanatory variables

Table 4 shows the summary statistics of the demographic variables. In the full sample, the average age of the respondents is 66. There are more women (55.6%) than men (44.4%) in the full sample. The education level of the respondents is the upper secondary, on average. Among the people who are fully retired the average age is higher (around 72 years old). On the other hand, this number is much lower (around 62 years of age) among those who are partially retired. The fraction of men among those who are partially retired is lower compared to the other samples. Furthermore, their education level, income and wealth are higher on average. These may explain their willingness to be working part-time rather than fully retire.

Table 4: Summary statistics of the demographic variables

Variable	**Mean**	**St.dev**	**Min**	**Max**
Full sample				
Age	66.007	9.901	50	104
Male	0.444	0.497	0	1
Education	2.789	1.497	0	6
Log(Income)	8.648	2.094	0	15.64206
Wealth	32190.220	81728.170	0	6773741
Fully Retired				
Age	71.670	7.947	50	104
Male	0.493	0.500	0	1
Education	2.635	1.485	0	6
Log(Income)	8.631	1.930	0	15.64206
Wealth	27618.780	74195.380	0	5579051
Partially Retired				
Age	61.576	7.581	50	93
Male	0.347	0.476	0	1
Education	3.214	1.392	0	6
Log(Income)	8.940	2.094	0	15.29952
Wealth	49810.770	105405.900	0	2520000

Table 5 shows the summary statistics of the instrumental variables for retirement behavior. In the full sample, a respondent has on average 2 children and 3 grandchildren. Around 65% of the respondents receive pension benefits, while 76% have a partner. Almost all of the fully retired respondents (96%) have pension(s). Among people who are partially retired, the number of grandchildren is lower and the fraction that receives pension benefits is also lower. This may be due to the fact that people who are working part-time are younger than an average respondent in the full sample as we saw in table 4.

Table 5: Summary statistics of the instrumental variables for retirement behavior

Variable	**Mean**	**St.dev**	**Min**	**Max**
Full sample				
Number of children	2.188	1.383	0	19
Number of grandchildren	2.611	3.041	0	20
Pension	0.648	0.478	0	1
Marital status	0.762	0.426	0	1
Fully Retired				
Number of children	2.160	1.371	0	19
Number of grandchildren	3.274	3.124	0	20
Pension	0.961	0.193	0	1
Marital status	0.723	0.447	0	1
Partially Retired				
Number of children	2.239	1.352	0	17
Number of grandchildren	2.154	2.791	0	20
Pension	0.533	0.499	0	1
Marital status	0.813	0.390	0	1

The frequency tables of the respondents per wave and per country are given in tables 6 and 7, respectively. In the full sample, the largest fraction of respondents comes from wave 6 and the smallest fraction comes from wave 1. The largest number of fully retired respondents comes from wave 6, while the largest number of partially retirees comes from wave 5. Belgium has the highest number of respondents in the full sample, while Ireland has the lowest number. Among people who are fully retired, Czech Republic has the highest number followed closely by France and Belgium. Among the people who are working part-time, Germany has the highest number followed by Switzerland and Belgium.

Table 6: Frequency table of the respondents per wave

Wave	Full Sample Frequency	Percent	Fully Retired Frequency	Percent	Partially Retired Frequency	Percent
1	27,341	11.61	12,256	10.10	1,580	12.43
2	33,639	14.28	15,734	12.97	1,792	14.10
4	52,671	22.36	28,079	23.14	2,767	21.77
5	60,781	25.80	31,644	26.08	3,582	28.18
6	61,113	25.95	33,616	27.71	2,991	23.53
Total	235,545		121,329		12,712	

Table 7: Frequency table of the respondents per country

Country	Full Sample Frequency	Percent	Fully Retired Frequency	Percent	Partially Retired Frequency	Percent
Austria	13,774	5.85	8,707	7.18	718	5.65
Germany	16,094	6.83	7,773	6.41	1,521	11.97
Sweden	15,398	6.54	8,606	7.09	1,146	9.02
Netherlands	11,739	4.98	4,588	3.78	1,117	8.79
Spain	18,027	7.65	7,532	6.21	491	3.86
Italy	16,804	7.13	8,498	7.00	454	3.57
France	18,543	7.87	10,695	8.81	650	5.11
Denmark	13,889	5.90	6,231	5.14	732	5.76
Greece	9,948	4.22	4,194	3.46	331	2.60
Switzerland	11,682	4.96	4,931	4.06	1,411	11.10
Belgium	22,319	9.48	10,688	8.81	1,256	9.88
Israel	7,784	3.30	3,105	2.56	600	4.72
Czech Republic	17,250	7.32	11,110	9.16	818	6.43
Poland	5,431	2.31	3,229	2.66	136	1.07
Ireland	772	0.33	261	0.22	65	0.51
Luxembourg	2,892	1.23	1,337	1.10	167	1.31
Hungary	2,716	1.15	1,657	1.37	83	0.65
Portugal	3,150	1.34	1,719	1.42	181	1.42
Slovenia	9,003	3.82	6,151	5.07	173	1.36
Estonia	16,172	6.87	9,084	7.49	626	4.92
Croatia	2,158	0.92	1,233	1.02	36	0.28
Total	235,545		121,329		12,712	

5.2 Physical health variables

In table 8, the frequency of each category of self-perceived health in the full sample is given. Good health is the most frequent self-perceived health and Excellent health is the least frequent one in the full sample. This is also the case among the people who are fully retired (51.5% of the full sample). Among respondents who are partially retired (5.39% of the full sample), this is not the case. The least frequent self-perceived health category among partially retired respondents is poor health. Respondents who have poor health are not likely to continue working part-time.

Table 8: Frequency table of Self-perceived health

	Full Sample		Fully Retired		Partially Retired	
	Frequency	**Percent**	**Frequency**	**Percent**	**Frequency**	**Percent**
Excellent	19,524	8.29	6,861	5.65	1,476	11.61
Very Good	44,300	18.81	17,789	14.66	3,113	24.49
Good	87,962	37.34	45,037	37.12	5,023	39.51
Fair	63,373	26.90	38,439	31.68	2,617	20.59
Poor	20,386	8.65	13,203	10.88	483	3.80
Total	235,545		121,329		12,712	

In table 9, the instrumental variables for self-perceived health are summarized. The average respondent has less than one limitation with daily activities, more than one chronic disease, a BMI[5] measure of around 27 and a grip strength[6] measure of around 34.

[5]According to the World Health Organization, an adult over 20 years old is underweight if the BMI is below 18.5. A person has a normal weight if the BMI is between 18.5 and 25. If the BMI is between 25 and 30, then the person is overweight. A BMI over 30 falls into the obese category. So the average respondent is a bit overweight.

[6]Lauretani et al (2003) established cut-off values of grip strength measures of 30 kg and 20 kg for men and women respectively, to describe those with mobility limitations.

Table 9: Summary statistics of the instrumental variables for self-perceived health

Variable	**Mean**	**St.dev**	**Min**	**Max**
Full sample				
Limitation with activities	0.175	0.668	0	6
Chronic diseases	1.675	1.531	0	14
BMI	26.844	4.587	12.753	157.394
Maxgrip	33.677	11.936	1	99
Fully Retired				
Limitation with activities	0.221	0.740	0	6
Chronic diseases	1.988	1.586	0	14
BMI	26.991	4.507	12.754	99.088
Maxgrip	32.101	11.187	1	99
Partially Retired				
Limitation with activities	0.072	0.369	0	6
Chronic diseases	1.360	1.332	0	9
BMI	26.364	4.529	15.427	99.088
Maxgrip	34.372	10.473	1	80

5.3 Mental health variables

In table 10 the summary statics of the mental health variables are given. The average respondent has an EURO-D score of 2, a fluency score of 20, a numeracy score of 3, an immediate recall memory score of 5 and a delayed recall memory score of 4 in the full sample. So the average respondent is not depressed according to the definition by Denny (2011) and is able to answer the first question of the numeracy test correctly. The average respondent is able to name 20 animals, recall 5 out of the list of ten words immediately and 4 after a while. Around 26% of the respondents meet Denny's definition of depressed in the full sample. The cognitive abilities (fluency, numeracy and memory) are all lower among people who are fully retired. According to Salthouse (2009); cognitive abilities tend to decline over time. These measures are however higher among the people who are partially retired. People who are still working part-time are younger than the average respondent in the full sample (as we saw in table 4).

Table 10: Summary statistics of the mental health variables

Variable	**Mean**	**St.dev**	**Min**	**Max**
Full sample				
EURO-D	2.338	2.199	0	12
Depressed	0.255	0.436	0	1
Verbal fluency	20.040	7.573	0	100
Numeracy	3.469	0.912	1	5
Immediate recall memory	5.238	1.774	0	10
Delayed recall memory	3.847	2.132	0	10
Fully Retired				
EURO-D	2.387	2.192	0	12
Depressed	0.263	0.440	0	1
Verbal fluency	19.019	7.223	0	100
Numeracy	3.412	0.894	1	5
Immediate recall memory	4.915	1.751	0	10
Delayed recall memory	3.435	2.084	0	10
Partially Retired				
EURO-D	2.136	2.005	0	12
Depressed	0.220	0.414	0	1
Verbal fluency	22.330	7.232	0	99
Numeracy	3.638	0.830	1	5
Immediate recall memory	5.819	1.600	0	10
Delayed recall memory	4.593	1.993	0	10

6 Results

In this section, the results of the first-stage models and second-stage models are given[7]. This section focuses on analyzing the effects of the instrumental variables, the retirement variables and the other demographics on the dependent variables. The marginal effects of the explanatory variables: *BMI*, *maxgrip*, *log(income)* and *wealth* are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit.

6.1 First-stage models

In the first-stage, self-perceived health, full retirement and partial retirement decisions are modeled as a function of instrumental variables and demographic variables in an Ordered Probit model (equation (1)-(3)), a Probit model (equation (4)-(6)) and a Probit model (equation (7)-(9)), respectively.

Focusing on table 11, we can see that the explanatory variables are all statistically significant in explaining self-perceived health at the 1%-level. This means that the instrumental variables are all significant in explaining self-perceived health. Furthermore, the fact that these instrumental variables are very significant when considered together, shows that they are valid instruments for self-perceived health. This is important, because we would like to account for the endogeneity issue in order to analyze the association between health and retirement in the second-stage models.

[7] The coefficients and marginal effects of the country and wave dummies are left out from the tables in this section, since they are not the main focus of this paper. The complete tables including the coefficients and marginal effects of the country and wave dummies can be found in Appendix A

Table 11: Results of the first-stage Ordered Probit model of self-perceived health as a function of the Instrumental variables and Demographic variables

Variable	**Coeff.**	**P(Excellent)**	**P(VeryGood)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Instrumental variables:						
Limitation with activities	0.305 ***	-0.038 ***	-0.043 ***	-0.004 ***	0.052 ***	0.032 ***
	(0.004)	(0.001)	(0.001)	(1.7E-4)	(0.001)	(4.1E-4)
Chronic diseases	0.337 ***	-0.042 ***	-0.047 ***	-0.004 ***	0.058 ***	0.036 ***
	(0.002)	(3E-4)	(2.8E-4)	(1.9E-4)	(3E-4)	(2.3E-4)
BMI	0.018 ***	-0.010 ***	-0.012 ***	-0.001 ***	0.014 ***	0.009 ***
	(0.001)	(3E-4)	(3.3E-4)	(5.6E-5)	(4E-4)	(2.5E-4)
Maxgrip	-0.021 ***	0.031 ***	0.035 ***	0.003 ***	-0.043 ***	-0.027 ***
	(3.2E-4)	(5E-4)	(0.001)	(1.5E-4)	(0.001)	(4.2E-4)
Demographic variables:						
Age	0.002 ***	-2.7E-4 ***	-3E-4 ***	-2.8E-5 ***	3.7E-4 ***	2.3E-4 ***
	(2.7E-4)	(3.4E-5)	(3.8E-5)	(3.7E-6)	(4.6E-5)	(2.9E-5)
Male	0.345 ***	-0.043 ***	-0.048 ***	-0.004 ***	0.059 ***	0.037 ***
	(0.007)	(0.001)	(0.001)	(2.1E-4)	(0.001)	(0.001)
Education	-0.102 ***	0.013 ***	0.014 ***	0.001 ***	-0.017 ***	-0.011 ***
	(0.002)	(2.2E-4)	(2.4E-4)	(6.2E-5)	(2.9E-4)	(1.9E-4)
Log(Income)	-0.030 ***	0.008 ***	0.009 ***	0.001 ***	-0.011 ***	-0.007 ***
	(0.002)	(4.1E-4)	(4.6E-4)	(5.5E-5)	(0.001)	(3.5E-4)
Wealth	-6E-7 ***	0.006 ***	0.007 ***	0.001 ***	-0.008 ***	-0.005***
	(3E-8)	(3E-4)	(3.4E-4)	(4.3E-5)	(4.1E-4)	(2.6E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-1.479 ***					
	(0.033)					
Cut point 2	-0.519 ***					
	(0.033)					
Cut point 3	0.771 ***					
	(0.033)					
Cut point 4	2.141 ***					
	(0.033)					
F-tests:	Chi-squared	p-value		Observations	235,545	
Instrumental variables	59,349.780	0.000		LogL	-28,4847.56	
Country dummies	19,403.480	0.000		LRχ^2(33)	115,043.89	
Wave dummies	253.890	0.000		Psuedo R^2	0.168	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: BMI, maxgrip, log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

In addition to the coefficients, the average marginal effects are also given in table 11. These effects are calculated by taking the average over the marginal effects evaluated at the observed values of the explanatory variables for each person in the estimation sample[8]. The marginal effects of the explanatory variables are the change in the probability of observing each self-perceived health category, if the independent variable changes by one unit or one standard deviation, while all the other variables are held constant.

With each additional limitation in activities, the probability of reporting excellent health decreases on average by 3.8 percentage points, while it increases the probability of poor health by 3.2 percentage points. Each additional chronic disease makes it 4.2 percentage points less likely that someone reports excellent health, while the opposite effect is found on poor health (3.6 percentage points more likely). Similar to our approach, Disney, Emmerson and Wakefield (2006) used an ordered Probit model to estimate self–reported health status as a function of objective health measures and individual characteristics and found that difficulties with activities and diseases increase the probability of reporting worse self-perceived health, which are consistent with the signs of our findings. Due to the fact that there are only coefficients provided in Disney et al (2006), it is not possible to compare the magnitude of the marginal effects.

One standard deviation increase in BMI (an increase of 4.587 kg/m^2) decreases the probability of excellent health by 1.0 percentage points, while the opposite effect is slightly smaller (0.9 percentage points) on poor health. One standard deviation increase in grip strength (an increase of 11.936 kg) increases the probability of reporting excellent health by 3.1 percentage points, while it decreases the case of poor health by 2.7 percentage points. Kanagae et al (2006) dichotomized self-rated health into good ("excellent" and "good") and poor ("fair" and "poor"). They found that higher BMI and weaker grip strength increase the probability of poor self-perceived health, which is consistent with the signs of our findings. In terms of magnitude, they found that the risk of poor health increases by

[8] One could also take the marginal effects at the mean of the explanatory variables. For a dichotomous variable the marginal effect is typically computed when the variable is equal to 1 minus the effect when it is equal to 0, with the other variables evaluated at their mean. However, due to multiple dummy variables in the models, this approach is not chosen.

1.07 fold (marginal effect of 7 percentage point) with every 5 kg/m^2 increase in BMI, which is larger than the 2.3 percentage points increase (computed as the sum of the 1.4 percentage point increase from the fair health category and the 0.9 percentage point increase from poor health in table 11) with every 4.6 kg/m^2 increase in our case. In addition, Kanagae et al (2006) found that the risk of poor health decreases by 1.72 fold (marginal effect of 72 percentage point) with every 10 kg increase in grip strength, which is larger than the 7 percentage points decrease (computed as the sum of the 4.3 percentage point decrease from the fair health category and the 2.7 percentage point decrease from poor health in table 11) with every 11.9 kg increase in our case. The difference in magnitudes may due to the fact that Kanagae et al (2006) used logistic regression analysis with a sample of 294 respondents consisting of Japanese women of age 65 or more, which is different than ours (our sample consists of 235,545 respondents from 20 European countries and Israel).

Looking at the instrumental variables in table 12, all are statistically significant in explaining full retirement decisions. On partial retirement, this is not the case. The number of grandchildren is insignificant in explaining partial retirement decisions. Lumsdaine and Vermeer (2015) examined the interaction between women caring for grandchildren and retirement decision. They found that retirement-age women who are both working and caring for grandchildren, are more likely to give up caring responsibilities than their work responsibilities. This result is in line with their finding. Despite the insignificance of the number of grandchildren on partial retirement, the instrumental variables are jointly significant in both models. This indicates that they are valid instruments for retirement decisions.

Having pension benefits and having a partner increase the probability of being retired (fully or partially), all things being equal, with bigger effects on full retirement (38.2 and 8.7 percentage points, respectively) than on partial retirement (1.1 and 0.4 percentage points). These findings are directionally in line with the findings in Schirle (2010), which found positive influences of pension wealth and partner on retirement. Schirle found that for each $10,000 increase in pension wealth, the probability of retirement increases by 1.8 percentage points. Unfortunately, we cannot compare this number to the marginal effect in our case, since pension is included as a dichotomous variable in our model rather than a continuous (monetary) variable in Schirle's. Furthermore, Schirle found that having a partner increases the probability of retiring by 5.3 percentage points, which is smaller than the marginal effect (8.3 percentage points increase) in our case. Another difference between this paper and Schirle's research is the fact that Schirle used a sample of 25,810 individuals age 50 and older from the Canadian Survey of Labour and Income Dynamics instead of SHARE. As in Bouman (2015), our results show that each additional child has a negative effect on full retirement (1.7 percentage points) compared to 1.3 percentage points found by Bouman and a smaller positive effect (0.3 percentage points) is observed for each additional grandchild than child compared to 0.4 percentage points found by Bouman, who also used the SHARE dataset (but only wave 1, wave 2 and wave 4).

Regarding demographics, each additional year in age makes it 1.4 percentage points more likely that someone is fully retired and 0.3 percentage points less likely that someone is partially retired. Schirle (2010) found that age has a negative impact on retirement for those who are younger than the age of 60 (marginal effects range from 2.7 to 6.9 percentage points) and that age has a positive impact for those who are 64 and older (marginal effects range from 4.4 to 20 percentage points). This is consistent in terms of signs with our findings since the average age of partially and fully retired respondents in our sample are around 62 and 72, respectively. Male workers are 4.7 percentage points more likely to be fully retired and 2 percentage points less likely to be partially retired.

Contrary to Hochman and Lewin-Epstein (2013), where they found that people with academic education are 28 percentage points less likely to retire as early as they can, we find that each additional level in education makes the probability of being fully retired and partially retired increase by 0.7 and 0.5 percentage points, respectively. The difference between our result and Hochman and Lewin-Epstein's may be due to the fact that they made use of only wave 2 of SHARE and a logistic regression to find the determinants of early retirement preferences. Furthermore, Hochman and Lewin-Epstein compared respondents with academic education (first and second tertiary stage) to those with lower education level, while education ranges from pre-primary education (0) to second tertiary education (6) in our case. Hank and Korbmacher (2013) found that men with medium (upper secondary or post-secondary) education are 10 percentage points more likely to exit the labor force, which compares to 5.4 percentage points (computed as the sum of the 4.7 percentage point increase from male on full retirement and the 0.7 percentage point increase from education in table 12) in our case. On the other hand, Hank and Korbmacher found that women with medium education are 8 percentage points less likely to exit, but in our case this effect is in the opposite direction and equal to 0.7 percentage point. The difference between our research and Hank and Korbmacher's is that they only made use of wave 3 of SHARE[9] and a logistic regression in their analysis.

On wealth and income, opposite effects are found on full retirement and partial retirement. One standard deviation increase in log income (an increase of 2.094, which is equal to 8.117 euros) decreases the probability of being fully retired by 2.7 percentage points but it increases the probability of being partially retired by 0.5 percentage points. One standard deviation increase in wealth (an increase of 81728.170 euros) leads to 0.4 percentage point decrease in the probability of full retirement but it increases the probability of partial retirement by 0.3 percentage point. Due to financial incentives, one could expect that those with high income and wealth to be more willing to continue work part-time rather than to retire fully. Gurley-Calvez and Hill (2011) found that the marginal effect of the log of wages on retirement decisions is equal to -1.3 percentage point, which compares against the -2.7 percentage point (marginal effect of log income on full retirement) found in our case. Hong (2006) used data from the Health and Retirement Study (HRS) and found that one unit increase in log non-pension wealth increases the probability of partial retirement by 0.2 percentage point, whereas the marginal effect of wealth is equal to 0.3 percentage point in our case. Hong did not find significance of the log non-pension wealth variable on full retirement, whereas wealth is negatively significant in our case.

[9]Wave 3 (SHARELIFE) differs from the other waves of SHARE as it is a retrospective survey that focuses on people's life histories.

Table 12: Results of the first-stage Probit models of full retirement decisions and partial retirement decisions as a function of the Instrumental variables and Demographic variables

	Full retirement		Partial retirement	
Variable	**Coeff.**	**P(Fully Retired)**	**Coeff.**	**P(Partially Retired)**
Instrumental variables:				
Number of children	-0.085 ***	-0.017 ***	0.018 ***	0.002 ***
	(0.003)	(0.001)	(0.004)	(4.2E-4)
Number of grandchildren	0.014 ***	0.003 ***	-0.001	-7.2E-5
	(0.002)	(3.2E-4)	(0.002)	(2.2E-4)
Pension	1.913 ***	0.382 ***	0.106 ***	0.011 ***
	(0.010)	(0.001)	(0.011)	(0.001)
Marital status	0.436 ***	0.087 ***	0.040 ***	0.004 ***
	(0.009)	(0.002)	(0.012)	(0.001)
Demographic variables:				
Age	0.068 ***	0.014 ***	-0.028 ***	-0.003 ***
	(4.7E-4)	(8.5E-5)	(0.001)	(7.1E-5)
Male	0.237 ***	0.047 ***	-0.195 ***	-0.020 ***
	(0.007)	(0.001)	(0.009)	(0.001)
Education	0.033 ***	0.007 ***	0.049 ***	0.005 ***
	(0.003)	(0.001)	(0.003)	(3.5E-4)
Log(Income)	-0.066 ***	-0.027 ***	0.021 ***	0.005 ***
	(0.002)	(0.001)	(0.003)	(0.001)
Wealth	-2.1E-7***	-0.004 ***	3.4E-7***	0.003***
	(4.6E-8)	(0.001)	(4.6E-8)	(3.9E-4)
Country dummies	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes
Constant	-5.089 ***		-0.340 ***	
	(0.043)		(0.057)	
F-tests:	Chi-squared	p-value	Chi-squared	p-value
Instrumental variables	39310.800	0.000	124.590	0.000
Country dummies	5478.240	0.000	2412.140	0.000
Wave dummies	245.370	0.000	40.780	0.000
Observations	235545		235545	
LogL	-83884.029		-45915.986	
LRχ^2(33)	158551.82		7114.99	
Psuedo R^2	0.4859		0.0719	

*In the second and fourth column the coefficients are given. The third and fifth column give the marginal effects of the explanatory variables on the probability of being fully retired and partially retired, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

6.2 Second-stage models

In the second-stage, self-perceived health status predicted from the first-stage model (from equation (1)-(3)) is modeled in an ordered Probit model (equation (10)-(12)) as a function of the predicted retirement variables (from equation (4)-(9)) and other demographics. In addition, the mental health variables are also modeled as a function of predicted retirement variables and demographic variables various models. For "Depressed" a Probit model (equation (13)-(15)) is used. Both memory scores and verbal fluency (equation (19)-(21)) are modeled in ordinary least squares models. Lastly, for numeracy measures an ordered Probit model (equation (16)-(18)) is estimated.

The results from table 13 show that all explanatory variables are statistically significant in explaining the self-perceived health. The association between the retirement variables and self-perceived health is of our main interest. Full retirement has the largest marginal effects in the very good (negative 3.8 percentage points) and fair (positive 3.9 percentage points) health categories. Overall, being fully retired leads to a higher probability of having worse self-perceived health. This finding is directionally in line with previous studies (Curl and Townsend (2014), Bender and Theodossiou (2009), Sahlgren (2012)), where full retirement is found to be associated with worse self-perceived health. Bender and Theodossiou found that unemployment increases the probability of worse self-perceived health by 22 percentage points, which compares against the 5.3 percentage point increase (computed as the sum of the 3.9 percentage point increase from full retirement on fair health and the 1.4 percentage point increase from full retirement on poor health in table 13),the marginal effect of full retirement on the probability of fair or poor health in our case. The difference in magnitude may be due to the fact that Bender and Theodossiou used data from the U.S. Health and Retirement Study (HRS) and that unemployment is not necessarily equal to full retirement.

Regarding partial retirement, the largest marginal effects are also in the very good (positive 0.5 percentage points) and fair (negative 0.5 percentage points) health categories. Noticeable is that in terms of magnitude, the marginal effects of partial retirement on self-perceived health are smaller (from 0.1 to 3.4 percentage points) than those of full retirement in all five health categories and the effects go in the opposite direction. Overall, people who are working part-time, have a higher probability of having better self-perceived health. This result is directionally consistent with previous studies (Ettner and Grzywacz (2001), Kantarci and Kolodziej (2016)), where working part time is found to be associated with better self-perceived health. Unfortunately, a thorough comparison of magnitudes is not possible due to the fact that these previous studies do not report marginal effects.

Of the demographic variables, education has the largest marginal effects on self-perceived health. One higher level of education increases the probability of reporting very good health by 4.3 percentage points, while it decreases the likelihood of fair health by 4.5 percentage points. As one would expect, older people are more likely to report worse health. Each additional year in age decreases the likelihood of reporting excellent health by 0.01 percentage points but it improves the likelihood of poor health by 0.2 percentage points. Bonner et al. (2017) made use of data from the 2010 Canadian Community Health Survey (CCHS) and a logistic regression in their analysis. They dichotomized self-perceived health categorized as good health (excellent, very good or good) and less than good health (fair or poor) and found that older people (age 65 or more) with secondary education or higher are around 70 percent more likely to have good self-perceived health than those with lower level of education. This compares against a marginal effect of 6.1 percentage point increase (computed as the sum of the 0.1 percentage point increase from education on excellent health and the 4.3 percentage point increase on very good health and the 1.7 percentage point increase on good health in table 13) for each additional level of education on good self-perceived health or better in our case. Liu and Hummer (2008) also found that the less educated have higher odds of reporting fair/poor health compared to college graduates. Our results are directionally consistent with these previous findings, but due to the fact that there are only coefficients provided in Liu and Hummer (2008), it is not possible to compare the magnitude of the marginal effects.

Regarding the incentive effects (income and wealth), our results indicate that those with higher incentives are more likely to report better health than those with lower incentives.

For example, one standard deviation increase (increase of 81728.170 euros) in wealth increases the likelihood of excellent health by 0.1 percentage points but decreases the likelihood of poor health by 1.1 percentage points. In terms of the signs of these effects, Hajat et al (2010) found a similar result. Hajat et al divided wealth into six quintiles with the sixth quintile being the wealthiest quintile and found that as wealth increased the risk of poor self-perceived declined. Those in the first quintile are 42 percent more likely to be in risk of poor self-perceived health relative to those in the wealthiest quintile. This compares against a 4 percentage point increase (computed as the sum of the 2.9 percentage point on fair health and the 1.1 percentage point on poor health) in the probability of fair or poor self-perceived health for each standard deviation decrease (decrease of 81728.170 euros) in wealth in our case. Furthermore, Wu et al (2013) found that people with low income (less than 2000 Chinese Yuan) are 50 percent more likely to have unhealthy self-perceived health than a healthy one. This compares against a 3.4 percentage point increase (computed as the sum of the 2.5 percentage point increase on fair health and the 0.9 percentage point increase on poor health in table 13) in the probability of fair or poor self-perceived health for each standard deviation decrease (decrease of 2.094, which equals to 8.117 euros) in log income in our case. The differences between these two previous studies and our approach are that they both made use of datasets that are not obtained from SHARE and both made use of logistic regression analyses.

While the signs of the most of the explanatory variables remain the same when we compare the result from the first-stage and the second-stage model of self-perceived health, the only difference is that women are no more likely to report worse health in the second model rather than men in the first model. By accounting for the reverse causality nature of the self-perceived health and retirement, the gender effect is shifted to the opposite direction. In the first-stage model, there may be some correlation between self-perceived health and an unobserved factor. The instrumental variables take this correlation into account and causes the sign of the gender effect to flip.

Table 13: Results of the second-stage Ordered Probit model of the predicted self-perceived health as a function of the predicted retirement variables and Demographic variables

Variable	**Coeff.**	**P(Excellent)**	**P(Very good)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Retirement variables:						
Full retirement	0.240 ***	-0.001 ***	-0.038 ***	-0.015 ***	0.039 ***	0.014 ***
	(0.007)	(4.1E-5)	(0.001)	(4.9E-4)	(0.001)	(4.4E-4)
Partial retirement	-0.030 ***	9E-5***	0.005 ***	0.002 ***	-0.005 ***	-0.002 ***
	(0.008)	(3E-5)	(0.001)	(0.001)	(0.001)	(5E-4)
Demographic variables:						
Age	0.039 ***	-1.1E-4***	-0.006 ***	-0.002 ***	0.006 ***	0.002 ***
	(4.1E-4)	(5.6E-6)	(6.6E-5)	(3.2E-5)	(6.6E-5)	(2.9E-5)
Male	-0.100 ***	3E-4***	0.016 ***	0.006 ***	-0.016 ***	-0.006 ***
	(0.005)	(2.1E-5)	(0.001)	(3.5E-4)	(0.001)	(3.2E-4)
Education	-0.272 ***	0.001 ***	0.043 ***	0.017 ***	-0.045 ***	-0.016 ***
	(0.002)	(4E-5)	(3.2E-4)	(2.1E-4)	(3.3E-4)	(1.7E-4)
Log(Income)	-0.073 ***	4.5E-4***	0.024 ***	0.010 ***	-0.025 ***	-0.009 ***
	(0.002)	(2.4E-5)	(0.001)	(2.5E-4)	(0.001)	(2.3E-4)
Wealth	-2.2E-6***	0.001 ***	0.028 ***	0.011 ***	-0.029 ***	-0.011 ***
	(3.6E-8)	(2.7E-5)	(4.6E-4)	(2.3E-4)	(5E-4)	(1.9E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.741 ***					
	(0.041)					
Cut point 2	0.410 ***					
	(0.033)					
Cut point 3	2.596 ***					
	(0.034)					
Cut point 4	4.179 ***					
	(0.035)					
F-tests:	Chi-square	p-value		Observations	235545	
Retirement variables	1126.630	0.000		LogL	-183,492.540	
Country dummies	49,722.340	0.000		LR$\chi^2(31)$	135,556.800	
Wave dummies	991.870	0.000		Psuedo R^2	0.270	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

The results of the second-stage models to analyze the interaction between mental health measures and the retirement decisions are given in table 14 and 15. Unlike Coe and Zamarro (2011) and Kolodziej and García-Gómez (2017), where full retirement is found to have positive significant impact on depression, full retirement is not significant in explaining the probability of being depressed in our model. However, Riumallo-Herl et al (2014) found no significant association between being retired and the depression scores from SHARE. Our finding is in line with their result. Furthermore, being fully retired is associated with higher fluency, immediate recall memory, delayed memory scores and numeracy scores. Being fully retired increases the probability of having a numeracy score of 4 and 5, by 0.3 and 0.2 percentage points, respectively, while it decreases the probability of having lower numeracy scores by 0.5 percentage points. This result is contrary to the findings in previous studies (Mazzonna and Peracchi (2012), Coe, von Gaudecker, Lindeboom and Maurer (2012), Rohwedder and Willis (2010)), where retirement is found to be associated with cognitive decline. Bingley and Martinello (2011) did find that after accounting for endogeneity of retirement from the labour market, retirement increases word recall by 3.132, which compares against a coefficient of 0.213 (computed as the sum of the 0.110 from immediate recall memory and the 0.103 from delayed recall memory in table 14) in our case since Bingley and Martinello defined word recall as the sum of immediate recall memory and delayed recall memory. Bingley and Martinello made use of data from 2004 across HRS, SHARE and ELSA (English Longitudinal Study of Ageing) to obtain their result.

Overall, in terms of magnitude, partial retirement has larger marginal effects (difference of 0.002 to 0.050 in coefficient) on fluency and numeracy scores than full retirement but smaller effects (difference of 0.020 to 0.068 in coefficient) on memory scores. Compared to Kantarci and Kolodziej (2016), where working part-time is found to decrease depression score (takes values from 0 to 12) by 5.612, being partially retired makes it 2.6 percentage points more likely that a respondent is depressed. While the sign of this effect is in line with Kantarci and Kolodziej's, the magnitudes are not really comparable since Kantarci and Kolodziej modeled depression as a continuous variable (EURO-D scores) in a fixed effects model rather than a binary variable in a Probit model. Being partially retired makes it 1.3 and 1 percentage point less likely of having numeracy scores of 4 and 5, respectively. Positive marginal effects are found for the probability of the lower numeracy scores. Overall, partial retirement decreases the probability of higher numeracy scores, which is in line in terms of sign with the finding in Kantarci and Kolodziej, where working part time decreases the numeracy score by 2.373. Unfortunately, the magnitudes are not comparable since Kantarci and Kolodziej modeled numeracy as a continuous variable (1 to 5) in a fixed effects model rather than a categorical variable in an ordered Probit model in our case. Contrary to Kantarci and Kolodziej, where working part time is found to increase the sum of both memory scores by 10.122, our results indicate that being partially retired decreases the immediate and delayed memory scores together by 0.125 (computed as the sum of 0.090 from immediate recall memory and 0.035 from delayed recall memory in table 14). The difference between this paper and Kantarci and Kolodziej's is that they modeled the memory scores together as the sum of the two scores (values range from 0 to 20) in a fixed effects model rather than separately in OLS models in our case.

Regarding the demographics, older people have lower scores in fluency, memory and numeracy. Each additional year in age decreases the fluency, immediate recall memory and delayed recall memory scores by 0.189, 0.061 and 0.070, respectively. Salthouse (2009) shows that cognitive measures tend to decline as one ages. This finding is in line with the results from our models. In addition to the negative effect of age on the cognitive measures, noticeable is that the marginal effect is of similar (in some cases around 0.001 to 0.020 bigger) magnitude to the retirement marginal effect so this offsets the positive effect of full retirement. Perhaps there is an interaction between these two variables, which will be investigated as a robustness check. Older women are 12.6 percentage points more likely to be depressed according to the results from table 14. Men have lower memory scores (difference of 0.324 and 0.428 in coefficient) than women, but higher numeracy scores (marginal effects range from 1.4 to 5.5 percentage points). A lower level of education increases the probability of being depressed by 2.5 percentage points and it also decreases the scores for fluency, immediate recall memory and delayed memory by 1.168, 0.302 and 0.335, respectively. People with a lower level of education are 7.8 percentage points less likely to have high scores of numeracy (4 or 5) and are 7.8 percentage points more likely to have lower scores (3 or less). Past studies (Bjelland et al (2008), Reuser et al (2010) and Villarreal et al (2015)) have found that lower level of education leads to depression and cognitive impairment. Reuser et al divided education into three groups: less than high school (low), high school graduate and some college (medium), and college graduate and above (high). They found that people with medium or high education are around 35 to 86 percent less likely to have cognitive impairment than those with low education. Villarreal et al found that people with 6 years or less of education are 2.8 times more likely to have cognitive impairment and 2.2 times more likely to be depressed than those with more education. So our results regarding education are in line with the previous findings in terms of the signs. The main difference between this paper and the previous studies is that the past studies made use of different datasets (Reuser et al (2010) used HRS data and Villarreal et al (2015) used data the Panama Aging Research Initiative study) and their results came from different models (Reuser et al estimated a Cox proportional hazard model and Villarreal et al made use of a logistic regression model).

Lastly, one standard deviation increase in log income and wealth (increase of 8.117 euros and 81728.170 euros, respectively) decreases the likelihood of depression by 2.2 and 1.3 percentage points, respectively. Higher (one standard deviation increases) incentives (income and wealth) increases all cognitive scores (fluency, memory and numeracy), with income having larger marginal effects (differences of around 0.070 to 0.292 for fluency and memory and differences of 0.3 to 0.9 percentage points for the different numeracy scores) than wealth in terms of magnitude. These results are in line in terms of signs with the studies (Wagner et al (2012), Dos Santos et al (2012), Lee et al (2006) and Huppert et al (2006)), where low income and wealth are found to be associated with depression and cognitive decline. In terms of magnitudes, Dos Santos et al found that for one unit increase in log income, the risk of depression decreases by around 0.4 percentage points, which compares against 1.1 percentage points decrease (computed as 0.5 multiplied by 1.1 since the marginal effect of log income is given in increases of 2.091 in our model in table 14) in our case. The similar sign of this income effect may be due to the fact that Dos Santos et al also made use of a Probit model to estimate the effect of log income on depression (also a binary variable in their case). They obtained their result by making use of a dataset from three National

Household Sampling Surveys (1998, 2003, and 2008) and their special supplements on the health status of the Brazilian population. Lee et al, who analyzed the effect of income on cognitive decline by dividing income into five groups (the first group is an income of 20,000 US dollars or less and the highest group is 50,000 dollars or more), found that those with the highest income are 40 percent less likely to suffer from cognitive decline than those from the first group. This finding is consistent with the signs of our income marginal effects on the cognitive scores, but unfortunately the magnitudes of these effects are not comparable.

Table 14: Results of the second-stage models of several mental health measures as a function of the predicted retirement variables and Demographic variables

	Depressed (probit)			Fluency (OLS)	Imm. mem. (OLS)	Delayed mem. (OLS)
Variable	**Coeff.**	**P(Depressed)**		**Coeff.**	**Coeff.**	**Coeff.**
Retirement variables:						
Full retirement	-0.001	-3.6E-4		0.170 ***	0.110 ***	0.103 ***
	(0.008)	(0.003)		(0.038)	(0.009)	(0.011)
Partial retirement	0.086 ***	0.026 ***		0.172 ***	-0.090 ***	-0.035***
	(0.010)	(0.003)		(0.044)	(0.011)	(0.013)
Demographic variables:						
Age	0.009 ***	0.003 ***		-0.189 ***	-0.061 ***	-0.070 ***
	(4.6E-4)	(1.4E-4)		(0.002)	(0.001)	(0.001)
Male	-0.407 ***	-0.123 ***		0.007	-0.324 ***	-0.428 ***
	(0.006)	(0.002)		(0.028)	(0.007)	(0.008)
Education	-0.084 ***	-0.025 ***		1.168 ***	0.302 ***	0.335 ***
	(0.002)	(0.001)		(0.010)	(0.002)	(0.003)
Log(Income)	-0.034 ***	-0.022 ***		0.601 ***	0.137 ***	0.129 ***
	(0.002)	(0.001)		(0.020)	(0.005)	(0.006)
Wealth	-5.4E-7***	-0.013 ***		0.309 ***	0.061 ***	0.059 ***
	(4.3E-8)	(0.001)		(0.014)	(0.003)	(0.003)
Country dummies	yes	yes		yes	yes	yes
Wave dummies	yes	yes		yes	yes	yes
Constant	-0.837 ***			28.380 ***	7.856 ***	6.978 ***
	(0.038)			(0.177)	(0.042)	(0.051)
F-tests:	Chi-square			F	F	F
Retirement variables	79.110			18.270	108.020	47.060
	(0.000)			(0.000)	(0.000)	(0.000)
Country dummies	4296.620			1380.680	348.420	350.710
	(0.000)			(0.000)	(0.000)	(0.000)
Wave dummies	144.600			299.520	826.650	1024.640
	(0.000)			(0.000)	(0.000)	(0.000)
Observations	235545		Observations	235545	235545	235545
LogL	-126337.870		RMSE	6.362	1.519	1.836
LRχ^2(31)	15026.840		F(31,235513)	3168.29	2779.890	2646.250
Psuedo R^2	0.056		R^2	0.294	0.268	0.258

*In the second and third column the coefficients and the marginal effects of the probit model with depressed as the dependent variable are given. The fifth, sixth and seventh column give the coefficients of the OLS regressions with dependent variables: fluency, immediate recall memory and delayed recall memory, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table 15: Results of the second-stage Ordered Probit model of numeracy scores as a function of the predicted retirement variables and Demographic variables

Variable	**Coeff.**	**P(Num=1)**	**P(Num=2)**	**P(Num=3)**	**P(Num=4)**	**P(Num=5)**
Retirement:						
FR	0.015 **	-0.001 **	-0.002 **	-0.002 **	0.003 **	0.002**
	(0.006)	(3.4E-4)	(0.001)	(0.001)	(0.001)	(0.001)
PR	-0.065 ***	0.003 ***	0.010 ***	0.010 ***	-0.013 ***	-0.010 ***
	(0.008)	(4E-4)	(0.001)	(0.001)	(0.002)	(0.001)
Demographics:						
Age	-0.018 ***	0.001 ***	0.003 ***	0.003 ***	-0.004 ***	-0.003 ***
	(3.7E-4)	(2.1E-5)	(5.4E-5)	(5.4E-5)	(7.3E-5)	(5.5E-5)
Male	0.272 ***	-0.014 ***	-0.040 ***	-0.040 ***	0.055 ***	0.040 ***
	(0.005)	(2.9E-4)	(0.001)	(0.001)	(0.001)	(0.001)
Education	0.225 ***	-0.012 ***	-0.033 ***	-0.033 ***	0.045 ***	0.033 ***
	(0.002)	(1.5E-4)	(2.9E-4)	(2.6E-4)	(3.5E-4)	(2.9E-4)
Log(Income)	0.042 ***	-0.005 ***	-0.013 ***	-0.013 ***	0.018 ***	0.013 ***
	(0.002)	(1.8E-4)	(4.9E-4)	(4.9E-4)	(0.001)	(4.9E-4)
Wealth	5.3E-7***	-0.002 ***	-0.006 ***	-0.006 ***	0.009 ***	0.006 ***
	(3.1E-8)	(1.3E-4)	(3.7E-4)	(3.7E-4)	(0.001)	(3.7E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.494 ***					
	(0.031)					
Cut point 2	-1.467 ***					
	(0.030)					
Cut point 3	-0.410 ***					
	(0.030)					
Cut point 4	1.252 ***					
	(0.030)					
F-tests:	Chi-square	p-value		Observations	235545	
Retirement variables	80.390	0.000		LogL	-276034.660	
Country dummies	9589.810	0.000		LRχ^2(31)	53836.930	
Wave dummies	509.510	0.000		Psuedo R^2	0.089	

*In the second column the coefficients are given. The third through seventh column give the marginal effects of the explanatory variables on the probability of obtaining numeracy scores of 1-5, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

6.3 Robustness

In order to check whether the baseline results from the first- and second-stage models are reliable, some variations of the models are estimated to gain further insights. The results of the robustness checks can be found in Appendix B.

Models without adjusted data

As there are some adjustments made to the data, it is useful to check whether these influence the results by leaving the data where any adjustments are made out of the sample. These account for around 7% (16,477 out of 235,545) of the analysis-sample. After leaving out these observations, the sample now consists of 219,068 observations. The results from table B1-B5 show that the signs and the significance of the coefficients and marginal effects do not differ from those from the baseline results. In terms of magnitude, the marginal effects are slightly larger or smaller (0.001 or 0.002) than those from the baseline models but they are not drastically different. We can say that the results from the baseline models are robust to using the sample with adjusted data.

Using the second health response for the discordant individuals in wave 1

The discordant individuals who were asked the WHO version of the health question first, are assigned their first response (mapped to the US scale) as their score for self-perceived health. Table B6 and B7 show the first- and second-stage model of self-perceived health using the second health response for these individuals instead. The results show that the signs and the significance of the coefficients and marginal effects do not differ from those from the baseline results. In terms of magnitude, the marginal effects of the instrumental variables (in the first stage) and the retirement variables (in the second stage) are slightly larger (around 0.001 or 0.002) and the effects of the demographics (in both models) remain largely the same. It can be concluded that the baseline results are robust to using the first health response.

Models excluding respondents from wave 1

Due to the adjustments made to the self-perceived health for the respondents of wave 1 (in order to account for the two versions of the health questions), table B8-B10 show the first- and second stage models of self-perceived health estimated with a sample that excludes all the respondents from wave 1. The results show that the signs and the significance of the coefficients and marginal effects do not change from those from the baseline results. In terms of magnitude, the marginal effects do not differ (around 0.001 or 0.002) drastically than those from the baseline models. It can be concluded that the baseline results are robust to the inclusion of wave 1 respondents in the analysis.

Models without the retirement variables

In order to quantify the effects on the retirement variables on the health measures, second-stage models of the health measures without including the retirement variables as explanatory variables are estimated in table B11-B13. While the signs and the significance of the coefficients and marginal effects of the demographic variables do not change from those from the baseline results, the magnitudes do seem to be different (the difference in coefficients ranges from 0.001 to 0.485 and the difference in marginal effects ranges from 0.0003 to 0.004). To compare the results from table B11 and table 13, we can make use of a LR test: LR = 2(LogL(table 13)-LogL(table B11)) = 2*(-183,492.540+184,057.500)= 1129.92, where the test statistic LR is $\chi^2(2)$ distributed and the associated p-value is 0.000 indicating that the coefficients for full and partial retirement are not simultaneously equal to zero, meaning that including these variables create a statistically significant improvement in the fit of the model. Similarly, the comparisons between table 14 and table B12 and between table 15 and table B13, can show that the effects of the retirement variables are significant.

Using different thresholds for the predicted probabilities of the retirement variables

In the baseline model, thresholds of 50% and 5% are used for the predicted probabilities of the retirement decisions from the first-stage models in order to obtain the predicted (instrumented) retirement variables. Thresholds of 60% and 6% are now applied and the resulting models are given in table B14-B16. The signs and the significance of the coefficients and marginal effects of all the variables are not different from those from the baseline results. In terms of magnitude, the demographic variables do not differ (difference of around 0.001 to 0.002) drastically from the baseline models. However, the marginal effects of the retirement variables are now different (difference of 0.0008 to 0.015) due to the different thresholds. One would expect the marginal effects to be smaller, since with a larger threshold, a smaller fraction of the sample is fully or partially retired. We do observe smaller marginal effects of full retirement decisions on self-perceived compared to the baseline model, but larger effect of partial retirement. On mental health, the magnitude of both retirement variables are larger than the baseline results. It can be concluded that while the sign and significance of the effects remain unchanged, changing the thresholds results into different marginal effects in terms of magnitudes, which cannot be easily be explained by intuition. It makes more sense to use the thresholds of the baseline model, since they are closer to the sample full and partial retire rate.

Models including the interaction term between retirement and age

In order to determine the effects of full and partial retirement while taking into account possible interaction between retirement and age, models with an interaction are estimated in table B17-B19. The results show that the interaction between age and retirement is significant in all the models. For example, in the OLS model of fluency on retirement (in table B18) and other variables, the coefficient of full retirement is now around 3.178. The coefficient of the interaction term is -0.051. The effect of full retirement is now interpreted as 3.178+(-0.051)*age. For an average respondent in the sample with age 66, this effect becomes negative (-0.188). So the effect of full retirement can be negative in this case, which is in line with the intuition and the results from previous studies. The coefficients in the other models can be interpreted in similar ways. Figure B1 shows the calculated marginal effects of full and partial retirement on several cognitive measures for different values of age. It shows that full retirement has a negative effect on fluency, immediate recall memory and delayed recall memory scores for around age 65 or more. Partial retirement has a negative effect on cognition for around age 60 or less. People who are fully retired are likely to be in the age range of 65 or older and people typically start working part-time at a younger age as they transition into full retirement. These findings are now in line with previous studies, where retirement is found to be associated with cognitive decline.

7 Conclusion

In many countries, the aging population has put the sustainability of pension and public heath systems under pressure. Many countries try to solve this issue by raising the retirement ages. This makes understanding what kind of impact an individual's labor force exit may have on their health important to the successful implementation of the policy reforms. For some people, retirement may be a plus. After many years of daily work that is tiring for their bodies and minds, they get to relax. For others, retirement could be negative period filled with health issues and limitations. The aim of this research is to analyze the association between health problems and retirement decisions. Distinctions are made between full retirement and partial retirement and physical health and mental health. Due to the reverse causality issue between retirement and health condition, a two-stage instrumental variable approach is taken.

In the first stage, physical health (self-perceived health) is modeled in an ordered Probit model as a function of instrumental variables and other demographics. Both full retirement and partial retirement decisions are modeled as a function of instrumental variables and demographic variables in Probit models. The instrumental variables for self-perceived health are: Limitation with activities, Chronic diseases, BMI and Maxgrip. The instrumental variables for retirement decisions are: number of chilren, number of grandchildren, pension and marital status. From the first stage models we see that directionally consistent (the marginal effects are not given in their paper) with Disney, Emmerson and Wakefield (2006), each additional limitation in activities and each additional chronic disease increases the probability of poor self-perceived health by 3.2 and 3.6 percentage points, respectively. In line with the direction of the effects in Kanagae et al (2006), one standard deviation increase in BMI (an increase of 4.587 kg/m^2) is found to increase the risk of poor health by 0.9 percentage points, while one standard deviation increase in grip strength (an increase of 11.936 kg) decreases this risk by 2.7 percentage points. These magnitudes however, differ from Kanagae et al may be due to different sample and different methodology.

Regarding the first-stage models of the retirement decisions, the number of grandchildren is found to be insignificant in explaining partial retirement decisions (as in Lumsdaine and Vermeer (2015)). In addition, having pension benefits and having a partner increase the probability of being retired (fully or partially), all things being equal, with bigger effects on full retirement (38.2 and 8.7 percentage points, respectively) than on partial retirement (1.1 and 0.4 percentage points). In terms of the direction of the effects, these findings are in line with the findings in Schirle (2010), which found positive influences of pension wealth and partner on retirement. As in Bouman (2015), our first-stage results show that each additional child has a negative effect on full retirement (1.7 percentage points) compared to 1.3 percentage points found by Bouman and a smaller positive effect (0.3 percentage points) is observed for each additional grandchild than child compared to 0.4 percentage points found by Bouman (the magnitudes are very close may be because Bouman also made use SHARE dataset). These results from the first-stage models are directionally in line with the findings (with different magnitudes in some cases) from previous studies and the joint significance of the instrumental variables for both physical health and retirement behavior show that they are are valid instruments. This is of importance, since the instrumental variables were included to account for the endogeneity issue between health problems and retirement.

The goal of this paper is to describe the association between health problems and retirement decisions. After accounting for the potential problems of endogeneity with the first-stage models, we could finally model the association between health and retirement in several models. In the second-stage, self-perceived health predicted from the first-stage model (to account for endogeneity) is modeled in an ordered Probit model as a function of the predicted retirement variables and other demographics. In addition, the mental health variables are also modeled as a function of predicted retirement variables (to account for reverse causality) and demographic variables various models. For 'Depressed' a Probit model is used. Both memory scores and verbal fluency are modeled in ordinary least squares models. Lastly, for numeracy measures an ordered Probit model is estimated.

The first sub question was: "How are the partial and full retirement decisions associated with the physical health problems"? To answer this question, one of the key findings from the second-stage model of self-perceived (physical) health is that full retirement increases the probability of bad (fair or poor) health by 5.3 percentage points. This finding is in line with Curl and Townsend (2014), Bender and Theodossiou (2009), Sahlgren (2012) in terms of the direction of the effect. Furthermore, similar to Ettner and Grzywacz (2001) and Kantarci and Kolodziej (2016)), partial retirement is found to decrease the risk of bad health by 0.7 percentage points. In addition, the marginal effects of full retirement on self-perceived health are larger (difference from 0.1 to 3.4 percentage points) than those of partial retirement in all five health categories. Full retirement is one of the main determinants of self-perceived health.

The second sub question was: "How are the partial and full retirement decisions associated with the mental health problems"? To answer this question, one of the key findings from the second-stage models is that full retirement does not have a significant effect on depression, which is in line with the result in Riumallo-Herl et al (2014) but contrary to Coe and Zamarro (2011) and Kolodziej and Garcia-Gomez (2017). Furthermore, our model found that being partially retired makes it 2.6 percentage points more likely that a respondent is depressed, which is line with Kantarci and Kolodziej (2016) in terms of the direction of the effect (unfortunately, the magnitude is not comparable due to different method). On the cognitive measures, the initial results show that full retirement increases the fluency, immediate memory and delayed memory scores by 0.170, 0.110 and 0.103, respectively. Full retirement also increases the probability of a good numeracy score (4 or 5) by 0.5 percentage points. So overall, these initial results show that that full retirement leads to higher cognitive measures, which is contrary to the findings in the studies (Mazzonna and Peracchi (2012), Coe et al (2012), Rohwedder and Willis (2010)), where retirement is found to be associated with cognitive decline. However, Bingley and Martinello (2011) did find that after accounting for the endogeneity issue of retirement from the labour market, retirement increases word recall by 3.132, which compares against 0.213 in our case (the difference in magnitude may be due to different dataset). Furthermore, partial retirement is found to decrease immediate and delayed memory scores by 0.090 and 0.035 and it decreases the probability of a good numeracy score by 2.3 percentage points. Regarding the memory scores, the findings are contrary to Kantarci and Kolodziej (2016). Regarding numeracy however, the finding is directionally in line with Kantarci and Kolodziej's (the magnitudes are not comparable due to different methods). Overall, in terms of magnitude, partial retirement has larger marginal effects (difference of 0.002 to 0.050 in coefficient) on fluency and numeracy scores than full retirement but smaller effects (difference of 0.020 to 0.068 in coefficient) on memory scores.

After performing several robustness checks, the baseline results in this paper are robust to adjustments made to the dataset and robust to changes made to the sample due to the issue of two versions of self-perceived health for the respondents in wave 1 of SHARE. The effects of the retirement variables are jointly tested in an additional robustness check. Choosing different thresholds for the predicted probabilities of the retirement decisions does not lead to sensible results. Lastly, after including an interaction term between retirement and age in the second-stage models, the results show that full retirement now has a negative effect on mental health measures for respondents around age 65 or more. Partial retirement now has a negative effect on cognition for around age 60 or less. The results make more sense than the initial results (where retirement was found to be associated with better cognition) and are now in line with previous studies (Mazzonna and Peracchi (2012), Coe et al (2012), Rohwedder and Willis (2010)).

To conclude, in line with many previous studies this research finds evidence that full retirement is associated with worse physical health as well as declining mental health and that partial retirement is associated with better physical health but declining mental health. This paper contributes to the existing literature by confirming previous findings with different magnitudes (in some cases) using a different dataset. With the recent release of SHARE wave 6, this research had the luxury to carry out the analysis with a large (235,545 observations) dataset, which previous studies did not have access to. It is hoped that this paper has described the associations between (physical and mental) health problems and exit from the labor market (via partial retirement or full retirement), which in turn should help policy makers in successfully implementing policy reforms such as raising the retirement ages.

For further research, one could consider using lagged values of the variables to make use the longitudinal aspect of the SHARE dataset. In this case, the one-time respondents would not have lagged values and would not be included in the models. For the purpose of this research, the sample is chosen to be kept as large as possible. During the construction of the dataset, many respondents with missing values in a middle wave are kept in the sample with the idea in mind of estimating the effect of a treatment (shock to the family relation e.g. the death of a family member) on health conditions using a difference in the differences model (DiD). Unfortunately, this idea was changed in the beginning stage due to undesirable results. For further research, one could consider models such DiD or proportional hazard models that are suitable for panel data. For further research, one could also model retirement decisions in an ordered probit model, where the ordinal choices are fully retired, partially retired and not retired. In this paper a decision is made to model full retirement and partial retirement separately, in order to distinguish the effects of the two retirement variables on health measures in the second stage models.

Acknowledgments

I would like to thank Professor Lumsdaine and Professor Kole for providing their helpful feedback on previous versions of this paper. Progression of this research would not have been possible without their guidance. Their inputs are very much appreciated.

8 References

- Alavinia, S.M. and A. Burdorf. (2008). "Unemployment and retirement and ill-health: A cross-sectional analysis across European countries." International Archives of Occuppational and Environmental Health 82: 39-45. DOI: 10.1007/s00420-008-0304-6

- Avendano, M., Aro, A. R., and Mackenbach, J. (2005). "Socio-Economic Disparities in Physical Health in 10 European Countries." In A. Börsch-Span, A. Brugiavini, H. Juerges, J. Mackenbach, J. Siegrist, and G. Weber (Eds.), Health, Ageing and Retirement in Europe: Health and Health Care (pp. 89-94)

- Bender, K. , A. and Theodossiou, I. (2009). "Controlling for Endogeneity in the Health Socioeconomic Status Relationship of the Near Retired" in: Journal of Socio-Economics, 38(6), pp. 977-87.

- Bingley, P. and Martinello, A. (2011). "Retirement Improves Cognitive Performance", in: The Working Paper Series of The Danish National Centre for Social Research

- Bjelland, I. , Krokstad, S. , Mykletun, A. ,Dahl ,A. , Tell, G. and Tambs, K. (2008). "Does higher education protect against anxiety and depression? The HUNT study", Social Science & Medicine xx 1-12

- Boman, A. (2015). "Spending time together? Effects on the retirement decision from partner's labour market status". University of Gothenburg, Department of Economics in its series Working Papers in Economics with number 618.

- Bonner, W. , I. , A. ,Weiler, R. , Orisatoki, R. , Lu, X. , Andkhoie, M. , Ramsay, D. , Yaghoubi, M. , Steeves, M. , Szafron, M. and Farag, M. (2017). "Determinants of self-perceived health for Canadians aged 40 and older and policy implications", International Journal for Equity in Health 16:94 DOI 10.1186/s12939-017-0595-x

- Börsch-Supan, A., M. Brandt, C. Hunkler, T. Kneip, J. Korbmacher, F. Malter, B. Schaan, S. Stuck, S. Zuber (2013). "Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)". International Journal of Epidemiology. DOI: 10.1093/ije/dyt088

- Börsch-Supan, A. (2013). "Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 1". Release version: 2.6.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w1.260

- Börsch-Supan, A. (2013). "Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 2". Release version: 2.6.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w2.260

- Börsch-Supan, A. (2010). “Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 3 – SHARELIFE”. Release version: 1.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w3.100

- Börsch-Supan, A. (2013). “Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4”. Release version: 1.1.1. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w4.111

- Börsch-Supan, A. (2015). “Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5”. Release version: 1.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w5.100

- Börsch-Supan, A. (2017). “Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6”. Release version: 6.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w6.600

- Christelis D. (2011). “Imputation of missing data in waves 1 and 2 of SHARE”. SHARE WP Series: 01-2011.

- Coe, N. B., von Gaudecker, H., Lindeboom, M. and Maurer, J. (2012). “The Effect of Retirement on Cognitive Functioning”. Journal of Health Economics 21(8): 913-927. DOI: 10.1002/hec.1771

- Curl, A. L. and Townsend A. L. (2014). “ A Multilevel Dyadic Study of the Impact of Retirement on Self-Rated Health: Does Retirement Predict Worse Health in Married Couples?” in: Research on Aging Vol. 36(3) 297-321

- Coe, N. B. and G. Zamarro. (2011). “Retirement effects on health in Europe”. Journal of Health Economics 30(1): 77-86. DOI: 10.1016/j.jhealeco.2010.11.002.

- Cullinan, J. and B. Gannon. (2009). “Ill-Health and early retirement in Ireland”. Economics of Ageing Bulletin 2/2009: 1-2.

- De Luca G., Celidoni M., and Trevisan E. (2015). “Item nonresponse and imputation strategies in SHARE Wave 5” in Malter F. and Börsch-Supan A. (ed.), SHARE Wave 5: Innovations & Methodology, Munich: MEA, Max Planck Institute for Social Law and Social Policy.

- Denny, K. (2011). “Instrumental variable estimation of the effect of prayer on depression”. Social Science Medecine 73:1194–1199

- Dewey, M.E. and M.J. Prince. (2005). “Cognitive function”. In: A. Börsch-Supan et al. Health, Ageing and Retirement in Europe - First Results from the Survey of Health, Ageing and Retirement in Europe, 118-125. Mannheim: MEA.

- Disney, R., Emmerson, C. and Wakefield, M. (2006). “Ill-health and retirement in Britain: A panel data-based analysis”, Journal of Health Economics 25, 621-649.

- Dos Santos, M. , J. , Kawamura, H. , C. and Kassouf, A. , L. (2012), “Socioeconomic Conditions and Risk of Mental Depression: An Empirical Analysis for Brazilian Citizens”, in: Hindawi Publishing Corporation Economics Research International Volume 2012, Article ID 278906, 16 pages doi:10.1155/2012/278906

- Ettner, S. , L. and Grzywacz, J. , G. (2001). “Workers’ Perceptions of How Jobs Affect Health: A Social Ecological Perspective”, Journal of Occupational Health Psychology Vol. 6, No. 2, 101-11

- Gurley-Calvez, T. and Hill, B. (2011) “Time to Retire? The Effect of State Fiscal Policies on Retirement Decisions”, American Economic Review Vol. 101, No. 3, pp. 35-39

- Hagan, R., Jones, A. and Rice, N. (2009). “Health and Retirement in Europe”, Int. J. Environ. Res. Public Health 6, 2676-2695. doi: 10.3390/ijerph6102676

- Hajat, A. , Kaufman, J. , S. , Rose, K. , M. , Siddiqi, A. and Thomas, J. , C. (2010). “Long-Term Effects of Wealth on Mortality and Self-rated Health Status”, American Journal of Epidemiology Vol. 173, No. 2 DOI: 10.1093/aje/kwq348

- Hank, K. and Korbmacher, J.M. (2013). “Parenthood and Retirement”, European Societies, 15:3, 446-461, DOI: 10.1080/14616696.2012.750731

- Hochman, O. and Lewin-Epstein, N. (2013). “Determinants of early retirement preferences in Europe: The role of grandparenthood”, International Journal of Comparative Sociology 54(1) 29-47

- Hong, W. (2006). “The Effect of Defined Contribution Plans on the Retirement Decision.” Dissertation, Georgia State University

- Huppert, F. A., Gardener, E., and McWilliams, B. (2006). “Cognitive function”. in: J. Banks, E. Breeze, C. Lessof and J. Nazroo (Eds.), Retirement, health and relationships of the older person in England: The English Longitudinal Study on Ageing (Wave 2) (pp. 217-230). London: Institute for Fiscal Studies.

- Kanagae, M., Abe, Y., Honda, S., Takamura, N., Kusano, Y., Takemoto, T. and Aoyagi, K. (2006) “Determinants of self-rated health among community-dwelling women aged 40 years and over in Japan”, Tohoku J Exp Med, 210(1):11-9.

- Kantarci, T. and Kolodziej, I. (2016) “Effects of working part-time and full-time on physical and mental health in old age in Europe”, Network for Studies on Pensions, Aging and Retirement (Netspar)

- Kerkhofs, M., Lindeboom, M. and Theeuwes, J. (1999). “Retirement, financial incentives and health”. Labour Economics 6, 203-227

- Kolodziej, I. and García-Gómez, P. (2017). “The Causal Effects of Retirement on Mental Health: Looking Beyond the Mean Effects”, Ruhr Economic Papers No. 668

- Lauretani, F. , Russo, C. , R. , Bandinelli, S. ,Bartali, B. , Cavazzini, C. , Iorio, A. , D., Corsi, A. , M. , Rantanen, T. , Guralnik, J. , M. , Ferrucci, L. (2003). “Age-associated changes in skeletal muscles and their effect on mobility: an operational diagnosis of sarcopenia” in: Journal of Applied Physiology Vol. 95 no. 5, 1851-1860

- Lee, S. , Buring, J. , E. , Cook, N. , R. and Grodstein, F. (2006), “The Relation of Education and Income to Cognitive Function among Professional Women”, in: Neuro epidemiology 26(2): 93–101.

- Lumsdaine, R. and Exterkate, A. (2013) “How survey design affects self-assessed health responses in the Survey of Health, Ageing, and Retirement in Europe(SHARE)”. European Economic Review 63 (2013) 299–307.

- Lumsdaine, R. and Vermeer, S. (2015) “Retirement Timing of Women and the Role of Care Responsibilities for Grandchildren”. Demography 52, pp 433-454.

- Mazzonna, F. and F. Peracchi. (2012). “Ageing, cognitive abilities and retirement”. European Economic Review 56(4): 691-710. DOI: 10.1016/j.euroecorev.2012.03.004

- McGarry, K. (2004). “Health and Retirement: Do Changes in Health Affect Retirement Expectations?” The Journal of Human Resources 39, 624-648.

- Pagán, R. (2012). “Transitions to part-time work at older ages: The case of people with disabilities in Europe”. Disability and Society 26(1): 95-115.

- Reuser, M., Bonneux, L. and Willekens, F. (2010), “The effect of risk factors on the duration of cognitive impairment: A multistate life table analysis of the U.S. Health and Retirement Survey”, Network for Studies on Pensions, Aging and Retirement (Netspar)

- Rohwedder, S. and R. J. Willis. (2010). “Mental retirement”. Journal of Economic Perspectives 24(1): 119-138.DOI: 10.1257/jep.24.1.119.

- Riumallo-Herl, C. , Basu, S. , Stuckler ,D. , Courtin, E. and Avendano, M. (2014), "Job loss, wealth and depression during the Great Recession in the USA and Europe", International Journal of Epidemiology, 2014, 1508–1517 doi: 10.1093/ije/dyu048

- Sahlgren, G. , H. (2012). "Work 'til You Drop: Short- and Longer-Term Health Effects of Retirement in Europe ", IFN Working Paper No. 928

- Salthouse, T. (2009). "When does age-related cognitive decline begin?". Neurobiol Aging 30(4): 507–514. doi:10.1016/j.neurobiolaging.2008.09.023.

- Schirle, T. (2010). "Health, pensions, and the retirement decision: evidence from Canada". in: Canadian Journal on Aging 29 (4): 519-27. doi: 10.1017/S0714980810000541.

- Schneeweis, N., Skirbekk, V. and Winter-Ebmer, R. (2014). "Does Education Improve Cognitive Performance Four Decades After School Completion?" in: Demography, 2014, 51 (2), 619-643

- Shields, M. and Shooshtari S. (2001). "Determinants of self-perceived health" in: Health Reports, Vol. 13, No. 1, 35-52

- Solon, G., Haider, S. and Wooldridge, J. (2015). "What are we weighting for?" in: The Journal of Human Resources vol. 50 no. 2 301-316

- Villarreal, A. ,E., Grajales, S. , Lopez, L. , Britton, G. , B. and Panama Aging Research Initiative (2015), "Cognitive Impairment, Depression, and Cooccurrence of Both among the Elderly in Panama: Differential Associations with Multimorbidity and Functional Limitations", in: BioMed Research International Volume 2015, Article ID 718701, 7 pages

- Wagner, G. , J. , Ghosh-Dastidar, B. , Dickens, A. , Nakasujja, N. , Okello, E. , Luyirika, E. and Musisi, S. (2012), "Depression and its Relationship to Work Status and Income Among HIV Clients in Uganda", in: World J AIDS. 2012 ; 2(3):126–134. doi:10.4236/wja.2012.23018.

- Wu, S. ,Wang , R. ,Zhao, Y. , Ma, X. , Wu, M. , Yan, X. and He, J. (2013). "The relationship between self-rated health and objective health status: a population-based study", BMC Public Health 13:320

9 Appendix A: Complete table of the results section

Table A1: Results of the first-stage Ordered Probit model of self-perceived health as a function of the Instrumental variables and Demographic variables

Variable	**Coeff.**	**P(Excellent)**	**P(VeryGood)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Instrumental variables:						
Limitation with activities	0.305 ***	-0.038 ***	-0.043 ***	-0.004 ***	0.052 ***	0.032 ***
	(0.004)	(0.001)	(0.001)	(1.7E-4)	(0.001)	(4.1E-4)
Chronic diseases	0.337 ***	-0.042 ***	-0.047 ***	-0.004 ***	0.058 ***	0.036 ***
	(0.002)	(3E-4)	(2.8E-4)	(1.9E-4)	(3E-4)	(2.3E-4)
BMI	0.018 ***	-0.010 ***	-0.012 ***	-0.001 ***	0.014 ***	0.009 ***
	(0.001)	(3E-4)	(3.3E-4)	(5.6E-5)	(4E-4)	(2.5E-4)
Maxgrip	-0.021 ***	0.031 ***	0.035 ***	0.003 ***	-0.043 ***	-0.027 ***
	(3.2E-4)	(5E-4)	(0.001)	(1.5E-4)	(0.001)	(4.2E-4)
Demographic variables:						
Age	0.002 ***	-2.7E-4 ***	-3E-4 ***	-2.8E-5 ***	3.7E-4 ***	2.3E-4 ***
	(2.7E-4)	(3.4E-5)	(3.8E-5)	(3.7E-6)	(4.6E-5)	(2.9E-5)
Male	0.345 ***	-0.043 ***	-0.048 ***	-0.004 ***	0.059 ***	0.037 ***
	(0.007)	(0.001)	(0.001)	(2.1E-4)	(0.001)	(0.001)
Education	-0.102 ***	0.013 ***	0.014 ***	0.001 ***	-0.017 ***	-0.011 ***
	(0.002)	(2.2E-4)	(2.4E-4)	(6.2E-5)	(2.9E-4)	(1.9E-4)
Log(Income)	-0.030 ***	0.008 ***	0.009 ***	0.001 ***	-0.011 ***	-0.007 ***
	(0.002)	(4.1E-4)	(4.6E-4)	(5.5E-5)	(0.001)	(3.5E-4)
Wealth	-6E-7 ***	0.006 ***	0.007 ***	0.001 ***	-0.008 ***	-0.005***
	(3E-8)	(3E-4)	(3.4E-4)	(4.3E-5)	(4.1E-4)	(2.6E-4)
Country dummies:						
Germany	0.355 ***	-0.038 ***	-0.054 ***	-0.013 ***	0.066 ***	0.039 ***
	(0.013)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Sweden	-0.319 ***	0.050 ***	0.046 ***	-0.014 ***	-0.058 ***	-0.024 ***
	(0.013)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
Netherlands	0.060 ***	-0.008 ***	-0.009 ***	6.1E-6 ***	0.011 ***	0.006 ***
	(0.014)	(0.002)	(0.002)	(1.1E-4)	(0.003)	(0.001)
Spain	0.038 ***	-0.005 ***	-0.006 ***	1.1E-4***	0.007 ***	0.003 ***
	(0.013)	(0.002)	(0.002)	(7.2E-5)	(0.002)	(0.001)
Italy	0.121 ***	-0.015 ***	-0.019 ***	-0.001 ***	0.023 ***	0.012 ***
	(0.013)	(0.002)	(0.002)	(2E-4)	(0.002)	(0.001)
France	0.188 ***	-0.022 ***	-0.029 ***	-0.003 ***	0.035 ***	0.019 ***
	(0.012)	(0.001)	(0.002)	(3.3E-4)	(0.002)	(0.001)
Denmark	-0.444 ***	0.075 ***	0.061 ***	-0.026 ***	-0.079 ***	-0.031 ***
	(0.013)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
Greece	-0.210 ***	0.031 ***	0.031 ***	-0.007 ***	-0.038 ***	-0.017 ***
	(0.015)	(0.002)	(0.002)	(0.001)	(0.003)	(0.001)
Switzerland	-0.113 ***	0.016 ***	0.017 ***	-0.002 ***	-0.021 ***	-0.010 ***
	(0.014)	(0.002)	(0.002)	(3.7E-4)	(0.003)	(0.001)
Belgium	-0.041 ***	0.006 ***	0.006 ***	-0.001 ***	-0.008 ***	-0.004 ***
	(0.012)	(0.002)	(0.002)	(1.5E-4)	(0.002)	(0.001)

Table A1 continued

Variable	**Coeff.**	**P(Excellent)**	**P(VeryGood)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Israel	-0.207 ***	0.031 ***	0.030 ***	-0.007 ***	-0.038 ***	-0.017 ***
	(0.017)	(0.003)	(0.002)	(0.001)	(0.003)	(0.001)
Czech Republic	0.192 ***	-0.023 ***	-0.030 ***	-0.003 ***	0.036 ***	0.019 ***
	(0.014)	(0.002)	(0.002)	(3.9E-4)	(0.003)	(0.001)
Poland	0.657 ***	-0.058 ***	-0.097 ***	-0.047 ***	0.116 ***	0.085 ***
	(0.018)	(0.001)	(0.003)	(0.002)	(0.003)	(0.003)
Ireland	-0.853 ***	0.173 ***	0.095 ***	-0.083 ***	-0.137 ***	-0.048 ***
	(0.043)	(0.011)	(0.002)	(0.007)	(0.006)	(0.002)
Luxembourg	-0.081 ***	0.011 ***	0.012 ***	-0.001 ***	-0.015 ***	-0.007 ***
	(0.022)	(0.003)	(0.003)	(0.001)	(0.004)	(0.002)
Hungary	0.483 ***	-0.048 ***	-0.073 ***	-0.025 ***	0.088 ***	0.057 ***
	(0.026)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Portugal	0.536 ***	-0.051 ***	-0.081 ***	-0.031 ***	0.097 ***	0.065 ***
	(0.022)	(0.002)	(0.003)	(0.002)	(0.004)	(0.003)
Slovenia	0.342 ***	-0.037 ***	-0.053 ***	-0.012 ***	0.064 ***	0.038 ***
	(0.015)	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)
Estonia	0.970 ***	-0.070 ***	-0.134 ***	-0.098 ***	0.155 ***	0.147 ***
	(0.013)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Croatia	0.206 ***	-0.024 ***	-0.032 ***	-0.004 ***	0.039 ***	0.021 ***
	(0.026)	(0.003)	(0.004)	(0.001)	(0.005)	(0.003)
Wave dummies:						
Wave 2	0.130 ***	-0.016 ***	-0.018 ***	-0.001 ***	0.022 ***	0.013 ***
	(0.009)	(0.001)	(0.001)	(1.5E-4)	(0.002)	(0.001)
Wave 4	0.113 ***	-0.015 ***	-0.016 ***	-0.001 ***	0.019 ***	0.012 ***
	(0.009)	(0.001)	(0.001)	(1.1E-4)	(0.001)	(0.001)
Wave 5	0.077 ***	-0.010 ***	-0.011 ***	-2.4E-4 ***	0.013 ***	0.008 ***
	(0.008)	(0.001)	(0.001)	(6.5E-5)	(0.001)	(0.001)
Wave 6	0.080 ***	-0.010 ***	-0.011 ***	-2.8E ***	0.014 ***	0.008 ***
	(0.008)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Cutpoint 1	-1.479 ***					
	(0.033)					
Cutpoint 2	-0.519 ***					
	(0.033)					
Cutpoint 3	0.771 ***					
	(0.033)					
Cutpoint 4	2.141 ***					
	(0.033)					
F-tests:	Chi-squared	p-value		Observations	235,545	
Instrumental variables	59,349.780	0.000		LogL	-28,4847.56	
Country dummies	19,403.480	0.000		LR$\chi^2(33)$	115,043.89	
Wave dummies	253.890	0.000		Psuedo R^2	0.168	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: BMI, maxgrip, log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table A2: Results of the first-stage Probit models of full retirement decisions and partial retirement decisions as a function of the Instrumental variables and Demographic variables

	Full retirement		Partial retirement	
Variable	**Coeff.**	**P(Fully Retired)**	**Coeff.**	**P(Partially Retired)**
Instrumental variables:				
Number of children	-0.085 ***	-0.017 ***	0.018 ***	0.002 ***
	(0.003)	(0.001)	(0.004)	(4.2E-4)
Number of grandchildren	0.014 ***	0.003 ***	-0.001	-7.2E-5
	(0.002)	(0.000)	(0.002)	(2.2E-4)
Pension	1.913 ***	0.382 ***	0.106 ***	0.011 ***
	(0.010)	(0.001)	(0.011)	(0.001)
Marital status	0.436 ***	0.087 ***	0.040 ***	0.004 ***
	(0.009)	(0.002)	(0.012)	(0.001)
Demographic variables:				
Age	0.068 ***	0.014 ***	-0.028 ***	-0.003 ***
	(4.7E-4)	(8.5E-5)	(0.001)	(7.1E-5)
Male	0.237 ***	0.047 ***	-0.195 ***	-0.020 ***
	(0.007)	(0.001)	(0.009)	(0.001)
Education	0.033 ***	0.007 ***	0.049 ***	0.005 ***
	(0.003)	(0.001)	(0.003)	(3.5E-4)
Log(Income)	-0.066 ***	-0.027 ***	0.021 ***	0.005 ***
	(0.002)	(0.001)	(0.003)	(0.001)
Wealth	-2.1E-7***	-0.004 ***	3.4E-7***	0.003***
	(4.6E-8)	(0.001)	(4.6E-8)	(3.9E-4)
Country dummies:				
Germany	-0.396 ***	-0.075 ***	0.295 ***	0.036 ***
	(0.020)	(0.004)	(0.023)	(0.003)
Sweden	-0.574 ***	-0.112 ***	0.285 ***	0.035 ***
	(0.020)	(0.004)	(0.025)	(0.003)
Netherlands	-0.796 ***	-0.160 ***	0.312 ***	0.039 ***
	(0.021)	(0.004)	(0.025)	(0.003)
Spain	-0.654 ***	-0.129 ***	-0.189 ***	-0.016 ***
	(0.020)	(0.004)	(0.028)	(0.002)
Italy	-0.289 ***	-0.054 ***	-0.228 ***	-0.018 ***
	(0.020)	(0.004)	(0.028)	(0.002)
France	-0.062 ***	-0.011 ***	-0.192 ***	-0.016 ***
	(0.019)	(0.004)	(0.026)	(0.002)
Denmark	-0.600 ***	-0.118 ***	-0.003	-2.9E-4
	(0.022)	(0.004)	(0.027)	(0.003)
Greece	-0.360 ***	-0.068 ***	-0.151 ***	-0.013 ***
	(0.023)	(0.004)	(0.032)	(0.003)
Switzerland	-0.682 ***	-0.135 ***	0.440 ***	0.060 ***
	(0.022)	(0.004)	(0.025)	(0.003)
Belgium	-0.389 ***	-0.074 ***	0.004	4.2E-4
	(0.018)	(0.003)	(0.023)	(0.002)
Israel	-1.097 ***	-0.228 ***	0.296 ***	0.036 ***
	(0.026)	(0.005)	(0.031)	(0.004)

Table A2 continued

Variable	Full retirement Coeff.	P(Fully Retired)	Partial retirement Coeff.	P(Partially Retired)
Czech Republic	-0.275 ***	-0.051 ***	0.062 **	0.006 **
	(0.021)	(0.004)	(0.028)	(0.003)
Poland	-0.300 ***	-0.056 ***	-0.274 ***	-0.021 ***
	(0.027)	(0.005)	(0.043)	(0.003)
Ireland	-1.321 ***	-0.278 ***	0.438 ***	0.060 ***
	(0.066)	(0.015)	(0.076)	(0.013)
Luxembourg	-0.367 ***	-0.069 ***	-0.013	-0.001
	(0.034)	(0.007)	(0.043)	(0.004)
Hungary	-0.484 ***	-0.093 ***	-0.140 **	-0.012 **
	(0.039)	(0.008)	(0.058)	(0.005)
Portugal	-0.111 ***	-0.020 ***	0.094 **	0.010 **
	(0.033)	(0.006)	(0.042)	(0.005)
Slovenia	0.123 ***	0.022 ***	-0.451 ***	-0.031 ***
	(0.023)	(0.004)	(0.037)	(0.002)
Estonia	-0.540 ***	-0.105 ***	-0.118 ***	-0.010 ***
	(0.019)	(0.004)	(0.026)	(0.002)
Croatia	-0.072 *	-0.013 *	-0.438 ***	-0.030 ***
	(0.040)	(0.007)	(0.072)	(0.004)
Wave dummies:				
Wave 2	-0.116 ***	-0.023 ***	0.013	0.001
	(0.014)	(0.003)	(0.018)	(0.002)
Wave 4	-0.102 ***	-0.020 ***	0.060 ***	0.006 ***
	(0.013)	(0.003)	(0.017)	(0.002)
Wave 5	-0.126 ***	-0.025 ***	0.077 ***	0.008 ***
	(0.013)	(0.002)	(0.016)	(0.002)
Wave 6	-0.197 ***	-0.039 ***	0.081 ***	0.008 ***
	(0.013)	(0.003)	(0.016)	(0.002)
Constant	-5.089 ***		-0.340 ***	
	(0.043)		(0.057)	
F-tests:	Chi-squared	p-value	Chi-squared	p-value
Instrumental variables	39310.800	0.000	124.590	0.000
Country dummies	5478.240	0.000	2412.140	0.000
Wave dummies	245.370	0.000	40.780	0.000
Observations	235545		235545	
LogL	-83884.029		-45915.986	
LRχ^2(33)	158551.82		7114.99	
Psuedo R^2	0.4859		0.0719	

*In the second and fourth column the coefficients are given. The third and fifth column give the marginal effects of the explanatory variables on the probability of being fully retired and partially retired, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table A3: Results of the second-stage Ordered Probit model of the predicted self-perceived health as a function of the predicted retirement variables and Demographic variables

Variable	**Coeff.**	**P(Excellent)**	**P(Very good)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Retirement variables:						
Full retirement	0.240 ***	-0.001 ***	-0.038 ***	-0.015 ***	0.039 ***	0.014 ***
	(0.007)	(4.1E-5)	(0.001)	(4.9E-4)	(0.001)	(4.4E-4)
Partial retirement	-0.030 ***	9E-5***	0.005 ***	0.002 ***	-0.005 ***	-0.002 ***
	(0.008)	(3E-5)	(0.001)	(0.001)	(0.001)	(5E-4)
Demographic variables:						
Age	0.039 ***	-1.1E-4***	-0.006 ***	-0.002 ***	0.006 ***	0.002 ***
	(4.1E-4)	(5.6E-6)	(6.6E-5)	(3.2E-5)	(6.6E-5)	(2.9E-5)
Male	-0.100 ***	3E-4***	0.016 ***	0.006 ***	-0.016 ***	-0.006 ***
	(0.005)	(2.1E-5)	(0.001)	(3.5E-4)	(0.001)	(3.2E-4)
Education	-0.272 ***	0.001 ***	0.043 ***	0.017 ***	-0.045 ***	-0.016 ***
	(0.002)	(4E-5)	(3.2E-4)	(2.1E-4)	(3.3E-4)	(1.7E-4)
Log(Income)	-0.073 ***	4.5E-4***	0.024 ***	0.010 ***	-0.025 ***	-0.009 ***
	(0.002)	(2.4E-5)	(0.001)	(2.5E-4)	(0.001)	(2.3E-4)
Wealth	-2.2E-6***	0.001 ***	0.028 ***	0.011 ***	-0.029 ***	-0.011 ***
	(3.6E-8)	(2.7E-5)	(4.6E-4)	(2.3E-4)	(5E-4)	(1.9E-4)
Country dummies:						
Germany	0.878 ***	-0.001 ***	-0.117 ***	-0.109 ***	0.175 ***	0.052 ***
	(0.015)	(3.5E-5)	(0.002)	(0.002)	(0.003)	(0.001)
Sweden	-0.751 ***	0.003 ***	0.183 ***	-0.076 ***	-0.099 ***	-0.011 ***
	(0.015)	(1.9E-4)	(0.004)	(0.002)	(0.002)	(3.3E-4)
Netherlands	0.048 ***	-7E-5***	-0.009 ***	-0.001 ***	0.008 ***	0.001 ***
	(0.016)	(2.3E-5)	(0.003)	(2.3E-4)	(0.003)	(4.5E-4)
Spain	0.306 ***	-3.3E-4***	-0.053 ***	-0.015 ***	0.057 ***	0.011 ***
	(0.014)	(2.6E-5)	(0.003)	(0.001)	(0.003)	(0.001)
Italy	0.267 ***	-3E-4***	-0.047 ***	-0.012 ***	0.049 ***	0.009 ***
	(0.014)	(2.5E-5)	(0.003)	(0.001)	(0.003)	(0.001)
France	0.436 ***	-4.1E-4***	-0.071 ***	-0.030 ***	0.084 ***	0.017 ***
	(0.014)	(2.9E-5)	(0.002)	(0.001)	(0.003)	(0.001)
Denmark	-0.883 ***	0.005 ***	0.220 ***	-0.104 ***	-0.110 ***	-0.011 ***
	(0.015)	(2.5E-4)	(0.004)	(0.002)	(0.002)	(3.2E-4)
Greece	-0.379 ***	0.001 ***	0.084 ***	-0.019 ***	-0.058 ***	-0.007 ***
	(0.017)	(7.4E-5)	(0.004)	(0.001)	(0.003)	(3.4E-4)
Switzerland	-0.407 ***	0.001 ***	0.091 ***	-0.022 ***	-0.062 ***	-0.008 ***
	(0.016)	(7.8E-5)	(0.004)	(0.001)	(0.002)	(3.3E-4)
Belgium	0.153 ***	-2E-4 ***	-0.028 ***	-0.004 ***	0.028 ***	0.005 ***
	(0.013)	(-2.1E-5)	(0.002)	(4E-4)	(0.002)	(4.1E-4)
Israel	0.069 ***	-9.6E-5***	-0.013 ***	-0.001 ***	0.012 ***	0.002 ***
	(0.019)	(2.6E-5)	(0.004)	(3.8E-4)	(0.003)	(0.001)
Czech Republic	0.480 ***	-4.3E-4***	-0.077 ***	-0.036 ***	0.093 ***	0.020 ***
	(0.015)	(3E-5)	(0.003)	(0.001)	(0.003)	(0.001)
Poland	1.237 ***	-0.001 ***	-0.139 ***	-0.198 ***	0.241 ***	0.096 ***
	(0.020)	(3.7E-5)	(0.002)	(0.004)	(0.003)	(0.002)
Ireland	-1.601 ***	0.022 ***	0.424 ***	-0.289 ***	-0.145 ***	-0.012 ***
	(0.049)	(0.002)	(0.013)	(0.013)	(0.002)	(3.3E-4)

Table A3 continued

Variable	Coeff.	P(Excellent)	P(Very good)	P(Good)	P(Fair)	P(Poor)
Luxembourg	0.227 ***	-2.6E-4***	-0.040 ***	-0.009 ***	0.042 ***	0.008 ***
	(0.025)	(3E-5)	(0.004)	(0.002)	(0.005)	(0.001)
Hungary	1.018 ***	-0.001 ***	-0.127 ***	-0.142 ***	0.203 ***	0.067 ***
	(0.027)	(3.6E-5)	(0.003)	(0.006)	(0.005)	(0.003)
Portugal	1.331 ***	-0.001 ***	-0.143 ***	-0.223 ***	0.256 ***	0.111 ***
	(0.023)	(3.7E-5)	(0.002)	(0.006)	(0.004)	(0.003)
Slovenia	0.657 ***	-5E-4***	-0.097 ***	-0.064 ***	0.130 ***	0.032 ***
	(0.017)	(3.3E-5)	(0.002)	(0.002)	(0.003)	(0.001)
Estonia	1.718 ***	-0.001 ***	-0.154 ***	-0.329 ***	0.299 ***	0.185 ***
	(0.014)	(3.8E-5)	(0.002)	(0.003)	(0.002)	(0.002)
Croatia	0.476 ***	-4.2E-4***	-0.076 ***	-0.035 ***	0.092 ***	0.020 ***
	(0.029)	(3.2E-5)	(0.004)	(0.004)	(0.006)	(0.002)
Wave dummies:						
Wave 2	0.267 ***	-0.001 ***	-0.044 ***	-0.013 ***	0.044 ***	0.014 ***
	(0.010)	(5.7E-5)	(0.002)	(0.001)	(0.002)	(0.001)
Wave 4	0.280 ***	-0.001 ***	-0.046 ***	-0.014 ***	0.046 ***	0.015 ***
	(0.010)	(5.8E-5)	(0.002)	(5E-4)	(0.002)	(0.001)
Wave 5	0.225 ***	-0.001 ***	-0.038 ***	-0.010 ***	0.037 ***	0.012 ***
	(0.009)	(5.1E-5)	(0.002)	(3.7E-4)	(0.001)	(4.6E-4)
Wave 6	0.198 ***	-0.001 ***	-0.033 ***	-0.008 ***	0.032 ***	0.010 ***
	(0.009)	(5.1E-5)	(0.002)	(3.7E-4)	(0.002)	(4.6E-4)
Cut point 1	-2.741 ***					
	(0.041)					
Cut point 2	0.410 ***					
	(0.033)					
Cut point 3	2.596 ***					
	(0.034)					
Cut point 4	4.179 ***					
	(0.035)					
F-tests:	Chi-square	p-value		Observations	235545	
Retirement variables	1126.630	0.000		LogL	-183,492.540	
Country dummies	49,722.340	0.000		LRχ^2(31)	135,556.800	
Wave dummies	991.870	0.000		Psuedo R^2	0.270	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table A4: Results of the second-stage models of several mental health measures as a function of the predicted retirement variables and Demographic variables

Variable	Depressed (probit) Coeff.	P(Depressed)	Fluency (OLS) Coeff.	Imm. mem. (OLS) Coeff.	Delayed mem. (OLS) Coeff.
Retirement variables:					
Full retirement	-0.001	-3.6E-4	0.170 ***	0.110 ***	0.103 ***
	(0.008)	(0.003)	(0.038)	(0.009)	(0.011)
Partial retirement	0.086 ***	0.026 ***	0.172 ***	-0.090 ***	-0.035***
	(0.010)	(0.003)	(0.044)	(0.011)	(0.013)
Demographic variables:					
Age	0.009 ***	0.003 ***	-0.189 ***	-0.061 ***	-0.070 ***
	(4.6E-4)	(1.4E-4)	(0.002)	(0.001)	(0.001)
Male	-0.407 ***	-0.123 ***	0.007	-0.324 ***	-0.428 ***
	(0.006)	(0.002)	(0.028)	(0.007)	(0.008)
Education	-0.084 ***	-0.025 ***	1.168 ***	0.302 ***	0.335 ***
	(0.002)	(0.001)	(0.010)	(0.002)	(0.003)
Log(Income)	-0.034 ***	-0.022 ***	0.601 ***	0.137 ***	0.129 ***
	(0.002)	(0.001)	(0.020)	(0.005)	(0.006)
Wealth	-5.4E-7***	-0.013 ***	0.309 ***	0.061 ***	0.059 ***
	(4.3E-8)	(0.001)	(0.014)	(0.003)	(0.003)
Country dummies:					
Germany	0.152 ***	0.043 ***	-1.981 ***	-0.195 ***	-0.319 ***
	(0.017)	(0.005)	(0.076)	(0.018)	(0.022)
Sweden	-0.084 ***	-0.022 ***	0.674 ***	-0.021	0.141 ***
	(0.018)	(0.005)	(0.079)	(0.019)	(0.023)
Netherlands	-0.028	-0.007	-2.414 ***	-0.113 ***	-0.062 ***
	(0.019)	(0.005)	(0.083)	(0.020)	(0.024)
Spain	0.231 ***	0.067 ***	-5.398 ***	-0.957 ***	-0.989 ***
	(0.017)	(0.005)	(0.075)	(0.018)	(0.022)
Italy	0.343 ***	0.103 ***	-6.037 ***	-0.601 ***	-0.757 ***
	(0.017)	(0.005)	(0.076)	(0.018)	(0.022)
France	0.459 ***	0.142 ***	-3.589 ***	-0.455 ***	-0.402 ***
	(0.016)	(0.005)	(0.073)	(0.017)	(0.021)
Denmark	-0.044 **	-0.012 **	-0.101	-0.085 ***	0.016
	(0.019)	(0.005)	(0.079)	(0.019)	(0.023)
Greece	0.153 ***	0.043 ***	-8.107 ***	-0.399 ***	-0.573 ***
	(0.020)	(0.006)	(0.087)	(0.021)	(0.025)
Switzerland	0.011	0.003	-2.647 ***	-0.009	0.114 ***
	(0.020)	(0.005)	(0.084)	(0.020)	(0.024)
Belgium	0.349 ***	0.105 ***	-2.689 ***	-0.346 ***	-0.431 ***
	(0.016)	(0.005)	(0.070)	(0.017)	(0.020)
Israel	0.179 ***	0.051 ***	-3.446 ***	-0.319 ***	-0.389 ***
	(0.022)	(0.006)	(0.099)	(0.024)	(0.029)
Czech Republic	-0.007	-0.002	0.696 ***	0.063 ***	-0.263 ***
	(0.018)	(0.005)	(0.081)	(0.019)	(0.023)
Poland	0.644 ***	0.210 ***	-4.912 ***	-0.816 ***	-1.107 ***
	(0.022)	(0.008)	(0.106)	(0.025)	(0.031)

Table A4 continued

	Depressed (probit)			Fluency (OLS)	Imm. mem. (OLS)	Delayed mem. (OLS)
Variable	**Coeff.**	**P(Depressed)**		**Coeff.**	**Coeff.**	**Coeff.**
Ireland	-0.285 ***	-0.066 ***		-4.864 ***	0.346 ***	0.549 ***
	(0.060)	(0.012)		(0.254)	(0.061)	(0.073)
Luxembourg	0.331 ***	0.099 ***		-5.234 ***	-0.386 ***	-0.111 ***
	(0.029)	(0.009)		(0.132)	(0.031)	(0.038)
Hungary	0.356 ***	0.107 ***		-3.652 ***	-0.158 ***	-0.419 ***
	(0.031)	(0.010)		(0.149)	(0.036)	(0.043)
Portugal	0.567 ***	0.181 ***		-7.082 ***	-1.080 ***	-1.145 ***
	(0.027)	(0.009)		(0.127)	(0.030)	(0.037)
Slovenia	0.187 ***	0.053 ***		-0.940 ***	-0.679 ***	-1.083 ***
	(0.020)	(0.006)		(0.089)	(0.021)	(0.026)
Estonia	0.530 ***	0.168 ***		-1.369 ***	-0.420 ***	-0.624 ***
	(0.016)	(0.005)		(0.075)	(0.018)	(0.022)
Croatia	0.252 ***	0.073 ***		-2.872 ***	-0.224 ***	-0.674 ***
	(0.033)	(0.010)		(0.154)	(0.037)	(0.045)
Wave dummies:						
Wave 2	-0.069 ***	-0.020 ***		0.396 ***	0.228 ***	0.318 ***
	(0.012)	(0.003)		(0.053)	(0.013)	(0.015)
Wave 4	0.057 ***	0.017 ***		0.337 ***	0.400 ***	0.548 ***
	(0.011)	(0.003)		(0.051)	(0.012)	(0.015)
Wave 5	0.021 **	0.006 **		0.911 ***	0.497 ***	0.664 ***
	(0.011)	(0.003)		(0.048)	(0.011)	(0.014)
Wave 6	0.011	0.003		1.374 ***	0.614 ***	0.834 ***
	(0.011)	(0.003)		(0.049)	(0.012)	(0.014)
Constant	-0.837 ***			28.380 ***	7.856 ***	6.978 ***
	(0.038)			(0.177)	(0.042)	(0.051)
F-tests:	Chi-square			F	F	F
Retirement variables	79.110			18.270	108.020	47.060
	(0.000)			(0.000)	(0.000)	(0.000)
Country dummies	4296.620			1380.680	348.420	350.710
	(0.000)			(0.000)	(0.000)	(0.000)
Wave dummies	144.600			299.520	826.650	1024.640
	(0.000)			(0.000)	(0.000)	(0.000)
Observations	235545		Observations	235545	235545	235545
LogL	-126337.870		RMSE	6.362	1.519	1.836
LRχ^2(31)	15026.840		F(31,235513)	3168.29	2779.890	2646.250
Psuedo R^2	0.056		R^2	0.294	0.268	0.258

*In the second and third column the coefficients and the marginal effects of the probit model with depressed as the dependent variable are given. The fifth, sixth and seventh column give the coefficients of the OLS regressions with dependent variables: fluency, immediate recall memory and delayed recall memory, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table A5: Results of the second-stage Ordered Probit model of numeracy scores as a function of the predicted retirement variables and Demographic variables

Variable	**Coeff.**	**P(Num=1)**	**P(Num=2)**	**P(Num=3)**	**P(Num=4)**	**P(Num=5)**
Retirement:						
FR	0.015 **	-0.001 **	-0.002 **	-0.002 **	0.003 **	0.002**
	(0.006)	(3.4E-4)	(0.001)	(0.001)	(0.001)	(0.001)
PR	-0.065 ***	0.003 ***	0.010 ***	0.010 ***	-0.013 ***	-0.010 ***
	(0.008)	(4E-4)	(0.001)	(0.001)	(0.002)	(0.001)
Demographics:						
Age	-0.018 ***	0.001 ***	0.003 ***	0.003 ***	-0.004 ***	-0.003 ***
	(3.7E-4)	(2.1E-5)	(5.4E-5)	(5.4E-5)	(7.3E-5)	(5.5E-5)
Male	0.272 ***	-0.014 ***	-0.040 ***	-0.040 ***	0.055 ***	0.040 ***
	(0.005)	(2.9E-4)	(0.001)	(0.001)	(0.001)	(0.001)
Education	0.225 ***	-0.012 ***	-0.033 ***	-0.033 ***	0.045 ***	0.033 ***
	(0.002)	(1.5E-4)	(2.9E-4)	(2.6E-4)	(3.5E-4)	(2.9E-4)
Log(Income)	0.042 ***	-0.005 ***	-0.013 ***	-0.013 ***	0.018 ***	0.013 ***
	(0.002)	(1.8E-4)	(4.9E-4)	(4.9E-4)	(0.001)	(4.9E-4)
Wealth	5.3E-7***	-0.002 ***	-0.006 ***	-0.006 ***	0.009 ***	0.006 ***
	(3.1E-8)	(1.3E-4)	(3.7E-4)	(3.7E-4)	(0.001)	(3.7E-4)
Country dummies:						
Germany	-0.233 ***	0.008 ***	0.031 ***	0.042 ***	-0.040 ***	-0.041 ***
	(0.013)	(4.5E-4)	(0.002)	(0.002)	(0.002)	(0.002)
Sweden	-0.034 **	0.001 **	0.004 **	0.006 **	-0.005 **	-0.007 **
	(0.014)	(3.6E-4)	(0.002)	(0.003)	(0.002)	(0.003)
Netherlands	0.014	-3.5E-4	-0.002	-0.003	0.002	0.003
	(0.014)	(3.6E-4)	(0.002)	(0.003)	(0.002)	(0.003)
Spain	-0.867 ***	0.054 ***	0.141 ***	0.111 ***	-0.200 ***	-0.106 ***
	(0.013)	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)
Italy	-0.484 ***	0.021 ***	0.070 ***	0.080 ***	-0.097 ***	-0.074 ***
	(0.013)	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)
France	-0.497 ***	0.021 ***	0.073 ***	0.082 ***	-0.100 ***	-0.075 ***
	(0.013)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Denmark	-0.242 ***	0.008 ***	0.032 ***	0.044 ***	-0.041 ***	-0.042 ***
	(0.014)	(4.8E-4)	(0.002)	(0.002)	(0.002)	(0.002)
Greece	-0.230 ***	0.007 ***	0.030 ***	0.042 ***	-0.039 ***	-0.040 ***
	(0.015)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)
Switzerland	-0.073 ***	0.002 ***	0.009 ***	0.014 ***	-0.011 ***	-0.014 ***
	(0.015)	(4.1E-4)	(0.002)	(0.003)	(0.002)	(0.003)
Belgium	-0.364 ***	0.014 ***	0.051 ***	0.063 ***	-0.068 ***	-0.060 ***
	(0.012)	(4.8E-4)	(0.002)	(0.002)	(0.002)	(0.002)
Israel	-0.212 ***	0.007 ***	0.027 ***	0.039 ***	-0.035 ***	-0.038 ***
	(0.017)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)
Czech Republic	-0.008	2.1E-4	0.001	0.002	-0.001	-0.002
	(0.014)	(3.6E-4)	(0.002)	(0.003)	(0.002)	(0.003)
Poland	-0.513 ***	0.022 ***	0.075 ***	0.084 ***	-0.104 ***	-0.077 ***
	(0.018)	(0.001)	(0.003)	(0.003)	(0.004)	(0.003)
Ireland	-0.066	0.002	0.008	0.013	-0.010	-0.013
	(0.043)	(0.001)	(0.005)	(0.008)	(0.007)	(0.008)

Table A5 continued

Variable	Coeff.	P(Num=1)	P(Num=2)	P(Num=3)	P(Num=4)	P(Num=5)
Luxembourg	-0.332 ***	0.012 ***	0.045 ***	0.058 ***	-0.061 ***	-0.055 ***
	(0.023)	(0.001)	(0.003)	(0.004)	(0.005)	(0.003)
Hungary	-0.150 ***	0.004 ***	0.019 ***	0.028 ***	-0.024 ***	-0.027 ***
	(0.025)	(0.001)	(0.003)	(0.005)	(0.004)	(0.004)
Portugal	-0.922 ***	0.061 ***	0.151 ***	0.113 ***	-0.215 ***	-0.109 ***
	(0.021)	(0.002)	(0.004)	(0.002)	(0.006)	(0.002)
Slovenia	-0.535 ***	0.024 ***	0.079 ***	0.086 ***	-0.110 ***	-0.079 ***
	(0.015)	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)
Estonia	-0.399 ***	0.015 ***	0.056 ***	0.069 ***	-0.076 ***	-0.064 ***
	(0.013)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Croatia	-0.395 ***	0.015 ***	0.056 ***	0.068 ***	-0.075 ***	-0.063 ***
	(0.026)	(0.001)	(0.004)	(0.004)	(0.006)	(0.004)
Wave dummies:						
Wave 2	0.113 ***	-0.006 ***	-0.017 ***	-0.016 ***	0.024 ***	0.015 ***
	(0.009)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Wave 4	0.055 ***	-0.003 ***	-0.008 ***	-0.008 ***	0.012 ***	0.007 ***
	(0.009)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Wave 5	0.128 ***	-0.007 ***	-0.019 ***	-0.018 ***	0.027 ***	0.018 ***
	(0.008)	(4.8E-4)	(0.001)	(0.001)	(0.002)	(0.001)
Wave 6	0.160 ***	-0.009 ***	-0.024 ***	-0.023 ***	0.033 ***	0.023 ***
	(0.008)	(4.8E-4)	(0.001)	(0.001)	(0.002)	(0.001)
Cut point 1	-2.494 ***					
	(0.031)					
Cut point 2	-1.467 ***					
	(0.030)					
Cut point 3	-0.410 ***					
	(0.030)					
Cut point 4	1.252 ***					
	(0.030)					
F-tests:	Chi-square	p-value		Observations	235545	
Retirement variables	80.390	0.000		LogL	-276034.660	
Country dummies	9589.810	0.000		LRχ^2(31)	53836.930	
Wave dummies	509.510	0.000		Psuedo R^2	0.089	

*In the second column the coefficients are given. The third through seventh column give the marginal effects of the explanatory variables on the probability of obtaining numeracy scores of 1-5, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

10 Appendix B: Robustness

Models without adjusted data

Table B1: Results of the first-stage Ordered Probit model of self-perceived health as a function of the Instrumental variables and Demographic variables (without adjusted data)

Variable	**Coeff.**	**P(Excellent)**	**P(VeryGood)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Instrumental variables:						
Limitation with activities	0.318 ***	-0.041 ***	-0.045 ***	-0.002 ***	0.055 ***	0.032 ***
	(0.004)	(0.001)	(0.001)	(1.9E-4)	(0.001)	(4.4E-4)
Chronic diseases	0.339 ***	-0.043 ***	-0.048 ***	-0.002 ***	0.059 ***	0.034 ***
	(0.002)	(3.1E-4)	(3E-4)	(2E-4)	(3.2E-4)	(2.4E-4)
BMI	0.019 ***	-0.011 ***	-0.012 ***	-0.001 ***	0.015 ***	0.009 ***
	(0.001)	(3.2E-4)	(3.3E-4)	(5.3E-5)	(4.3E-4)	(2.5E-4)
Maxgrip	-0.021 ***	0.032 ***	0.035 ***	0.001 ***	-0.043 ***	-0.025 ***
	(3.3E-4)	(0.001)	(0.001)	(1.5E-4)	(0.001)	(4.2E-4)
Demographic variables:						
Age	0.002 ***	-2.8E-4 ***	-3.1E-4 ***	-1.3E-5 ***	3.8E-4 ***	2.2E-4 ***
	(2.8E-4)	(3.6E-5)	(4E-5)	(2.2E-6)	(4.9E-5)	(2.9E-5)
Male	0.344 ***	-0.044 ***	-0.049 ***	-0.002 ***	0.060 ***	0.035 ***
	(0.007)	(0.001)	(0.001)	(2.1E-4)	(0.001)	(0.001)
Education	-0.101 ***	0.013 ***	0.014 ***	0.001 ***	-0.018 ***	-0.010 ***
	(0.002)	(2.3E-4)	(2.5E-4)	(6.1E-5)	(3.1E-4)	(1.9E-4)
Log(Income)	-0.031 ***	0.008 ***	0.009 ***	3.9E-4***	-0.011 ***	-0.007 ***
	(0.002)	(4.4E-4)	(4.9E-4)	(4.4E-5)	(0.001)	(3.5E-4)
Wealth	-6E-7 ***	0.006 ***	0.007 ***	3E-4 ***	-0.009 ***	-0.005***
	(3E-8)	(3.2E-4)	(3.5E-4)	(3.3E-5)	(4.3E-4)	(2.5E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-1.456 ***					
	(0.034)					
Cut point 2	-0.490 ***					
	(0.034)					
Cut point 3	0.805 ***					
	(0.034)					
Cut point 4	2.182 ***					
	(0.034)					
F-tests:	Chi-squared	p-value		Observations	219,068	
Instrumental variables	54,752.540	0.000		LogL	-265,045.48	
Country dummies	18,119.500	0.000		LR$\chi^2(33)$	105,180.03	
Wave dummies	211.370	0.000		Psuedo R^2	0.166	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: BMI, maxgrip, log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B2: Results of the first-stage Probit models of full retirement decisions and partial retirement decisions as a function of the Instrumental variables and Demographic variables (without adjusted data)

	Full retirement		Partial retirement	
Variable	**Coeff.**	**P(Fully Retired)**	**Coeff.**	**P(Partially Retired)**
Instrumental variables:				
Number of children	-0.088 ***	-0.017 ***	0.020 ***	0.002 ***
	(0.004)	(0.001)	(0.004)	(4.4E-4)
Number of grandchildren	0.016 ***	0.003 ***	-0.001	-7.1E-5
	(0.002)	(3.3E-4)	(0.002)	(2.3E-4)
Pension	1.910 ***	0.375 ***	0.113 ***	0.012 ***
	(0.010)	(0.001)	(0.012)	(0.001)
Marital status	0.421 ***	0.083 ***	0.035 ***	0.004 ***
	(0.009)	(0.002)	(0.012)	(0.001)
Demographic variables:				
Age	0.071 ***	0.014 ***	-0.027 ***	-0.003 ***
	(0.001)	(8.8E-5)	(0.001)	(7.5E-5)
Male	0.222 ***	0.044 ***	-0.198 ***	-0.021 ***
	(0.008)	(0.001)	(0.010)	(0.001)
Education	0.028 ***	0.006 ***	0.048 ***	0.005 ***
	(0.003)	(0.001)	(0.003)	(3.7E-4)
Log(Income)	-0.067 ***	-0.027 ***	0.021 ***	0.005 ***
	(0.003)	(0.001)	(0.004)	(0.001)
Wealth	-2.3E-7***	-0.004 ***	3.3E-7***	0.003***
	(4.7E-8)	(0.001)	(4.8E-8)	(4.1E-4)
Country dummies	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes
Constant	-5.208 ***		-0.340 ***	
	(0.045)		(0.057)	
F-tests:	Chi-squared	p-value	Chi-squared	p-value
Instrumental variables	36123.910	0.000	130.600	0.000
Country dummies	4977.890	0.000	2259.560	0.000
Wave dummies	232.620	0.000	44.640	0.000
Observations	219068		219068	
LogL	-76652.794		-43458.108	
LRχ^2(33)	150241.310		6468.670	
Psuedo R^2	0.495		0.069	

*In the second and fourth column the coefficients are given. The third and fifth column give the marginal effects of the explanatory variables on the probability of being fully retired and partially retired, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B3: Results of the second-stage Ordered Probit model of the predicted self-perceived health as a function of the predicted retirement variables and Demographic variables (without adjusted data)

Variable	**Coeff.**	**P(Excellent)**	**P(Very good)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Retirement variables:						
Full retirement	0.234 ***	-0.001 ***	-0.039 ***	-0.012 ***	0.038 ***	0.013 ***
	(0.008)	(4.4E-5)	(0.001)	(4.2E-4)	(0.001)	(4.3E-4)
Partial retirement	-0.025 ***	8E-5***	0.004 ***	0.001 ***	-0.004 ***	-0.001 ***
	(0.009)	(2.8E-5)	(0.001)	(4.4E-4)	(0.001)	(4.8E-4)
Demographic variables:						
Age	0.039 ***	-1.3E-4***	-0.006 ***	-0.002 ***	0.006 ***	0.002 ***
	(4.3E-4)	(6.1E-6)	(7.3E-5)	(3E-5)	(7E-5)	(2.9E-5)
Male	-0.092 ***	3E-4***	0.015 ***	0.005 ***	-0.015 ***	-0.005 ***
	(0.006)	(2.3E-5)	(0.001)	(2.9E-4)	(0.001)	(3.1E-4)
Education	-0.271 ***	0.001 ***	0.045 ***	0.014 ***	-0.044 ***	-0.015 ***
	(0.002)	(4.3E-5)	(3.4E-4)	(2E-4)	(3.4E-4)	(1.7E-4)
Log(Income)	-0.077 ***	0.001***	0.026 ***	0.008 ***	-0.026 ***	-0.009 ***
	(0.002)	(2.8E-5)	(0.001)	(2.2E-4)	(0.001)	(2.2E-4)
Wealth	-2.2E-6***	0.001 ***	0.029 ***	0.009 ***	-0.029 ***	-0.010 ***
	(3.8E-8)	(2.9E-5)	(0.001)	(2E-4)	(0.001)	(1.9E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.753 ***					
	(0.042)					
Cut point 2	0.135 ***					
	(0.035)					
Cut point 3	2.578 ***					
	(0.035)					
Cut point 4	4.156 ***					
	(0.036)					
F-tests:	Chi-square	p-value		Observations	219068	
Retirement variables	973.710	0.000		LogL	-170,062.29	
Country dummies	46,702.620	0.000		LRχ^2(31)	123,963.95	
Wave dummies	760.690	0.000		Psuedo R^2	0.267	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B4: Results of the second-stage models of several mental health measures as a function of the predicted retirement variables and Demographic variables (without adjusted data)

Variable	Depressed (probit) Coeff.	P(Depressed)		Fluency (OLS) Coeff.	Imm. mem. (OLS) Coeff.	Delayed mem. (OLS) Coeff.
Retirement variables:						
Full retirement	-0.007	-0.002		0.139 ***	0.100 ***	0.097 ***
	(0.009)	(0.003)		(0.039)	(0.009)	(0.011)
Partial retirement	0.088 ***	0.026 ***		0.198 ***	-0.075 ***	-0.032***
	(0.010)	(0.003)		(0.045)	(0.011)	(0.013)
Demographic variables:						
Age	0.009 ***	0.003 ***		-0.184 ***	-0.058 ***	-0.069 ***
	(4.9E-4)	(1.5E-4)		(0.002)	(0.001)	(0.001)
Male	-0.403 ***	-0.121 ***		-0.018	-0.329 ***	-0.443 ***
	(0.007)	(0.002)		(0.029)	(0.007)	(0.008)
Education	-0.084 ***	-0.025 ***		1.149 ***	0.291 ***	0.326 ***
	(0.002)	(0.001)		(0.010)	(0.002)	(0.003)
Log(Income)	-0.036 ***	-0.022 ***		0.595 ***	0.129 ***	0.123 ***
	(0.002)	(0.001)		(0.021)	(0.005)	(0.006)
Wealth	-5.3E-7***	-0.013 ***		0.296 ***	0.056 ***	0.056 ***
	(4.5E-8)	(0.001)		(0.015)	(0.003)	(0.004)
Country dummies	yes	yes		yes	yes	yes
Wave dummies	yes	yes		yes	yes	yes
Constant	-0.810 ***			30.721 ***	8.335 ***	7.496 ***
	(0.040)			(0.165)	(0.039)	(0.0548)
F-tests:	Chi-square			F	F	F
Retirement variables	76.900			16.520	79.300	38.450
	(0.000)			(0.000)	(0.000)	(0.000)
Country dummies	4005.660			1297.460	292.990	306.630
	(0.000)			(0.000)	(0.000)	(0.000)
Wave dummies	144.600			288.580	813.000	1012.880
	(0.000)			(0.000)	(0.000)	(0.000)
Observations	219068		Observations	219068	219068	219068
LogL	-116225.240		RMSE	6.329	1.496	1.830
LRχ^2(31)	13344.820		F(31,235513)	2848.21	2413.590	2313.420
Psuedo R^2	0.054		R^2	0.287	0.255	0.247

*In the second and third column the coefficients and the marginal effects of the probit model with depressed as the dependent variable are given. The fifth, sixth and seventh column give the coefficients of the OLS regressions with dependent variables: fluency, immediate recall memory and delayed recall memory, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B5: Results of the second-stage Ordered Probit model of numeracy scores as a function of the predicted retirement variables and Demographic variables (without adjusted data)

Variable	Coeff.	P(Num=1)	P(Num=2)	P(Num=3)	P(Num=4)	P(Num=5)
Retirement:						
FR	0.008 **	-3.6E-4**	-0.001 **	-0.001 **	0.002 **	0.001**
	(0.007)	(2.9E-4)	(0.001)	(0.001)	(0.001)	(0.001)
PR	-0.057 ***	0.002 ***	0.008 ***	0.009 ***	-0.011 ***	-0.009 ***
	(0.008)	(3.3E-4)	(0.001)	(0.001)	(0.002)	(0.001)
Demographics:						
Age	-0.017 ***	0.001 ***	0.002 ***	0.003 ***	-0.003 ***	-0.003 ***
	(3.9E-4)	(1.9E-5)	(5.7E-5)	(6.3E-5)	(7.6E-5)	(6E-5)
Male	0.277 ***	-0.012 ***	-0.040 ***	-0.045 ***	0.055 ***	0.042 ***
	(0.005)	(2.6E-4)	(0.001)	(0.001)	(0.001)	(0.001)
Education	0.219 ***	-0.009 ***	-0.032 ***	-0.035 ***	0.043 ***	0.033 ***
	(0.002)	(1.4E-4)	(2.9E-4)	(3E-4)	(3.6E-4)	(3.1E-4)
Log(Income)	0.040 ***	-0.004 ***	-0.012 ***	-0.014 ***	0.017 ***	0.013 ***
	(0.002)	(1.6E-4)	(0.001)	(0.001)	(0.001)	(5.3E-4)
Wealth	5.2E-7***	-0.002 ***	-0.006 ***	-0.007 ***	0.008 ***	0.006 ***
	(3.2E-8)	(1.1E-4)	(3.8E-4)	(4.2E-4)	(0.001)	(4E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.528 ***					
	(0.032)					
Cut point 2	-1.466 ***					
	(0.032)					
Cut point 3	-0.383 ***					
	(0.032)					
Cut point 4	1.277 ***					
	(0.032)					
F-tests:	Chi-square	p-value		Observations	219068	
Retirement variables	54.260	0.000		LogL	-253669.430	
Country dummies	8034.570	0.000		LRχ^2(31)	45621.540	
Wave dummies	388.080	0.000		Psuedo R^2	0.083	

*In the second column the coefficients are given. The third through seventh column give the marginal effects of the explanatory variables on the probability of obtaining numeracy scores of 1-5, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Models with the second health response for the discordant respondents in w1

Table B6: Results of the first-stage Ordered Probit model of self-perceived health as a function of the Instrumental variables and Demographic variables (using the second health response for the discordant respondents in wave 1)

Variable	**Coeff.**	**P(Excellent)**	**P(VeryGood)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Instrumental variables:						
Limitation with activities	0.311 ***	-0.039 ***	-0.043 ***	-0.004 ***	0.053 ***	0.034 ***
	(0.004)	(0.001)	(0.001)	(1.8E-4)	(0.001)	(4.2E-4)
Chronic diseases	0.337 ***	-0.042 ***	-0.047 ***	-0.004 ***	0.057 ***	0.036 ***
	(0.002)	(2.9E-4)	(2.8E-4)	(1.9E-4)	(3E-4)	(2.3E-4)
BMI	0.018 ***	-0.010 ***	-0.012 ***	-0.001 ***	0.014 ***	0.009 ***
	(0.001)	(3E-4)	(3.3E-4)	(5.6E-5)	(4E-4)	(2.6E-4)
Maxgrip	-0.021 ***	0.031 ***	0.035 ***	0.003 ***	-0.043 ***	-0.027 ***
	(3.2E-4)	(5E-4)	(0.001)	(1.5E-4)	(0.001)	(4.2E-4)
Demographic variables:						
Age	0.002 ***	-2.7E-4 ***	-3E-4 ***	-2.9E-5 ***	3.7E-4 ***	2.3E-4 ***
	(2.7E-4)	(3.4E-5)	(3.8E-5)	(3.8E-6)	(4.6E-5)	(2.9E-5)
Male	0.346 ***	-0.043 ***	-0.048 ***	-0.005 ***	0.059 ***	0.037 ***
	(0.007)	(0.001)	(0.001)	(2.2E-4)	(0.001)	(0.001)
Education	-0.101 ***	0.013 ***	0.014 ***	0.001 ***	-0.017 ***	-0.011 ***
	(0.002)	(2.2E-4)	(2.4E-4)	(6.2E-5)	(2.9E-4)	(1.9E-4)
Log(Income)	-0.030 ***	0.008 ***	0.009 ***	0.001 ***	-0.010 ***	-0.007 ***
	(0.002)	(4.1E-4)	(4.6E-4)	(5.6E-5)	(0.001)	(3.6E-4)
Wealth	-6E-7 ***	0.006 ***	0.007 ***	0.001 ***	-0.008 ***	-0.005***
	(3E-8)	(3E-4)	(3.4E-4)	(4.4E-5)	(4.1E-4)	(2.6E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-1.505 ***					
	(0.033)					
Cut point 2	-0.549 ***					
	(0.033)					
Cut point 3	0.741 ***					
	(0.033)					
Cut point 4	2.097 ***					
	(0.033)					
F-tests:	Chi-squared	p-value		Observations	235,545	
Instrumental variables	59,613.770	0.000		LogL	-285,247.13	
Country dummies	19,424.300	0.000		LR$\chi^2(33)$	115,206.91	
Wave dummies	146.590	0.000		Psuedo R^2	0.168	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: BMI, maxgrip, log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B7: Results of the second-stage Ordered Probit model of the predicted self-perceived health as a function of the predicted retirement variables and Demographic variables (using the second health response for the discordant respondents in wave 1)

Variable	**Coeff.**	**P(Excellent)**	**P(Very good)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Retirement variables:						
Full retirement	0.241 ***	-0.001 ***	-0.038 ***	-0.016 ***	0.039 ***	0.015 ***
	(0.007)	(4.1E-5)	(0.001)	(0.001)	(0.001)	(4.6E-4)
Partial retirement	-0.033 ***	1E-4***	0.005 ***	0.002 ***	-0.005 ***	-0.002 ***
	(0.008)	(2.6E-5)	(0.001)	(0.001)	(0.001)	(0.001)
Demographic variables:						
Age	0.039 ***	-1.2E-4***	-0.006 ***	-0.003 ***	0.006 ***	0.002 ***
	(4.1E-4)	(5.7E-6)	(6.6E-5)	(3.2E-5)	(6.6E-5)	(3E-5)
Male	-0.100 ***	3.1E-4***	0.016 ***	0.007 ***	-0.016 ***	-0.006 ***
	(0.005)	(2.2E-5)	(0.001)	(3.5E-4)	(0.001)	(3.3E-4)
Education	-0.271 ***	0.001 ***	0.042 ***	0.018 ***	-0.044 ***	-0.017 ***
	(0.002)	(4E-5)	(3.2E-4)	(2.1E-4)	(3.3E-4)	(1.7E-4)
Log(Income)	-0.072 ***	4.6E-4***	0.024 ***	0.010 ***	-0.025 ***	-0.009 ***
	(0.002)	(2.5E-5)	(0.001)	(2.6E-4)	(0.001)	(2.4E-4)
Wealth	-2.2E-6***	0.001 ***	0.028 ***	0.012 ***	-0.029 ***	-0.011 ***
	(3.6E-8)	(2.7E-5)	(4.6E-4)	(2.4E-4)	(5E-4)	(2E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.774 ***					
	(0.041)					
Cut point 2	0.086 ***					
	(0.033)					
Cut point 3	2.543 ***					
	(0.034)					
Cut point 4	4.101 ***					
	(0.035)					
F-tests:	Chi-square	p-value		Observations	235545	
Retirement variables	1136.560	0.000		LogL	-184,121.66	
Country dummies	49,796.620	0.000		LRχ^2(31)	134811.91	
Wave dummies	591.220	0.000		Psuedo R^2	0.268	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Models excluding respondents from wave 1

Table B8: Results of the first-stage Ordered Probit model of self-perceived health as a function of the Instrumental variables and Demographic variables (excluding respondents from wave 1)

Variable	**Coeff.**	**P(Excellent)**	**P(VeryGood)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Instrumental variables:						
Limitation with activities	0.307 ***	-0.037 ***	-0.043 ***	-0.004 ***	0.052 ***	0.034 ***
	(0.004)	(0.001)	(0.001)	(1.7E-4)	(0.001)	(4.6E-4)
Chronic diseases	0.335 ***	-0.040 ***	-0.047 ***	-0.004 ***	0.058 ***	0.037 ***
	(0.002)	(3E-4)	(3E-4)	(1.9E-4)	(3E-4)	(2.6E-4)
BMI	0.018 ***	-0.010 ***	-0.011 ***	-0.001 ***	0.014 ***	0.009 ***
	(0.001)	(3E-4)	(3.3E-4)	(5.6E-5)	(4E-4)	(2.8E-4)
Maxgrip	-0.021 ***	0.030 ***	0.035 ***	0.003 ***	-0.043 ***	-0.028 ***
	(3.3E-4)	(0.001)	(0.001)	(1.5E-4)	(0.001)	(4.7E-4)
Demographic variables:						
Age	0.002 ***	-2.8E-4 ***	-3.2E-4 ***	-2.8E-5 ***	3.7E-4 ***	2.6E-4 ***
	(2.9E-4)	(3.5E-5)	(4E-5)	(3.7E-6)	(4.6E-5)	(3.2E-5)
Male	0.349 ***	-0.042 ***	-0.049 ***	-0.004 ***	0.059 ***	0.039 ***
	(0.007)	(0.001)	(0.001)	(2.1E-4)	(0.001)	(0.001)
Education	-0.102 ***	0.012 ***	0.014 ***	0.001 ***	-0.017 ***	-0.011 ***
	(0.002)	(2.2E-4)	(2.5E-4)	(6.2E-5)	(2.9E-4)	(2.1E-4)
Log(Income)	-0.031 ***	0.008 ***	0.009 ***	0.001 ***	-0.011 ***	-0.007 ***
	(0.002)	(4.2E-4)	(4.9E-4)	(5.5E-5)	(0.001)	(3.9E-4)
Wealth	-5.7E-7 ***	0.006 ***	0.007 ***	0.001 ***	-0.008 ***	-0.005***
	(3E-8)	(3E-4)	(3.4E-4)	(4.3E-5)	(4.1E-4)	(2.8E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-1.602 ***					
	(0.034)					
Cut point 2	-0.644 ***					
	(0.034)					
Cut point 3	0.641 ***					
	(0.034)					
Cut point 4	2.003 ***					
	(0.034)					
F-tests:	Chi-squared	p-value		Observations	208,204	
Instrumental variables	52,774.410	0.000		LogL	-251,648.82	
Country dummies	18,211.390	0.000		LR$\chi^2(32)$	102,661.88	
Wave dummies	70.390	0.000		Psuedo R^2	0.169	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: BMI, maxgrip, log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B9: Results of the first-stage Probit models of full retirement decisions and partial retirement decisions as a function of the Instrumental variables and Demographic variables (excluding respondents from wave 1)

	Full retirement		Partial retirement	
Variable	**Coeff.**	**P(Fully Retired)**	**Coeff.**	**P(Partially Retired)**
Instrumental variables:				
Number of children	-0.085 ***	-0.018 ***	0.016 ***	0.002 ***
	(0.003)	(0.001)	(0.004)	(4.5E-4)
Number of grandchildren	0.014 ***	0.003 ***	-0.001	-5E-5
	(0.002)	(3.4E-4)	(0.002)	(2.4E-4)
Pension	1.913 ***	0.387 ***	0.118 ***	0.012 ***
	(0.010)	(0.002)	(0.012)	(0.001)
Marital status	0.436 ***	0.087 ***	0.044 ***	0.005 ***
	(0.009)	(0.002)	(0.013)	(0.001)
Demographic variables:				
Age	0.068 ***	0.014 ***	-0.028 ***	-0.003 ***
	(4.7E-4)	(9E-5)	(0.001)	(7.5E-5)
Male	0.237 ***	0.042 ***	-0.193 ***	-0.020 ***
	(0.007)	(0.002)	(0.010)	(0.001)
Education	0.033 ***	0.007 ***	0.051 ***	0.005 ***
	(0.003)	(0.001)	(0.004)	(3.7E-4)
Log(Income)	-0.066 ***	-0.028 ***	0.019 ***	0.004 ***
	(0.002)	(0.001)	(0.004)	(0.001)
Wealth	-2.1E-7***	-0.004 ***	3.4E-7***	0.003***
	(4.6E-8)	(0.001)	(4.7E-8)	(3.9E-4)
Country dummies	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes
Constant	-5.089 ***		-0.302 ***	
	(0.043)		(0.059)	
F-tests:	Chi-squared	p-value	Chi-squared	p-value
Instrumental variables	34483.720	0.000	121.840	0.000
Country dummies	4861.840	0.000	2332.690	0.000
Wave dummies	96.060	0.000	20.090	0.000
Observations	208204		208204	
LogL	-74276.2140		-40209.502	
LRχ^2(33)	139604.6800		6443.36	
Psuedo R^2	0.485		0.074	

*In the second and fourth column the coefficients are given. The third and fifth column give the marginal effects of the explanatory variables on the probability of being fully retired and partially retired, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B10: Results of the second-stage Ordered Probit model of the predicted self-perceived health as a function of the predicted retirement variables and Demographic variables (excluding respondents from wave 1)

Variable	**Coeff.**	**P(Excellent)**	**P(Very good)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Retirement variables:						
Full retirement	0.243 ***	-0.001 ***	-0.036 ***	-0.019 ***	0.040 ***	0.016 ***
	(0.008)	(4.4E-5)	(0.001)	(0.001)	(0.001)	(0.001)
Partial retirement	-0.024 ***	7E-5***	0.004 ***	0.002 ***	-0.004 ***	-0.002 ***
	(0.009)	(2.7E-5)	(0.001)	(0.001)	(0.001)	(0.001)
Demographic variables:						
Age	0.039 ***	-1.2E-4***	-0.006 ***	-0.003 ***	0.006 ***	0.002 ***
	(4.3E-4)	(6.1E-6)	(6.6E-6)	(3.8E-5)	(7E-5)	(3.3E-5)
Male	-0.090 ***	2.8E-4***	0.013 ***	0.007 ***	-0.015 ***	-0.006 ***
	(0.006)	(2.2E-5)	(0.001)	(4.4E-4)	(0.001)	(3.6E-4)
Education	-0.274 ***	0.001 ***	0.041 ***	0.021 ***	-0.045 ***	-0.018 ***
	(0.002)	(4.3E-5)	(3.2E-4)	(2.4E-4)	(3.6E-4)	(1.9E-4)
Log(Income)	-0.076 ***	5E-4***	0.024 ***	0.012 ***	-0.027 ***	-0.010 ***
	(0.002)	(2.8E-5)	(0.001)	(3.2E-4)	(0.001)	(2.6E-4)
Wealth	-2.2E-6***	0.001 ***	0.026 ***	0.013 ***	-0.029 ***	-0.011 ***
	(3.7E-8)	(2.8E-5)	(4.5E-4)	(2.3E-4)	(0.001)	(2.1E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.938 ***					
	(0.043)					
Cut point 2	-0.114 ***					
	(0.035)					
Cut point 3	2.349 ***					
	(0.035)					
Cut point 4	3.938 ***					
	(0.036)					
F-tests:	Chi-square	p-value		Observations	208204	
Retirement variables	1014.210	0.000		LogL	-163,061.83	
Country dummies	46,292.280	0.000		LRχ^2(30)	120,672.54	
Wave dummies	145.260	0.000		Psuedo R^2	0.270	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Models without the retirement variables

Table B11: Results of the second-stage Ordered Probit model of the predicted self-perceived health as a function of the Demographic variables (without the retirement variables)

Variable	**Coeff.**	**P(Excellent)**	**P(Very good)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Demographic variables:						
Age	0.047 ***	-1.4E-4***	-0.007 ***	-0.003 ***	0.008 ***	0.003 ***
	(2.7E-4)	(6.8E-6)	(4.6E-5)	(3.1E-5)	(4.4E-5)	(2.6E-5)
Male	-0.072 ***	2.1E-4***	0.012 ***	0.004 ***	-0.012 ***	-0.004 ***
	(0.005)	(1.8E-5)	(0.001)	(3.1E-4)	(0.001)	(3E-4)
Education	-0.273 ***	0.001 ***	0.043 ***	0.017 ***	-0.045 ***	-0.016 ***
	(0.002)	(4E-5)	(3.1E-4)	(2.1E-4)	(3.3E-4)	(1.7E-4)
Log(Income)	-0.067 ***	4.2E-4***	0.022 ***	0.009 ***	-0.023 ***	-0.008 ***
	(0.002)	(2.3E-5)	(0.001)	(2.4E-4)	(0.001)	(2.3E-4)
Wealth	-2.2E-6***	0.001 ***	0.029 ***	0.011 ***	-0.030 ***	-0.011 ***
	(3.6E-8)	(2.7E-5)	(4.6E-4)	(2.3E-4)	(0.001)	(1.9E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.256 ***					
	(0.036)					
Cut point 2	0.602 ***					
	(0.028)					
Cut point 3	3.044 ***					
	(0.028)					
Cut point 4	4.628 ***					
	(0.029)					
F-tests:	Chi-square	p-value		Observations	235545	
				LogL	-184,057.5	
Country dummies	52,432.430	0.000		LRχ^2(29)	134,426.87	
Wave dummies	962.740	0.000		Psuedo R^2	0.268	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B12: Results of the second-stage models of several mental health measures as a function of the Demographic variables (without the retirement variables)

Variable	Depressed (probit) Coeff.	P(Depressed)		Fluency (OLS) Coeff.	Imm. mem. (OLS) Coeff.	Delayed mem. (OLS) Coeff.
Demographic variables:						
Age	0.007 ***	0.002 ***		-0.188 ***	-0.055 ***	-0.066 ***
	(2.9E-4)	(8.9E-5)		(0.001)	(3.3E-4)	(3.9E-4)
Male	-0.424 ***	-0.128 ***		-0.013	-0.296 ***	-0.412 ***
	(0.006)	(0.002)		(0.027)	(0.006)	(0.008)
Education	-0.080 ***	-0.024 ***		1.176 ***	0.297 ***	0.333 ***
	(0.002)	(0.001)		(0.010)	(0.002)	(0.003)
Log(Income)	-0.032 ***	-0.020 ***		0.619 ***	0.138 ***	0.133 ***
	(0.002)	(0.001)		(0.019)	(0.005)	(0.006)
Wealth	-5.1E-7***	-0.013 ***		0.311 ***	0.057 ***	0.056 ***
	(4.3E-8)	(0.001)		(0.014)	(0.003)	(0.004)
Country dummies	yes	yes		yes	yes	yes
Wave dummies	yes	yes		yes	yes	yes
Constant	-0.692 ***			31.054 ***	8.104 ***	7.311 ***
	(0.032)			(0.119)	(0.028)	(0.034)
F-tests:	Chi-square			F	F	F
Country dummies	4408.760			1530.760	391.550	425.000
	(0.000)			(0.000)	(0.000)	(0.000)
Wave dummies	156.650			303.300	800.560	1015.410
	(0.000)			(0.000)	(0.000)	(0.000)
Observations	235545		Observations	235545	235545	235545
LogL	-126377.43		RMSE	6.362	1.519	1.837
LRχ^2(29)	14947.71		F(29,235513)	3385.03	2961.460	2824.400
Psuedo R^2	0.056		R^2	0.294	0.267	0.258

*In the second and third column the coefficients and the marginal effects of the probit model with depressed as the dependent variable are given. The fifth, sixth and seventh column give the coefficients of the OLS regressions with dependent variables: fluency, immediate recall memory and delayed recall memory, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B13: Results of the second-stage Ordered Probit model of numeracy scores as a function of the Demographic variables (without the retirement variables)

Variable	Coeff.	P(Num=1)	P(Num=2)	P(Num=3)	P(Num=4)	P(Num=5)
Demographics:						
Age	-0.016 ***	0.001 ***	0.002 ***	0.002 ***	-0.003 ***	-0.002 ***
	(2.3E-4)	(1.5E-5)	(3.5E-5)	(3.5E-5)	(4.6E-5)	(3.5E-5)
Male	0.287 ***	-0.015 ***	-0.042 ***	-0.042 ***	0.057 ***	0.042 ***
	(0.005)	(2.9E-4)	(0.001)	(0.001)	(0.001)	(0.001)
Education	0.222 ***	-0.012 ***	-0.033 ***	-0.033 ***	0.045 ***	0.032 ***
	(0.002)	(1.5E-4)	(2.8E-4)	(2.6E-4)	(3.5E-4)	(2.9E-4)
Log(Income)	0.041 ***	-0.005 ***	-0.013 ***	-0.013 ***	0.017 ***	0.012 ***
	(0.002)	(1.8E-4)	(4.8E-4)	(4.8E-4)	(0.001)	(4.8E-4)
Wealth	5E-7***	-0.002 ***	-0.006 ***	-0.006 ***	0.008 ***	0.006 ***
	(3.1E-8)	(1.4E-4)	(3.7E-4)	(3.7E-4)	(0.001)	(3.7E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.358 ***					
	(0.025)					
Cut point 2	-1.331 ***					
	(0.025)					
Cut point 3	-0.276 ***					
	(0.025)					
Cut point 4	1.388 ***					
	(0.025)					
F-tests:	Chi-square	p-value		Observations	235545	
				LogL	-276074.86	
Country dummies	10599.13	0.000		LR$\chi^2(29)$	53756.54	
Wave dummies	486.790	0.000		Psuedo R^2	0.089	

*In the second column the coefficients are given. The third through seventh column give the marginal effects of the explanatory variables on the probability of obtaining numeracy scores of 1-5, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Using different thresholds for the predicted probabilities of the retirement variables

Table B14: Results of the second-stage Ordered Probit model of the predicted self-perceived health as a function of the predicted retirement variables (using thresholds of 60% and 6% for the predicted probabilities of full retirement and partial retirement, respectively) and Demographic variables

Variable	**Coeff.**	**P(Excellent)**	**P(Very good)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Retirement variables:						
Full retirement	0.149 ***	-4.4E-4 ***	-0.023 ***	-0.009 ***	0.024 ***	0.014 ***
	(0.007)	(3.1E-5)	(0.001)	(4.8E-4)	(0.001)	(4.4E-4)
Partial retirement	-0.050 ***	1.5E-4***	0.008 ***	0.003 ***	-0.008 ***	-0.002 ***
	(0.008)	(2.6E-5)	(0.001)	(0.001)	(0.001)	(5E-4)
Demographic variables:						
Age	0.041 ***	-1.2E-4***	-0.006 ***	-0.003 ***	0.007 ***	0.002 ***
	(4E-4)	(6E-6)	(6.6E-5)	(3.2E-5)	(6.5E-5)	(2.9E-5)
Male	-0.099 ***	2.9E-4***	0.016 ***	0.006 ***	-0.016 ***	-0.006 ***
	(0.005)	(2.1E-5)	(0.001)	(3.4E-4)	(0.001)	(3.2E-4)
Education	-0.272 ***	0.001 ***	0.043 ***	0.017 ***	-0.045 ***	-0.016 ***
	(0.002)	(4E-5)	(3.2E-4)	(2.1E-4)	(3.3E-4)	(1.7E-4)
Log(Income)	-0.070 ***	4.3E-4***	0.023 ***	0.009 ***	-0.024 ***	-0.009 ***
	(0.002)	(2.4E-5)	(0.001)	(2.5E-4)	(0.001)	(2.3E-4)
Wealth	-2.2E-6***	0.001 ***	0.028 ***	0.011 ***	-0.029 ***	-0.011 ***
	(3.6E-8)	(2.7E-5)	(4.6E-4)	(2.3E-4)	(5E-4)	(1.9E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.617 ***					
	(0.040)					
Cut point 2	0.253 ***					
	(0.033)					
Cut point 3	2.702 ***					
	(0.033)					
Cut point 4	4.284 ***					
	(0.034)					
F-tests:	Chi-square	p-value		Observations	235545	
Retirement variables	452.810	0.000		LogL	-183,830.820	
Country dummies	49,946.99	0.000		LR$\chi^2(31)$	134,880.24	
Wave dummies	994.400	0.000		Psuedo R^2	0.268	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B15: Results of the second-stage models of several mental health measures as a function of the predicted retirement variables (using thresholds of 60% and 6% for the predicted probabilities of full retirement and partial retirement, respectively) and Demographic variables

	Depressed (probit)			Fluency (OLS)	Imm. mem. (OLS)	Delayed mem. (OLS)
Variable	**Coeff.**	**P(Depressed)**		**Coeff.**	**Coeff.**	**Coeff.**
Retirement variables:						
Full retirement	-0.044	-0.013		0.257 ***	0.113 ***	0.101 ***
	(0.009)	(0.003)		(0.039)	(0.009)	(0.011)
Partial retirement	0.101 ***	0.031 ***		0.219 ***	-0.127 ***	-0.056***
	(0.010)	(0.003)		(0.044)	(0.011)	(0.013)
Demographic variables:						
Age	0.011 ***	0.003 ***		-0.192 ***	-0.062 ***	-0.071 ***
	(4.5E-4)	(1.4E-4)		(0.002)	(0.001)	(0.001)
Male	-0.400 ***	-0.121 ***		0.002	-0.334 ***	-0.434 ***
	(0.006)	(0.002)		(0.028)	(0.007)	(0.008)
Education	-0.084 ***	-0.025 ***		1.166 ***	0.302 ***	0.335 ***
	(0.002)	(0.001)		(0.010)	(0.002)	(0.003)
Log(Income)	-0.033 ***	-0.021 ***		0.596 ***	0.139 ***	0.130 ***
	(0.002)	(0.001)		(0.020)	(0.005)	(0.006)
Wealth	-5.6E-7***	-0.014 ***		0.309 ***	0.062 ***	0.060 ***
	(4.3E-8)	(0.001)		(0.014)	(0.003)	(0.004)
Country dummies	yes	yes		yes	yes	yes
Wave dummies	yes	yes		yes	yes	yes
Constant	-0.915 ***			31.131 ***	8.503 ***	7.575 ***
	(0.038)			(0.153)	(0.036)	(0.044)
F-tests:	Chi-square			F	F	F
Retirement variables	141.370			32.350	152.710	52.050
	(0.000)			(0.000)	(0.000)	(0.000)
Country dummies	4347.470			1385.760	361.880	365.360
	(0.000)			(0.000)	(0.000)	(0.000)
Wave dummies	140.880			301.010	838.770	1032.190
	(0.000)			(0.000)	(0.000)	(0.000)
Observations	235545		Observations	235545	235545	235545
LogL	-126,306.730		RMSE	6.362	1.5198	1.836
LRχ^2(31)	15,089.120		F(31,235513)	3169.570	2783.820	2646.690
Psuedo R^2	0.056		R^2	0.294	0.268	0.258

*In the second and third column the coefficients and the marginal effects of the probit model with depressed as the dependent variable are given. The fifth, sixth and seventh column give the coefficients of the OLS regressions with dependent variables: fluency, immediate recall memory and delayed recall memory, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B16: Results of the second-stage Ordered Probit model of numeracy scores as a function of the predicted retirement variables (using thresholds of 60% and 6% for the predicted probabilities of full retirement and partial retirement, respectively) and Demographic variables

Variable	**Coeff.**	**P(Num=1)**	**P(Num=2)**	**P(Num=3)**	**P(Num=4)**	**P(Num=5)**
Retirement:						
FR	0.033 **	-0.002 **	-0.005 **	-0.005 **	0.007 **	0.005**
	(0.007)	(3.5E-4)	(0.001)	(0.001)	(0.001)	(0.001)
PR	-0.090 ***	0.005 ***	0.013 ***	0.013 ***	-0.018 ***	-0.013 ***
	(0.008)	(4E-4)	(0.001)	(0.001)	(0.002)	(0.001)
Demographics:						
Age	-0.019 ***	0.001 ***	0.003 ***	0.003 ***	-0.004 ***	-0.003 ***
	(3.6E-4)	(2.1E-5)	(5.3E-5)	(5.4E-5)	(7.2E-5)	(5.4E-5)
Male	0.265 ***	-0.014 ***	-0.040 ***	-0.039 ***	0.053 ***	0.039 ***
	(0.005)	(2.9E-4)	(0.001)	(0.001)	(0.001)	(0.001)
Education	0.226 ***	-0.012 ***	-0.033 ***	-0.033 ***	0.045 ***	0.033 ***
	(0.002)	(1.5E-4)	(2.8E-4)	(2.6E-4)	(3.5E-4)	(2.9E-4)
Log(Income)	0.042 ***	-0.005 ***	-0.013 ***	-0.013 ***	0.018 ***	0.013 ***
	(0.002)	(1.8E-4)	(4.9E-4)	(4.9E-4)	(0.001)	(4.9E-4)
Wealth	5.4E-7***	-0.002 ***	-0.006 ***	-0.007 ***	0.009 ***	0.006 ***
	(3.1E-8)	(1.4E-4)	(3.7E-4)	(3.7E-4)	(0.001)	(3.7E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.549 ***					
	(0.030)					
Cut point 2	-1.521 ***					
	(0.030)					
Cut point 3	-0.464 ***					
	(0.030)					
Cut point 4	1.198 ***					
	(0.030)					
F-tests:	Chi-square	p-value		Observations	235545	
Retirement variables	174.110	0.000		LogL	-275,987.800	
Country dummies	9826.350	0.000		LR$\chi^2(31)$	53,930.650	
Wave dummies	523.240	0.000		Psuedo R^2	0.089	

*In the second column the coefficients are given. The third through seventh column give the marginal effects of the explanatory variables on the probability of obtaining numeracy scores of 1-5, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Models including the interaction term between retirement and age

Table B17: Results of the second-stage Ordered Probit model of the predicted self-perceived health as a function of the predicted retirement variables and Demographic variables (including interaction between retirement and age)

Variable	**Coeff.**	**P(Excellent)**	**P(Very good)**	**P(Good)**	**P(Fair)**	**P(Poor)**
Retirement variables:						
Full retirement	0.729 ***	-0.002 ***	-0.115 ***	-0.047 ***	0.120 ***	0.044 ***
	(0.054)	(1.9E-4)	(0.008)	(0.003)	(0.009)	(0.003)
Partial retirement	0.143 ***	-4.2E-4***	-0.022 ***	-0.009 ***	0.023 ***	0.009 ***
	(0.049)	(1.5E-4)	(0.008)	(0.003)	(0.008)	(0.003)
Full retirement*age	-0.008 ***	-0.001 ***	-0.039 ***	-0.006 ***	0.036 ***	0.009 ***
	(0.001)	(3.7E-5)	(0.001)	(0.001)	(0.001)	(0.001)
Partial retirement*age	-0.003 ***	3.3E-5***	0.004 ***	0.006 ***	-0.006 ***	-0.003 ***
	(0.001)	(2.8E-5)	(0.001)	(0.001)	(0.002)	(0.001)
Demographic variables:						
Age	0.045 ***	-1.2E-4***	-0.006 ***	-0.002 ***	0.006 ***	0.002 ***
	(0.001)	(6.3E-6)	(8.3E-5)	(5E-5)	(6.9E-5)	(3E-5)
Male	-0.099 ***	2.9E-4***	0.015 ***	0.006 ***	-0.016 ***	-0.006 ***
	(0.005)	(2.1E-5)	(0.001)	(3.5E-4)	(0.001)	(3.2E-4)
Education	-0.272 ***	0.001 ***	0.043 ***	0.017 ***	-0.045 ***	-0.016 ***
	(0.002)	(4E-5)	(3.2E-4)	(2.1E-4)	(3.3E-4)	(1.7E-4)
Log(Income)	-0.073 ***	4.5E-4***	0.024 ***	0.010 ***	-0.025 ***	-0.009 ***
	(0.002)	(2.4E-5)	(0.001)	(2.5E-4)	(0.001)	(2.3E-4)
Wealth	-2.2E-6***	0.001 ***	0.028 ***	0.011 ***	-0.029 ***	-0.011 ***
	(3.6E-8)	(2.7E-5)	(4.6E-4)	(2.3E-4)	(5E-4)	(1.9E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.362 ***					
	(0.060)					
Cut point 2	0.522 ***					
	(0.055)					
Cut point 3	2.979 ***					
	(0.056)					
Cut point 4	4.562 ***					
	(0.056)					
F-tests:	Chi-square	p-value		Observations	235545	
Retirement variables	1213.31	0.000		LogL	-183449.17	
Country dummies	49687.46	0.000		LR$\chi^2(31)$	135643.54	
Wave dummies	983.06	0.000		Psuedo R^2	0.270	

*In the second column the coefficients are given. The third, fourth, fifth, sixth and seventh column give the marginal effects of the explanatory variables on the probability of excellent health, very good health, good health, fair health and poor health, respectively. The marginal effects of the explanatory variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Table B18: Results of the second-stage models of several mental health measures as a function of the predicted retirement variables and Demographic variables (including interaction between retirement and age)

Variable	Depressed (probit) Coeff.	P(Depressed)		Fluency (OLS) Coeff.	Imm. mem. (OLS) Coeff.	Delayed mem. (OLS) Coeff.
Retirement variables:						
Full retirement	-0.010	-0.003		3.178 ***	0.954 ***	1.152 ***
	(0.061)	(0.019)		(0.285)	(0.068)	(0.082)
Partial retirement	1.177 ***	0.355 ***		-3.272 ***	-1.342 ***	-1.035***
	(0.057)	(0.017)		(0.258)	(0.062)	(0.075)
Full retirement*age	0.001	0.015		-0.051 ***	-0.015 ***	-0.018 ***
	(0.001)	(0.003)		(0.004)	(0.001)	(0.001)
Partial retirement*age	-0.018 ***	-0.003 ***		0.056 ***	0.020 ***	0.016 ***
	(0.001)	(0.003)		(0.004)	(0.001)	(0.001)
Demographic variables:						
Age	0.012 ***	0.002 ***		-0.163 ***	-0.054 ***	-0.061 ***
	(0.001)	(1.5E-4)		(0.004)	(0.001)	(0.001)
Male	-0.408 ***	-0.123 ***		0.014	-0.323 ***	-0.425 ***
	(0.006)	(0.002)		(0.028)	(0.007)	(0.008)
Education	-0.082 ***	-0.025 ***		1.161 ***	0.299 ***	0.333 ***
	(0.002)	(0.001)		(0.010)	(0.002)	(0.003)
Log(Income)	-0.036 ***	-0.023 ***		0.615 ***	0.42 ***	0.133 ***
	(0.002)	(0.001)		(0.020)	(0.005)	(0.006)
Wealth	-4.8E-7***	-0.012 ***		0.289 ***	0.054 ***	0.052 ***
	(4.4E-8)	(0.001)		(0.014)	(0.003)	(0.003)
Country dummies	yes	yes		yes	yes	yes
Wave dummies	yes	yes		yes	yes	yes
Constant	-1.073 ***			29.800 ***	8.185 ***	7.071 ***
	(0.063)			(0.281)	(0.067)	(0.081)
F-tests:	Chi-square			F	F	F
Retirement variables	605.200			163.860	349.340	208.500
	(0.000)			(0.000)	(0.000)	(0.000)
Country dummies	4133.780			1404.250	361.590	358.690
	(0.000)			(0.000)	(0.000)	(0.000)
Wave dummies	161.550			276.860	766.630	972.070
	(0.000)			(0.000)	(0.000)	(0.000)
Observations	235545		Observations	235545	235545	235545
LogL	-126074.46		RMSE	6.353	1.515	1.834
LRχ^2(31)	15553.64		F(31,235513)	3002.820	2660.240	2516.070
Psuedo R^2	0.058		R^2	0.296	0.272	0.261

*In the second and third column the coefficients and the marginal effects of the probit model with depressed as the dependent variable are given. The fifth, sixth and seventh column give the coefficients of the OLS regressions with dependent variables: fluency, immediate recall memory and delayed recall memory, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%*

Table B19: Results of the second-stage Ordered Probit model of numeracy scores as a function of the predicted retirement variables and Demographic variables (including interaction between retirement and age)

Variable	**Coeff.**	**P(Num=1)**	**P(Num=2)**	**P(Num=3)**	**P(Num=4)**	**P(Num=5)**
Retirement variables:						
Full retirement	0.085 **	-0.005 **	-0.013 **	-0.013 **	0.017 **	0.012**
	(0.048)	(0.003)	(0.007)	(0.007)	(0.010)	(0.007)
Partial retirement	-0.396 ***	0.021 ***	0.058 ***	0.058 ***	-0.079 ***	-0.058 ***
	(0.044)	(0.002)	(0.006)	(0.006)	(0.009)	(0.006)
Full retirement*age	-0.001 **	0.001 **	0.001 **	1.4E-4**	-0.001 **	-2.2E-4**
	(0.001)	(4.5E-4)	(0.001)	(0.001)	(0.001)	(0.001)
Partial retirement*age	0.005 ***	0.001 ***	0.005 ***	0.009 ***	-0.007 ***	-0.009 ***
	(0.001)	(4.8E-4)	(0.001)	(0.001)	(0.002)	(0.001)
Demographics:						
Age	-0.018 ***	0.001 ***	0.002 ***	0.002 ***	-0.003 ***	-0.002 ***
	(0.001)	(2.1E-5)	(5.5E-5)	(6.4E-5)	(7.4E-5)	(6.4E-5)
Male	0.272 ***	-0.014 ***	-0.040 ***	-0.040 ***	0.055 ***	0.040 ***
	(0.005)	(2.9E-4)	(0.001)	(0.001)	(0.001)	(0.001)
Education	0.225 ***	-0.012 ***	-0.033 ***	-0.033 ***	0.045 ***	0.033 ***
	(0.002)	(1.5E-4)	(2.9E-4)	(2.6E-4)	(3.5E-4)	(2.9E-4)
Log(Income)	0.042 ***	-0.005 ***	-0.013 ***	-0.013 ***	0.018 ***	0.013 ***
	(0.002)	(1.8E-4)	(4.9E-4)	(4.9E-4)	(0.001)	(4.9E-4)
Wealth	5.1E-7***	-0.002 ***	-0.006 ***	-0.006 ***	0.008 ***	0.006 ***
	(3.1E-8)	(1.4E-4)	(3.7E-4)	(3.7E-4)	(0.001)	(3.7E-4)
Country dummies	yes	yes	yes	yes	yes	yes
Wave dummies	yes	yes	yes	yes	yes	yes
Cut point 1	-2.525 ***					
	(0.050)					
Cut point 2	-1.496 ***					
	(0.050)					
Cut point 3	-0.439 ***					
	(0.050)					
Cut point 4	1.223 ***					
	(0.050)					
F-tests:	Chi-square	p-value		Observations	235545	
Retirement variables	179.120	0.000		LogL	-275985.300	
Country dummies	9579.030	0.000		LR$\chi^2(31)$	53935.660	
Wave dummies	486.760	0.000		Psuedo R^2	0.089	

*In the second column the coefficients are given. The third through seventh column give the marginal effects of the explanatory variables on the probability of obtaining numeracy scores of 1-5, respectively. The marginal effects of the variables: log(income) and wealth are given in terms of one standard deviation increase. The remaining control variables are given in terms of increase in one unit. The standard errors for the marginal effects are calculated via Stata. Stata obtains the standard errors of the average marginal effects by applying a Jacobian matrix to the estimated variance matrix of the fitted model parameters. In order to avoid multicollinearity, the dummy variables for Austria and Wave 1 are suppressed. The standard errors of the coefficients are given in parenthesis below the value of the coefficients. '***', '**' and' *' refer to the significant level of the coefficient at 1%, 5% and 10%, respectively.*

Figure B1: The effects of retirement variables on cognitive measures for different ages (calculated using results from table B18)