WORKING
PAPER

Retirement Effects on Health In Europe

NORMA B. COE, GEMA ZAMARRO

WR-588

June 2008

This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the NICHD funded RAND Population Research Center (R24HD050906).

RAND LABOR AND POPULATION

Retirement Effects on Health In Europe

Norma B. Coe[1]
Tilburg University & Netspar

Gema Zamarro[2]
RAND Corporation

June 2008

[1]E-mail: N.Coe@uvt.nl.

[2]E-mail: gzamarro@rand.org. We gratefully acknowledge funding for this project from the Marie Curie Incoming International Fellowship (Coe), Roybal Center for Health Policy Simulation funded by the National Institute on Aging (5P30 AG024968-02) (Zamarro) and grant number 1 P01 AG022481-01A1 from the National Institute of Ageing (Zamarro). We thank Anamaria Camacho for able research assistance. We are also grateful to Arthur van Soest for replying all our questions about vignettes and to Arie Kapteyn for all his support. We would like to thank participants of seminars from RAND Corporation, Tilburg University, the Workshop on the Economics of Ageing (Turin), and the Netspar Health Workshop (Maastricht) for very helpful comments. Errors, if any, are our own.

Abstract

What are the health impacts of retirement? As talk of raising retirement ages in pensions and social security schemes continues around the world, it is important to know both the costs and benefits for the individual as well as the governments' budgets. In this paper we use the Survey of Health, Aging and Retirement in Europe (SHARE) dataset to address this question in a multi-country setting. We use country-specific early and full retirement ages as an instrument for retirement behavior in a regression discontinuity design approach. These statutory retirement ages clearly induce retirement, but are not related to an individual's health.

Exploiting the discontinuities in retirement behavior across countries, we find significant evidence that retirement has a health-preserving effect on overall general health. Our estimates indicate that retirement leads to a 0.35 decrease in the probability of reporting to be in fair, bad or very bad health, and an almost 1 standard deviation drop in the health index (in the direction of better health). In addition, these results are being driven by retirement at age 65. We find no evidence of a health-preserving effect of retiring at younger ages.

1 Introduction

The notion that retirement harms health is an old, and persistent hypothesis (See Minkler (1981) for a review). Many argue that retirement itself is a stressful event (Carp (1967), Eisdorfer and Wilkie (1977), MacBride (1976), Sheppard (1976)). Retirement can also lead to a break with support networks and friends, and may be accompanied by emotional or mental impacts of "loneliness," "obsolesce," or "feeling old" (Bradford (1979), MacBride (1976)). Others believe that retirement is a health-preserving life change. Anecdotal evidence suggests that many discussions about the retirement decision include the idea that work is taxing to the individual, thus retirement would remove this stress and preserve the health of the retiree (Ekerdt et al. (1983)).

Despite the long-standing debate, there is little conclusive empirical evidence thus far. The inherent problem is that retirement is often a choice, and is often based on health characteristics before retirement. Many of the early studies do not address this, thus they can only infer correlation, not causality. Compounding the problem is that some of the studies find a positive correlation with health (Thomson et al. (1958)), no correlation with health (Carp (1967), Atchley (1976), Kasl (1980), Rowland (1977), Haynes et al. (1978), Niemi (1980), Adams et al. (1981)), or a negative correlation with health (Casscells et al. (1980), Gonzales (1980)).

A few recent papers try to address the endogeneity of the retirement decision in examining future health. Charles (2002) and Neuman (2008) use age specific retirement incentives provided by the US Social Security regulations as instrumental variables in the US context. Coe and Lindeboom (2008) also use early retirement window offers as an instrument. The results from these papers combined indicate that retirement has a positive effect on subjective measures of health but no effect on objective measures of health in the United States. There is no a priori reason to assume that findings from the US situation will hold for European countries considering the numerous differences in the labor markets, health insurance, and social policies.

Kerkhofs & Lindeboom (1997) assess the effects of work history on the health status of older Dutch workers using fixed effects regressions. This accounts for time invariant factors that may confound the results, but it does not control for time varying factors such as a sudden change in the individual environment. Their results suggest that health deteriorates with increased work effort and that increasing retirement ages may negatively influence later life health outcomes. Lindeboom et al. (2002) use a fixed effect control function to assess the effect of life events, such as retirement, on mental health of older individuals, also in the Dutch setting. They try to control for all transitory changes as well as individual fixed-effects. They find no statistical effect of loss of work on mental health two-years later. Their approach does not address any physical health effects of retirement. In fact, they control for all physical health deterioration that is observed in the data. Thus, this result may fail to measure the true cumulative impact of retirement on overall health.

Bound and Waidmann (2007) examine the health effects of retirement in the UK using one wave of the English Longitudinal Study of Aging (ELSA). They examine both self-reported measures of health and objective measures of health measured through blood samples. They find some evidence of a positive health effect of retirement, although temporarily, for men, and no corresponding relationship for women.

We examine the effect retirement has on health in a multi-country setting using the Survey of Health and Retirement in Europe (SHARE) dataset. In addition to demographic information, the survey collects detailed information concerning retirement behavior. The health information is rich, and includes self-reported health, diagnoses of different diseases, the Euro-D depression index, as well as newer, more powerful predictors of mortality such as grip strength. We have supplemented the data with information on early and full statutory retirement ages in 11 different countries.

In addition to the new setting, we also use an empirical methodology that allows us to get at the causal relationship between retirement and health. Due to the lack of long panel data that is comparable across many countries, we use a single cross-section of data and use the differing retirement ages across countries as exogenous instruments for the retirement decision. Use of these kind of instruments, under some assumptions, can be viewed as an application of a regression discontinuity design. Unlike single-country analysis, we can exploit the exogenous variation in retirement policies to explore the effect of retirement on health at different ages, not just age 65 as in UK and US studies.

The paper proceeds as follows. Section 2 discusses the empirical model, while the data and the definition of key variables are introduced in section 3. In section 4 we present the results and conclude in section 5.

2 Empirical Approach

2.1 IV and Regression discontinuity designs

We aim to determine the effect of the binary decision of being retired ($D_i = 1$) on a measure of health (Y_i) (i.e. general health measures, cognitive ability or depression measures). The traditional approach consists on estimating the following equation by OLS:

$$Y_i = \beta + \alpha D_i + u_i$$

A selection problem may arise; namely D_i can be correlated with the unobservables u_i. For example, this could happen if people self-select into retirement earlier based on their gains to being retired, i.e. those who enjoy their jobs the least retire early to be happier, or those with the most physically demanding jobs retire earlier to relieve themselves of the daily strain. If this is the case OLS estimates of α are not consistent.

The econometric approach we use to solve this problem exploits the fact that the regressor of interest (retirement) is partly determined by a known discontinuous function of an observed covariate (age). In particular, we use early and full retirement ages in different countries in Europe to construct IV estimates of retirement effects. Although retirement and the instruments derived are both a function of age themselves, this function is nonlinear and non-monotonic. We can therefore control for a wide range of smooth age effects in health when using early and full retirement ages as instruments.

This sort of identification argument has a long history in social science and can be viewed as an application of Regression Discontinuity (RD) designs for evaluating the effect of retirement on health. Campbell (1969), analyzed the impact of student scholarships on career aspirations, exploiting the fact that the awards are made only when a student's test scores exceed a threshold; Angrist and Lavy (1999) study the effect of class size on student test scores, taking advantage of the "Maimonides Rule", which stipulates that the class be split when it reaches a specific threshold size; Van der Klaauw (2003) estimates the effect of financial aid offers on the student's decision to attend a college, exploiting the identifying information provided by a discontinuity in the administrative rule that relates the aid to the student's scores.

2.1.1 Regression discontinuity design (RD)

RD is a quasi-experimental design with the main characteristic that the probability of receiving a policy or treatment changes discontinuously as a function of one or more underlying variables. Let S_i be the age of the individual and $\bar{S}$ the full retirement age. The identification strategy exploits the fact that around the threshold $\bar{S}$ individuals are assigned to treatment randomly. Thus, at $\bar{S}$ the effect of retirement is identified by the discontinuity caused by D. There are two main types of discontinuity designs considered in the literature: the sharp design and the so-called fuzzy design.

In the sharp design, being above the threshold $(\bar{S})$ completely determines the treatment:

$$D_i = D(S_i) = 1\left\{S_i \geq \bar{S}\right\}$$

Note that S_i can affect directly Y_i too. In that case, it is possible that D_i affects Y_i even if the causal effect is zero. This is a key difference between RD sharp design and randomization. A random experiment guarantees that treatment and controls are similar whereas a RD sharp design they are different; at least in their value of S_i.

As a first approximation assume that the effect of retirement on health is homogenous; α is constant among individuals. Thus, this is the special case of selection on observables where we should control for the dependence between Y_i and S_i. One approach to estimate the causal effect is to specify the dependence $(E\left(u|D,S\right) = K(S))$, include it as a "control function" and obtain estimates of the following equation by OLS (Heckman and Robb (1985)):

$$Y_i = \beta + \alpha D_i + K(S_i) + w_i \tag{1}$$

If the control function is well specified, this procedure will lead to consistent estimators of α. That is because in this specification $E(u|D,S) = E(u|S)$ since S is the only determinant of treatment status, it will capture any correlation between D and u. Goldberger (1972) and Cain (1975) considered $K(S_i)$ a linear function. In this case, α will be estimated by the distance between the two linear parallel regression lines at the cutoff point, which equals the difference in the intercepts of the two lines.

In the case of a fuzzy design, a threshold does not perfectly determine treatment exposure, but it creates a discontinuity in its probability. This is the appropriate case in this application, since the individual retirement decision is not completely determined by eligibility for retirement programs, but being of eligible retirement age does change the probability of retirement. In this case, the selection process is not totally determined by S:

$$D_i = E(D_i|S_i) + e_i$$

where e_i are a set of unobservables. Estimates of α using equation (1) would be biased if e_i is correlated with u_i even given S_i. This problem can be solved estimating the following equation instead:

$$Y_i = \beta + \alpha E(D_i|S_i) + K(S_i) + w_i$$

In order to do so, we employ a two step estimation procedure. In the first step, we estimate:

$$E(D_i|S_i) = f(S_i) + \gamma 1 \left\{ S_i \geq \bar{S} \right\} \tag{2}$$

where $f(S_i)$ is a general function of S continuous at $\bar{S}$ and γ measures the discontinuity in the propensity score function at $\bar{S}$. Note, that to guarantee identification of $\hat{\alpha}$ we need $\gamma \neq 0$. We use the predicted values from the first step in order to estimate:

$$Y_i = \beta + \alpha \widehat{E(D_i|S_i)} + K(S_i) + w_i \tag{3}$$

In this paper, we consider a parametric estimation approach. Thus, we assume the same functional form for $K(S_i)$ and $f(S_i)$. Both depend on S_i and the square of S_i in a linear way. In this case, this two-stage procedure is equivalent as two-stage least squares where $1 \left\{ S_i \geq \bar{S} \right\}$ and the terms in $f(S_i)$ are used as instruments. To recap, being above early and full retirement ages are used as external instruments for retirement decisions.

2.1.2 Heterogeneity

In this application, we are concerned that retirement might have different effects on health for different individuals. For example, retirement from a job that requires strenuous physical labor might effect ones health different than retiring from a desk job. Due to data limitations, we cannot control for many characteristics of the job an individual retires from, and thus we have limited ability to identify the underlying distribution of health effects. While the individual effects are important for individual decisions, policy makers should also be interested in the average effect of retirement on health, which our method allows us to identify.

More formally, let Y_{1i} denote the health of an individual when he is retired ($D_i = 1$) and Y_{0i} his health in case of not being retired ($D_i = 0$). The evaluation problem arises because people either are retired or not and no individual is observed in both states at the same time. We would like to know what is the effect of retirement on health ($Y_{1i} - Y_{0i}$) but what we observe is:

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i}$$

So far we assumed that the effect was homogeneous among individuals (($Y_{1i} - Y_{0i}$) is constant $\forall\ i$). If there is a distribution of effects in the population, the average effect of retirement at the threshold $\bar{S}$ ($E(Y_{1i} - Y_{0i})|S_i = \bar{S}$) is nonparametrically identified using the regression discontinuity approach[1].

3 Data

For this paper we use release 2 of the Survey of Health and Retirement in Europe (SHARE) dataset[2]. SHARE is a new multidisciplinary and cross-national individual-level dataset on health, socioeconomic status, social and

[1] See Hahn, Todd and Van der Klaauw (2001).

[2] For more information about SHARE see http://www.share-project.org/. This paper uses data from release 2.0.1 of SHARE 2004. The SHARE data collection has been primarily funded by the European Commission through the 5th framework programme (project QLK6-CT-2001- 00360 in the thematic programme Quality of Life). Additional funding came from the US National Institute on Ageing (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04-064). Data collection for wave 1 was nationally funded in Austria (through the Austrian Science Foundation, FWF), Belgium (through the Belgian Science Policy Office), France (through CNAM, CNAV, COR, Drees, Dares, Caisse des Dépôts et Consignations et le Commissariat Général du Plan) and Switzerland (through BBW/OFES/UFES. The SHARE data collection in Israel was funded by the US National Institute on Aging (R21 AG025169), by the German-Israeli Foundation for Scientific Research and Development (G.I.F.), and by the National Insurance Institute of Israel. Further support by the European Commission through the 6th framework program (projects SHARE-I3, RII-CT- 2006-062193, and COMPARE, CIT5-CT-2005-028857) is gratefully acknowledged. For methodological details see Boersch-Supan and Juerges (2005).

family networks. SHARE is unique in three important ways: multidisciplinary, cross-country comparability, and eventually longitudinal. The data we use is the first wave of the survey, 2004, containing information on approximately 31,000 individuals over the age of 50, in the following countries: Austria, Germany, Sweden, the Netherlands, Spain, Italy, France, Denmark, Greece, Switzerland, and Belgium[3]. In addition to demographic information, there is also a wealth of socioeconomic variables (labor situation, wealth, consumption, housing, etc.) and detailed health status information, including self-reported health, diagnoses of different diseases, the Euro-D depression index, as well as newer, more powerful predictors of mortality such as grip strength and walking speed.

There are very few sample restrictions necessary for this analysis. First we eliminate incomplete survey records. Second, we keep those individuals who are between 50 and 69 years old. We eliminate those individuals who have never worked, and those who have not worked since age 50, either due to individual choice or physical or mental limitations. Finally, we limit our analysis to men, since we are less worried about the potential for cohort effects in the characteristics of the working population for men than for women. Otherwise this data set is perfectly suited to our needs, sampling near- and post-retirement aged individuals in a variety of institutional settings. The final sample consists on 5,282 men.

We supplemented this dataset with the early and full retirement ages in place in each country in 2004[4] (see table 1). Retirement ages vary by country, and sometimes by gender, by as much as 8 years. The earliest age for retirement is 57, and the latest is age 65.

3.1 Measuring Health

3.1.1 Self-Reported Health and Health Index

Health is a difficult concept to measure in survey data. Objective measures of health often only ask about specific diseases, and thus are incomplete. Self-

[3]SHARE data also collects information in Israel that was not considered in our analysis.

[4]The main source for this data was Natali (2004), but was supplemented with information from OECD (2003), the Bartelsmann foundation, Sundén (2004), Preesman (2006), and OECD (2005). Slightly differences can be found between our retirement ages and those from other OECD publications (for example, OECD (2005)). This is because these publications describe the current law, while what we are concerned with is the law that was in place when these individuals were facing the retirement decision.

reported general health questions may provide a more complete picture of ones overall well-being, and consistently is found to be a significant predictor of mortality even after one controls for many more objective measures of health. In SHARE, respondents are asked to rate their health on a 5-point scale - very good, good, fair, bad and very bad[5]. We then condense these responses to a 2-point scale, bad health (poor, bad, or very bad) and good health (good and very good).

Self-reported health may have other problems, though. It may suffer from justification bias, and thus may be itself endogenous to retirement behavior. There is a large literature discussing how to create a reliable, meaningful, measure of health, but it varies widely conceptually and methodologically. We will follow Bound et al. (1999) to create a health index to address some of these issues.

Bound et al. (1999) predict self-reported health using all available objective measures of health in the Health and Retirement Study (HRS). This "health stock" variable ameliorates the reporting bias in the self-reported health variable, and does not suffer from multi-collinearity problems that arise when including both objective and subjective measures of health as controls[6]. We create a health index by estimating

$$S_i = \beta + \alpha L_i + w_i \tag{4}$$

where S is self-reported health, ranging on a five-point scale from very good (1) to very bad (5) health. L includes individual objective measures of health. The SHARE dataset has extensive health measures. We include variables measuring limitations to activities due to health reasons, the number of limitations in Activities of Daily Living (ADLs)[7], the number of limitations in the Instrumental Activities of Daily Living (IADLs)[8], the number of chronic

[5]A second version of the 5-point scale is also recorded. This version includes the options excellent, very good, good, fair and poor. This latter version is used in American surveys collecting health measures such as the Health and Retirement Study. We do not use this second version in the paper.

[6]Bound et al (1999) also use changes in the health stock, or "health shocks" to predict retirement behavior, but as of now, we do not have a panel dataset for SHARE.

[7]ADL's measure limitiations in the following skills: Bathing; Dressing; Toilet Use; Transferring (in and out of bed or chair); Urine and Bowel Continence; Eating.

[8]IADL's measure the following skills: Use of the telephone; Traveling via car or public transportation; Food or clothes shopping; Meal preparation; Housework; Medication use (Preparing and taking correct dose); Management of money (write checks, pays bills).

diseases[9], the number of limitations with mobility, including arm function and fine motor function[10], grip strength, indicators for over-weight and obesity as measured by body mass index (BMI), a depression index as measured by the Euro-D scale[11], number of chronic symptoms[12], physical inactivity, and hospitalizations within the last year.

Different countries may have different "norms" in reporting self-assessed measures of health (Kapteyn et al. (2008)). Thus we alter the Bound approach slightly, estimating equation (4) separately for each of the 11 countries in our sample. This controls for different means and different cut-points in self-reported health across countries. Table 2 shows the ordered probit results. A positive coefficient indicates that the condition leads to reporting of worse health. The coefficients are intuitively consistent; more health problems lead to lower reported health. Reporting that one has limitations with activities and the number of chronic illnesses are consistently significant regressors for self-reported health. Interestingly, despite the concern about different reporting standards used across countries, in most cases the estimates for the marginal effect of the various health conditions on self-reported health are not statistically different between countries.

In order to control for country differences in reporting, we create an international health index by predicting self-reported health using Germany's country-specific coefficients, regardless of where the individual might live. Implicitly, this technique assumes that countries vary on reporting self-reported health only at the mean, and not the cut-points. This also implicitly assumes that the probability of diagnosis, given a certain condition, is the same in all countries. While we agree that these are strong assumptions, the ability to

[9]The numer of chronic diseases is a count of the following diseases an individual might have: heart problems; high blood pressure; high cholesterol; crebral vascular disease; diabetes; lung diseases; asthma; arthritis; osteoporosis; cancer; stomach ulcer; parkinson disease; cataracts; hip fracure or femoral fracture.

[10]The mobility measure includes information on the ability to walk 100 meters, the ability to sit in a chair for 2 hours or more, and the ability to get up from a chair unaided after sitting.

[11]The euro-d scale runs from 0-12, measuring the number of depression symptoms including: depression; pessimism; suicidality; guilt; sleep; interest; irritability; appetite; fatigue; concentration; enjoyment; and tearfulness.

[12]The number of chronic symptoms is a count variable for individuals reporting any of the following conditions in the previous six months: pain in your back, knees, hips or any other joint; heart trouble; breathlessness; persistent cough; swollen legs; sleeping problems.

relax these assumptions is limited[13].

Table 3 shows the average predicted health index for each country using both their own estimated coefficients and when using Germany as reference. Using the coefficients from Germany, we find that the health index is surprisingly stable between countries. Greece and Switzerland have the best health on average, with a low health index, and Spain and France exhibit the worst health.

3.1.2 Depression

We use two measures of depression in this analysis. The first is the Euro-D depression scale. Unlike the measurement of health, the depression scale was specifically developed for detecting depression and has been validated for use in cross-country analysis (see Larraga et al. (2006) and Prince (2002)). The Euro-D scale runs from 0 to 12, counting whether the individual reported having problems with the following feelings in the prior month: sadness or depression; pessimism; suicidal thoughts; guilt; sleep trouble; lack of interest; concentration; appetite; irritability; fatigue; enjoyment; and tearfulness. The standard cut-off is 4 or more positive symptoms is depression. We also use a simple indicator variable for whether or not the individual reports being sad or depressed in the previous month.

3.1.3 Cognitive Ability

The cognitive reserve is defined in the neuro-psychological literature as the individuals' capacity to use brain networks more efficiently or, in other words, to process tasks in a more efficient manner (Stern (2002)). The decline in cognitive function with age is associated with structural changes in the brain (Raz (2004)). In addition, this cognitive decline is associated with diseases such as Alzheimer's.

Adam et al. (2006), using SHARE data, found that occupational activities, including paid-work and not paid-work as well as sport practice and other physical activities are highly correlated with cognitive ability. However, they cannot address whether this is due to a causal link between paid work and cognitive ability.

[13] In section 4.7, we explore the use of vignettes to control for country-specific reporting differences in depression.

In SHARE cognitive ability is measured through several questions. Memory is tested through word recall. The respondent was asked to memorize a list of 10 common words. They were then asked to recall this words two times, once immediately, and once again after a considerable delay spent answering more survey questions. The memory scale ranges from 0 to 20, with 20 being the highest. Verbal fluency is assessed by asking respondents to name as many animals as possible within a one-minute time frame. This variable ranges between 0 and 72.

Following Adam et al. (2006), we estimate the effect of retirement on memory and verbal fluency scores, but correct for the endogeneity in retirement using our early and full retirement ages as instrumental variables. For this analysis, we limit the sample further by dropping those individuals who had suffered a vascular accident, have Parkinson disease, had brain cancer, are taking medicine for depression and those who have ever been hospitalized in a psychiatric institution. Thus our sample size drops to 4,928.

3.2 Retirement Definition

There are two common ways of defining retirement: self-reported retirement status, or anyone who is not in the paid labor force. Often individuals report that they are retired even when working full- or part-time, simply because they have left their "career" job. Since we want to determine the effect of work status on health, we employ the latter definition. We consider this a cleaner measure of retirement behavior than the self-reported measure. Thus, while we limit our sample to those who are working in the paid labor market at the age of 50, we consider individuals that report themselves to be retired, a homemaker, sick and disabled, separated from the labor force (not temporary), and unemployed (not temporary) as retired.

3.3 Instruments

We use early and full statutory retirement ages in each country's social security scheme as instrumental variables. There is an 8-year difference in the ability to collect early retirement benefits in our sample, from age 57 in Italy to age 65 in Denmark (where there is no early retirement). These are highly significant in predicting retirement behavior, but since these are national eligibility ages, there is no reason to think they are linked to any particular individual's health. Further, it is difficult to image that there exists breaks in the health trajectory which begins at different ages in different countries. For example, there is no reason to believe that there is an independent, age-specific, non-linear change in health at age 57 for Italian men, age 60 for Belgian men, and at age 65 for Danish men.

3.4 Descriptive Statistics

Table 4 provides the descriptive statistics for the full sample. The average age of the sample is 59 years, with 24 percent being over the full retirement age and an additional 46 percent over the early retirement age of their country. 81 percent are still married or co-reside.

While only 25 percent report being in fair, bad or very bad health, we do see a number of health limitations. 28 percent report that health has limited their activities in the previous 6 months. The more severe measures of health decline, such as the number of ADLs, IADLs, or mobility limitations, are still, on average, quite low. While the average Euro-D depression measure is also low (only 1.6 out of 12), well under the 4-point cut-off for depression, 26 percent of our sample reported feeling depressed in the previous month. We also find that half of our sample is overweight, while another 17 percent is obese (as measured by BMI greater than 24).

On average, our sample can recall just under 22 animals in one minute, and can recall just under 10 words when asked to do so both immediately and again after some time has elapsed.

45 percent of our sample is out of the labor market. We also control for sector of employment in our regressions, where 18 percent are self-employed and another 12 percent work for the public sector.

One can also see the country-by-country breakdown of our sample. The most respondents come from Belgium, Germany, Sweden, the Netherlands, and France comprising most of the sample. Switzerland is the country with the fewest respondents.

4 Results

4.1 Instrument Validity

In order for statutory retirement ages to be valid instruments, they must be related to actual retirement behavior. Said differently, the propensity to retire must change at these given ages. Ideally we would illustrate this point with country-specific retirement hazard rates. Since we do not have panel data, we instead report the fraction of individuals retired before they are eligible for social security, when they are between the early and normal retirement ages, and after the normal retirement age in table 5. There is a dramatic increase in the fraction of the sample retired at both the early cut-off and the normal retirement age cutoff in every country. For example, only 20 percent of the sample in Austria is out of the labor market before they are eligible for the state pension. This fraction increases to 86 percent for those who are between the early retirement eligibility age and the normal retirement age. Virtually all of those over the full retirement age, 97 percent, are out of labor market.

The first stage regression indicates that early and full retirement ages are important predictors for retirement behavior. Table 6 presents the coefficients of estimating equation (2) for Europe. This indicates that being over the early retirement age increases the probability of being out of the labor force by 0.08 points, while being over the full-retirement age increases the probability by almost 0.15 points. Once controlling for these country-specific age-breaks, age itself is not statistically important for retirement behavior. The coefficients on the rest of the variables are as one would expect: self-employed and public-employed individuals are less likely to retire, and the more education one has the less likely they are to retire. The country control variables are also quite significant.

4.2 Self reported health, depression and cognitive ability

Table 7 presents the baseline results for Europe. The table presents the results for self-reported health (columns 1 and 2), the health index (columns 3 and 4), and the two measures of depression (columns 5-8). The first column of each pair presents the OLS results, which ignores the endogeneity of retirement and health, while the second column presents our preferred IV estimates.

Most of the effects of the control variables are statistically significant and as we would intuitively expect. Higher education tends to be associated with better health, and marriage is correlated with lower incidence of depression. Many of the country indicator variables are significant. A few surprises do occur. The more children one has the worse health they are in and the more likely they are to report depression. We interpret this result as a socioeconomic status or wealth indication, not that children themselves are bad for ones health. Age has a surprisingly little effect on any of our health measures, and we find no evidence of a non-linear relationship once we control for retirement behavior. We also added controls for the season of interview for the depression regressions, and find that these are insignificant.

The OLS results indicate that retirement is correlated with a 0.14 points increase in the likelihood of reporting in fair, bad or very bad health. Once we take into account the endogeneity of health and retirement, we find that retirement leads to a 0.35 point decrease in the probability of reporting in fair, bad or very bad health.

We find similar results when we use the health index instead of the self-reported health measure. Remember, the health index controls for both objective measures of health and country reporting differences at the mean. Here we find that there is a positive coefficient in the OLS, indicating a negative correlation between retirement and health, similar to the self-reported health findings. Again, once taking into account the endogeneity of the retirement decision, the sign reverse and we find that retirement leads to almost a 1 standard deviation decrease in the health index (indicating better health).

For both measures of depression, we find that there is a statistically significant and positive correlation between depression and retirement. However, we find no evidence of a causal mechanism since our IV coefficients are statistically insignificant.

As mentioned above, the sample for testing the effect of retirement on cognitive function is slightly smaller due to limited the sample to those without physical reasons to expect cognitive function decline. In addition to the control variables we included above, we also add a dummy for the individual being born out of the country, number of ADL limitations, number of IADL limitations, number of chronic diseases, number of non-professional activities done, a dummy for doing physical demanding activities (sports, heavy work) and less demanding activities (e.g. gardening) at least once per week in order to make our results directly comparable to Adam et al. (2006).

Much like Adam et al. (2006), we also find a strong negative correlation between retirement and cognitive function in both the memory and the verbal measures, as can be seen in table 8. Once we control for the endogeneity of the retirement decision, the coefficients become insignificant, and thus suggest there is no causal relationship between work status and decline of cognitive skills.

4.3 Age-Gradient of the Retirement Effect

By examining the health-retirement relationship in a multi-country setting, we can see whether or not there are differential health effects of working longer/retiring earlier. When our identification strategy is used in a single-country context, one can find the relationship between retirement and health at a particular age, say 65, but it is more difficult to then extrapolate what the effect of increasing the retirement age would be. Is there a different relationship between work status and health if one retires at age 62 or age 65?

Since we have varying ages of retirement, especially early retirement, between countries, we use this variation to examine whether there is a different relationship between retirement and health at different ages. We allow for different retirement effects depending on the age when the individuals retired allowing for 3 categories (ages 57-59, ages 60-64, and ages 65-69) and restrict the sample to those who retired at age 57 or older.

The results are presented in Table 9. We present the OLS results and the IV estimation for self-reported health and the health index[14]. Columns 1 and 2 present the results for self-reported health, and columns 3 and 4 the health index results.

The differences between the OLS and the IV estimates again emphasize the importance of accounting for the endogeneity of health in the retirement decision. For both the self-reported health and the health index, the OLS estimates indicate that retiring early is correlated with bad health. We also find a relatively strong age-gradient in this relationship, so the earlier one retires, the worse health he has.

Once we control for endogeneity, we find that there is a positive and significant effect on health of leaving the work force between ages 65-69, and only when using the health index measure. The estimate is actually larger than the baseline estimate, suggesting that retirement at age 65 leads to a 1.7 point decrease in the health index, or an overall improvement in general health. These findings suggest that governments should proceed with caution

[14] Results for depression and cognitive function are available from the authors, but much like the results presented in tables 7 and 8, we find no causal relationship between work status and depression or cognitive function.

when increasing the normal retirement ages in order to decrease the financial burden on their social security systems - this might in fact just shift the burden between social security and health care costs.

4.4 Robustness Checks

4.4.1 Conditions during the Lifecycle

One concern is that there might be other things that could cause a nonlinear relationship between health and age other than retirement at these particular ages. Lindeboom et al. (2003) find that macroeconomic events during early life and childhood have an effect on lifetime health, and thus these nonlinearities we find could be due to differences in childhood resources. Given the age of the individuals, the biggest concern is that World War II (or the Spanish Civil War), or famines related to these wars, could be driving the results, leading to one-time differences between these particular cohorts, and not a general effect of retirement on health.

We have added indicator variables if someone was born between the date of war declaration and peace treaty signing in their particular country. The results are presented in table 10[15]. We present the baseline IV results from table 7 again for easier comparison, as well as the regression with these additional controls. The estimated effect of retirement on self-reported health and the health index are statistically insignificant from the baseline estimation. It is also interesting to note that these indicator variables for being born during a war or a famine are insignificant themselves.

4.4.2 Vignettes

It is generally found that response scales tend to differ by country as a result of language and cultural differences. Hence, we can consider self-reported health and depression measures to be an indicator of relative health rather than absolute health. While absolute health would refer to one's health status as compared to the health of any other person in the world, relative health would be one's health relative to someone else in the same country and possibly other factors. In order to make broader comparisons of self-reported measures of health between individuals across countries, two strategies have

[15] We have also used date of invasion instead of date of war declaration, and the results are robust to the definition of the timing of war.

been developed. Kapteyn et al. (2008) use vignettes, and Meijer et al. (2007) use an objective physical measure to anchor the scales.

SHARE contains vignette questions in which a subsample of respondents[16] are asked to rate a group of health problems (pain, sleep, breath and mobility problems, self reported memory, depression and work disability) of hypothetical persons. While the sample size drops considerably, only 915 observations satisfied our sample requirements, we explore the use of vignettes to adjust the self-reported depression information to see if it changes our earlier results[17].

For this sample we use the following question about depression:

"Overall in the last 30 days, how much of a problem did you have with feeling sad, low, or depressed?"

Vignettes are hypothetical questions that ask the respondent to rate the health of hypothetical people based on a brief description of the abilities and limitations. An example in the depression domain follows:

"Henri feels nervous and anxious. He worries and thinks negatively about the future, but feels better in the company of people or when doing something that really interests him. When he is alone he tends to feel useless and empty. Overall in the last 30 days, how much of a problem did Henri have with feeling sad, low, or depressed?"

The responses to the vignettes are on the same 5-point scale (none, mild, moderate, severe, extreme) which is available when answering about themselves. In order to guarantee a sufficient number of observations in each category, responses were re-grouped to a 3- point scale (none, mild, moderate or severe or extreme)[18].

[16] SHARE respondents in Austria, Switzerland and Denmark were not asked the vignette questions, thus will be omitted from this analysis.

[17] We examine depression because there is no vignette measuring overall general health, comparable to the self-rated health information. This robustness check tests to see if our insignificant results are due to lack of cross-country comparability, or if there is truly no measurable effect of retirement on depression.

[18] The gender of the vignette person is randomly selected. While Kapteyn et al (2008) find it to be important when using the vignette information we cannot control for it in our estimates due to the fact that this information is not made publicly available in the second release of SHARE.

Self report depression is modeled as a function of respondent characteristics, including retirement, as and ordered probit model:

$$
\begin{aligned}
Y_i^* &= X_i\beta + \varepsilon_i \\
Y_i &= j \text{ if } \tau_i^{j-1} < Y_i^* \le \tau_i^j, j = 1, 2, 3 \\
& \quad \varepsilon_i \tilde{} N(0, \sigma^2) \\
\tau_i^0 &= -\infty; \tau_i^3 = \infty; \tau_i^1 = \gamma_1 V_i; \tau_i^2 = \tau_i^1 + \exp(\gamma_2 V_i)
\end{aligned}
$$

Each respondent answered 3 vignettes questions. The evaluations of the vignettes are modeled using similar ordered response equations:

$$
\begin{aligned}
Y_i^* &= \theta + \varepsilon_{vi} \\
Y_i &= j \text{ if } \tau_i^{j-1} < Y_i^* \le \tau_i^j, j = 0, 1, 2, 3 \\
& \quad \varepsilon_{vi} \tilde{} N(0, \sigma^2)
\end{aligned}
$$

ε_{vi} are assume to be independent of each other, independent of ε_i and of X_i. The set of explanatory variables in the thresholds (V_i) includes different country dummies. This model is identified under the assumption that respondents use the same scale for the self-reports and the vignettes evaluations which means that the thresholds τ_i^j are the same for the self-reports and the vignettes. Then, all parameters are estimated simultaneously by maximum likelihood. Since all the error terms are assumed independent, the likelihood contribution is a product of univariate normal probabilities over all vignette evaluations and the self-report.

In order to take into account the endogeneity of retirement decisions we include as explanatory variables predicted retirement probabilities obtained from a linear probability model of retirement on the instruments.

Table 11 shows the estimated coefficients both with and without using vignette information. For illustration, in the second part of this table we show the estimated coefficients of the country dummies included in the first thresholds. Similarly, we obtained estimated coefficients for each country dummy included in the second threshold. As we can see in these table there are significant differences of reporting patterns across countries. However, we still do not find a significant effect of retirement on depression even when we use the vignettes to control for these reporting differences across countries.

5 Conclusion

We find that there is a statistically significant and economically important effect of retirement on general health. We also illustrate the importance of looking for the causal effect instead of just raw correlations between retirement and health status, since we find no evidence of a causal link between work status and depression or cognitive function.

Our estimates indicate that retiring leads to a 0.35 points decrease in the probability of reporting in fair, bad or very bad health, and an almost 1 standard deviation drop in the health index. It is, however, difficult to quantify these results. Self-reported health has been found to be an important predictor for mortality, especially for death due to diabetes, infection or respiratory diseases. Benjamins et al. (2004) finds that individuals who rate themselves in fair (poor) health are four- (six-) times more likely to die from diabetes than those that report they are in excellent health. So while it is difficult to quantify our results for self-reported health and the health index in terms of quality of life improvements or health care expenditures, the health-preserving effect does suggest that there might be a fair amount of health expenditure savings due to retirement as well.

In addition, we find that the health effects are driven by "late-retires", those who retire at the age of 65 and above, and no health effect for those retiring beforehand. This could indicate a strong age-gradient in the relationship between retirement and health. This could give pause to policy makers wanting to increase the full retirement age to increase the solvency of the social security/pension system. While postponing retirement could be good for some, our estimates suggest increasing retirement ages could have a strong negative effect on the average health of males in the country. Thus increasing the retirement age may only shift the burden of payments between the social security and the health care system.

Next we would like to explore the mechanism in which causes the health increase once one is retired. It is important to determine if individuals are using the extra time to make further health investments, and which health investments are being made. This could help direct public policy measures to decrease the effect of working on health, and potentially increasing the retirement age without experiencing the declines in health or increases in health spending that are suggested by the results of this paper.

References

[1] Adam S., Bay C., Bonsang E., Germain S. and S. Perelman (2006) "Occupational Activities and Cognitive Reserve: A Frontier Approach Applied to the Survey on Health, Ageing, and Retirement in Europe (SHARE)", CREPP Working Paper 2006/05.

[2] Adams O. and L. Lefebvre (1981) "Retirement and Mortality." Aging and Work; 4:115-120.

[3] Angrist J. D. and V. Lavy (1999) "Using Maimonides Rule to Estimate the Effect of Class Size on Scholastic Achievement", *Quarterly Journal of Economics*, 114, 533-575.

[4] Atchley R.C. (1976) "The Sociology of Retirement." New York: Halsted Press.

[5] Benjamins M. R., Hummer R.A., Eberstien I.W. and C. B. Nam. (2004) "Self-reported Health and Adult Mortality Risk: An Analysis of Cause-Specific Mortality." *Social Science and Medicine*, 1297-1306.

[6] Bradford L. P. (1979) "Can You Survive Your Retirement?" *Harvard Business Review* 57(4), 103-109.

[7] Börsch-Supan A., Brugiavini A., Jürges H., Mackenbach J., Siegrist J. and G. Weber (eds.) (2005): Health, Ageing and Retirement in Europe – First Results from the Survey of Health, Ageing and Retirement in Europe, Mannheim Research Institute for the Economics of Aging, University of Mannheim.

[8] Börsch-Supan, A and H. Jürges (eds.) (2005): The Survey of Health, Ageing and Retirement in Europe - Methodology, Mannheim Research Institute for the Economics of Aging, University of Mannheim.

[9] Bound J. and T. Waidmann. (2007) "Estimating the Health Effects of Retirement", University of Michigan Retirement Research Center working paper 2007-168.

[10] Bound J., Schoenbaum M., Stinebrickner T. and T. Waidmann. (1999). "The Dynamic Effects of Health on the Labor Force Transitions of Older Workers," *Labour Economics* 6(2), 179-202.

[11] Cain G. (1975), "Regression and Selection Models to Improve Nonexperimental Comparisons", in: C. Bennett and A. Lumsdaine, eds., Evaluation and Experiment, (Academic Press, New York), 297-317.

[12] Campbell D. (1969) "Reforms and Experiments" *American Psychologist*, 24, 409-429.

[13] Carp F.M. (1967) "Retirement Crisis" *Science* 157, 102-3.

[14] Casscells W., Evans D., DeSilva R., Davies J.E., Hennekens C.H., Rosner B., Lown B. and M.J. Jesse. (1980) "Retirement and Coronary Mortality." *Lancet*; I(8181), 1288-1289.

[15] Charles K. K. (2002) "Is Retirement Depressing? Labor Force Inactivity and Psychological Well-being in Later Life." NBER working paper 9033.

[16] Coe N. and M. Lindeboom (2008) "Does Retirement Kill You? The Evidence from Early Retirement Windows." manuscript.

[17] Ekerdt D., Raymond Bosse J. and J.S. LoCastro. (1983) "Claims that Retirement Improves Health." *Journal of Gerontology* 38.

[18] Eisdorfer C. and F. Wilkie. (1977) "Stress, Disease, Aging and Behavior." In James E. Birren and K. Warnere Shaie (eds) Handbook of he Phsychology of Aging. New York: Van Nostrand, Reinhold. Pg.251-75.

[19] Gonzales E.R. (1980) "Retiring May Predispose to Fatal Heart Attack" *Journal of the American Medical Association* 243, 13-14.

[20] Goldberger A. S. (1972) "Selection Bias in Evaluating Treatment Effects: Some Formal Illustrations" Discussion Paper 129-72, Madison WI: University of Wisconsin, Institute for Research on Poverty, June.

[21] Hahn J., Todd P. and W. Van der Klaauw (2001), "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design", *Econometrica*, 69, 201-209.

[22] Haynes S.G., McMichael A.J. and H.A. Tyroler (1978) "Survival After Early and Normal Retirement" *Journal of Gerontology* 33, 269-278.

[23] Kapteyn A., Smith J.P. and A. van Soest (2008) "Vignettes and Self-Reported Work Disability in the Netherlands and the US" *American Economic Review* 97(1), 461-473.

[24] Kasl S.V. (1980) "The Impact of Retirement." in C.L. Cooper and R. Payne (eds) Current Concerns in Occupational Stress. New York: John Wiley.

[25] Kerkhofs M. and M. Lindeboom (1997) "Age Related Health Dynamics and Changes in Labor Market Status" *Health Economics*, 6, 407-424.

[26] Larraga, L., Saz P., Dewey M.E., Marcos G., Lobo A. and the ZARADEMP Workgroup (2006) "Validation of the Spanis version of the Euro-D Scale: an instrument for detecting depression in older people," *International Journal of Geriatirc Psychiatry* 21, 1199-1205.

[27] Lindeboom M., Portrait F. and G. J. van den Berg (2002) "An Econometric Analysis of the Mental-Health Effects of Major Events in the Life of Older Individuals" *Health Economics* 11(6), 505-520.

[28] Lindeboom M., Portrait F. and G. J. van den Berg (2003) "Individual Mortality and Macroeconomic Conditions From Birth to Death" IZA DP no 930.

[29] MacBride A. (1976) "Retirement as a Life Crisis: Myth or Reality?" *Canadian Psychiatric Association Journal* 72, 547-56.

[30] Meijer E., Kapteyn A. and T. Andreyeva (2007) "Health Indexes and Retirement Modelling in International Comparisons" Manuscript.

[31] Minkler M. (1981) "Research on he Health Effects of Retirement: An Uncertain Legacy" *Journal of Health and Social Behavior* 22 (2), 117-130.

[32] Natali D. (2004) The Pension System Observatoire Social Européen. Research Project. "LA METHODE OUVERTE DE COORDINATION (MOC) EN MATIERE DES PENSIONS ET DE L'INTEGRATION EUROPEENNE" Service Public Fédéral Sécurité Sociale.

[33] Neuman K. (2008) "Quit Your Job and Live Longer? The Effect of Retirement on Health" *Journal of Labor Research* 29 (2).

[34] Niemi T. (1980) "Retirement and Mortality" *Scandinavian Journal of Social Medicine* 8, 39-41.

[35] OECD (2003) "Economic Survey of Austria", chapter 3: Pensions. http://www.oecd.org/dataoecd/5/16/27424371.pdf

[36] OECD (2005). "Vieillissement et politiques de l'emploi". France. Ageing and Employment Policies. http://www.oecd.org/dataoecd/52/58/34591763.pdf

[37] Price M. (2002) "The Development of the Euro-D Scale" chapter in Principles and Practice of Geriatric Psychiatry (second edition), editors John R.M. Copeland, Mohammed T. Abou-Saleh and Dan G. Blazer, Wiley and Sons Ltd.

[38] Preesman L. (2006), "Dutch to Abolish Civil Service Retirement Age" IPE International Publishers Ltd. Netherlands, June 28.

[39] Raz N. (2004) "The Aging Brain: Structural Changes and their Implications for Cognitive Aging", In R.A. Dixon, L. Backman & L.G. Nilsson (Eds.), New Frontier in Cognitive Aging. Oxford: University Press, 115-133.

[40] Rowland K.F. (1977) "Environmental Events Predicting Death for the Elderly" *Psychological Bulletin* 84, 349-372.

[41] Sheppard H.L. (1976) "Work and Retirement." In Robert H. Binstock and Ethel Shanas, Handbook of Aging and the Social Sciences. New York: Van Nostrand, Reinhold. Pg 286-306.

[42] Stern Y. (2002) "What is Cognitive Reserve? Theory and Research Application of the Reserve Concept", Journal of the International Neuropsychology Society, 8, 448-460.

[43] Sundén A. (2004). "The Future of Retirement in Sweden" PRC WP 2004-16. Pension Research Council Working Paper. Pension Research Council.

[44] Tarricone R. (1997) "Metodologia di indagine e primi risultati ul costo sociale della depressione maggiore in Italia" *Mescosan* 6(22), 57-68; as summarized in Berto, Patrizia, Daniele D'Ilario, Pierfrancesco Ruffo,

Roberto Di Virgilio and Fortunato Rizzo (2000) "Depression: Cost-of-Illness Studies in the International Literature, a Review" *The Journal of Mental Health Policy and Economics* 3, 3-10.

[45] The Bartelsmann foundation, International Reform Monitor, Country info. http://www.reformmonitor.org/index.php3?mode=status

[46] Thompson W.E. and G.F. Streib (1958) “Situational Determinants: Health and Economic Deprivation in Retirement” *Journal of Social Issues* 14(2),18-24.

[47] Van der Klaauw W. (2003) "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach", *International Economic Review* 43, 1249-1287.

Table 1: Eligibility for Public Retirement Benefits

Country	Early (normal) retirement age for all workers	
	Males	Females
Austria	60 (65)	57 (60)
Belgium	60 (65)	60 (65)
Denmark	65 (65)	65 (65)
France	57 (60)	57 (60)
Germany	63 (65)	63 (65)
Greece	57 (65)	57 (65)
Italy	57 (65)	57 (65)
Netherlands	60 (65)	60 (65)
Spain	60 (65)	60 (65)
Sweden	61 (65)	61 (65)
Switzerland	63 (65)	62 (64)

Table 2: Ordered Probit Results for Health Index

Self-reported Health	AT	DE	SE	NL	ES	IT	FR	DK	GR	CH	BE
Limited due to Health	0.734***	1.107***	0.443***	0.675***	0.678***	0.905***	0.839***	0.538***	1.017***	0.861***	0.848***
	(0.149)	(0.125)	(0.115)	(0.132)	(0.169)	(0.158)	(0.135)	(0.163)	(0.171)	(0.248)	(0.124)
Chronic Diseases	0.230***	0.217***	0.255***	0.286***	0.234***	0.146***	0.281***	0.262***	0.340***	0.208**	0.296***
	(0.064)	(0.050)	(0.048)	(0.052)	(0.055)	(0.055)	(0.049)	(0.058)	(0.050)	(0.101)	(0.044)
Mobility Limitations	0.211***	0.198***	0.159**	0.173***	0.068	0.137**	0.173***	0.402***	0.118**	0.452**	0.111**
	(0.065)	(0.043)	(0.069)	(0.056)	(0.054)	(0.063)	(0.048)	(0.084)	(0.052)	(0.200)	(0.047)
ADL’s	-0.049	0.063	0.306	0.231	-0.095	0.242	-0.211	0.082	-0.057	-0.793	-0.023
	(0.188)	(0.154)	(0.211)	(0.224)	(0.205)	(0.216)	(0.157)	(0.217)	(0.170)	(0.513)	(0.143)
IADL’s	-0.118	0.090	0.270	0.053	-0.008	-0.022	0.412***	-0.828***	0.094	0.385	0.102
	(0.187)	(0.146)	(0.213)	(0.186)	(0.129)	(0.289)	(0.157)	(0.292)	(0.224)	(0.321)	(0.178)
grip strength	-0.016**	-0.009*	-0.020***	-0.008	-0.011*	-0.017***	-0.006	-0.001	-0.004	0.014	-0.015***
	(0.007)	(0.005)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.007)	(0.005)	(0.012)	(0.005)
Overweight	0.039	0.144	0.196*	-0.083	0.145	0.071	0.057	-0.002	-0.078	-0.360*	0.081
	(0.140)	(0.109)	(0.107)	(0.115)	(0.143)	(0.129)	(0.109)	(0.138)	(0.123)	(0.187)	(0.103)
Obese	0.208	0.203	0.272*	-0.066	0.471***	0.091	0.391**	0.019	0.134	0.360	0.229*
	(0.182)	(0.146)	(0.152)	(0.173)	(0.171)	(0.172)	(0.152)	(0.195)	(0.160)	(0.276)	(0.128)
euro-D	0.089*	0.169***	0.113***	0.096***	0.042	0.039	0.094***	0.101**	0.093***	0.018	0.075***
	(0.046)	(0.032)	(0.035)	(0.032)	(0.036)	(0.033)	(0.028)	(0.043)	(0.033)	(0.061)	(0.026)
Number of symptoms	0.299***	0.071	0.214***	0.221***	0.227***	0.161**	0.117**	0.234***	0.158***	0.204*	0.171***
	(0.076)	(0.052)	(0.055)	(0.069)	(0.065)	(0.065)	(0.052)	(0.071)	(0.059)	(0.117)	(0.051)
Physical inactivity	0.637*			0.063	0.376	0.112	0.814***				
	(0.347)			(0.337)	(0.242)	(0.211)	(0.223)				
Hospital stay	-0.046	0.039	-0.113	0.013	0.418*	0.166	0.274*	-0.093	0.053	0.255	0.358**
	(0.176)	(0.139)	(0.175)	(0.189)	(0.226)	(0.200)	(0.161)	(0.232)	(0.190)	(0.310)	(0.152)
Obs	390	603	600	516	386	437	559	351	533	216	691

Standard errors in parentheses * significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Health Index Averages

Country	Average Health Index Own-country coefficients	Average Health Index German Coefficients
Austria	.34	.79
Germany	.83	.82
Sweden	.04	.64
Netherlands	.52	.73
Spain	.65	.94
Italy	.04	.79
France	.93	.84
Denmark	1.06	.75
Greece	.60	.54
Switzerland	1.20	.45
Belgium	.36	.74

Table 4: Descriptive Statistics

Demographics	Average
age (50-69)	59
Over Full Retirement Age	0.24
Over Early Retirement Age	0.46
High Education	0.27
Medium Education	0.36
Married/Co-resident	0.81
Health Measures	
Fair/Bad/Very Bad Health	0.25
Felt Depressed in Last Month	0.26
Euro-d Depression Scale (0-12)	1.57
Maximum Grip Strength (0-100)	46.91
Limited due to Health	0.28
Number of Chronic Conditions (0-8)	1.16
Number of Symptoms in last 6 months (0-9)	0.97
Number of Mobility Limitations (0-10)	0.62
Number of ADL limitations (0-6)	0.06
Number of IADL limitations (0-7)	0.07
Hospital Stay in previous year	0.10
Overweight	0.51
Obese	0.17
Memory Measures	
Verbal Score (0-72)	21.80
Memory Recall (0-20)	9.45
Employment	
Retired	0.45
Self-employed	0.18
Public-sector Employment	0.12
Country Representation	
Austria	0.07
Germany	0.11
Sweden	0.11
the Netherlands	0.10
Spain	0.07
Italy	0.08
France	0.11
Denmark	0.07
Greece	0.10
Switzerland	0.04
Belgium	0.13

Table 5: Retirement Rates (%)

	Under Pension Age	Over Early Retirement Age	Over Full Retirement Age
Austria	20	86	97
Germany	28	69	91
Sweden	11	34	86
The Netherlands	10	69	97
Spain	13	60	93
Italy	11	65	89
France	11	49	95
Denmark	20	--	90
Greece	4	40	89
Switzerland	8	50	73
Belgium	26	77	98

Table 6: First Stage Results

Out of the Labor Force	OLS
Over Early Retirement Age	0.0814***
	(0.0231)
Over Full Retirement Age	0.1462***
	(0.0180)
Age	0.0399
	(0.0256)
Age squared	0.0000
	(0.0002)
Public-employment	-0.0615***
	(0.0151)
Self-employed	-0.2017***
	(0.0130)
married	-0.0090
	(0.0129)
High education	-0.0976***
	(0.0129)
Medium education	-0.0213*
	(0.0122)
number of children	-0.0027
	(0.0037)
Household income	-0.3909
	(0.2517)
Constant	-1.8946***
	(0.7338)
Observations	5282
R-squared	0.52

Standard errors in parentheses * significant at 10%; ** significant at 5%; *** significant at 1%
Note: This regression also includes country fixed effects.

Table 7: Results for Health and Depression

	Bad Health		Health Index		Euro-D Depression Scale		Felt Depressed in a Month	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
retired	0.1393***	-0.3545**	0.3473***	-0.9903***	0.2702***	-0.0691	0.0308*	-0.1183
	(0.0165)	(0.1505)	(0.0394)	(0.3667)	(0.0679)	(0.5740)	(0.0172)	(0.1465)
age	0.0100	0.0049	0.1662***	0.1524**	-0.0035	-0.0070	0.0001	-0.0015
	(0.0237)	(0.0257)	(0.0565)	(0.0625)	(0.0973)	(0.0977)	(0.0247)	(0.0249)
Age Squared	-0.0001	0.0002	-0.0013***	-0.0006	-0.0001	0.0001	-0.0000	0.0000
	(0.0002)	(0.0002)	(0.0005)	(0.0006)	(0.0008)	(0.0009)	(0.0002)	(0.0002)
Public employment	0.0030	-0.0271	0.0067	-0.0748	0.1055	0.0848	0.0203	0.0112
	(0.0182)	(0.0217)	(0.0435)	(0.0529)	(0.0749)	(0.0827)	(0.0190)	(0.0211)
Self employed	-0.0300*	-0.1304***	-0.0684*	-0.3404***	0.0004	-0.0685	0.0130	-0.0172
	(0.0160)	(0.0350)	(0.0383)	(0.0853)	(0.0660)	(0.1333)	(0.0168)	(0.0340)
married	-0.0130	-0.0171	-0.0727**	-0.0839**	-0.3991***	-0.4019***	-0.0847***	-0.0860***
	(0.0155)	(0.0168)	(0.0371)	(0.0410)	(0.0638)	(0.0642)	(0.0162)	(0.0164)
High Education	-0.1197***	-0.1676***	-0.2842***	-0.4139***	-0.2651***	-0.2981***	0.0057	-0.0088
	(0.0156)	(0.0222)	(0.0372)	(0.0542)	(0.0640)	(0.0848)	(0.0163)	(0.0217)
Medium Education	-0.0674***	-0.0775***	-0.1654***	-0.1927***	-0.1621***	-0.1691***	-0.0113	-0.0144
	(0.0147)	(0.0162)	(0.0351)	(0.0395)	(0.0604)	(0.0617)	(0.0154)	(0.0158)
number of children	0.0161***	0.0152***	0.0526***	0.0502***	0.1083***	0.1076***	0.0215***	0.0213***
	(0.0045)	(0.0049)	(0.0107)	(0.0118)	(0.0184)	(0.0185)	(0.0047)	(0.0047)
Household Income	-0.2508	-0.4442	0.1121	-0.4117	-1.8743	-2.0076	-0.2305	-0.2890
	(0.3036)	(0.3336)	(0.7248)	(0.8128)	(1.2477)	(1.2705)	(0.3170)	(0.3244)
spring					0.1638	0.1660	0.0408	0.0418
					(0.1138)	(0.1141)	(0.0289)	(0.0291)
summer					0.0599	0.0572	0.0034	0.0023
					(0.1048)	(0.1051)	(0.0266)	(0.0268)
autumn					-0.0822	-0.0853	-0.0115	-0.0128
					(0.1093)	(0.1097)	(0.0278)	(0.0280)
Constant	-0.0829	-0.4667*	-4.4501***	-5.4897***	1.8868	1.6279	0.3464	0.2327
	(0.6992)	(0.7652)	(1.6692)	(1.8645)	(2.8752)	(2.9147)	(0.7306)	(0.7441)
Obs	5282	5282	5282	5282	5282	5282	5282	5282
R-squared	0.07		0.08		0.04	0.04	0.02	0.01

Standard errors in parentheses * significant at 10%; ** significant at 5%; *** significant at 1%
Note: All regressions also includes country dummy variables. These are linear probability models, although the results do not change substantially using probit models for the regression equation or the bad health, or felt-depressed regressions.

Table 8: Results: Cognitive Function

	Memory		Verbal	
	OLS	IV	OLS	IV
retired	-0.2769**	-0.0390	-0.7544***	2.5647
	(0.1153)	(0.9798)	(0.2628)	(2.2682)
Age	0.0132	0.0203	-0.0973	0.0019
	(0.1631)	(0.1658)	(0.3718)	(0.3837)
Age squared	-0.0007	-0.0009	0.0002	-0.0023
	(0.0014)	(0.0015)	(0.0031)	(0.0036)
Public employment	-0.3441***	-0.3298**	-0.4915*	-0.2916
	(0.1259)	(0.1389)	(0.2869)	(0.3216)
Self-employed	0.0241	0.0705	-0.1769	0.4699
	(0.1100)	(0.2192)	(0.2506)	(0.5074)
married	0.1545	0.1566	0.5856**	0.6142**
	(0.1032)	(0.1036)	(0.2352)	(0.2398)
High education	1.7646***	1.7874***	3.5650***	3.8827***
	(0.1095)	(0.1438)	(0.2496)	(0.3329)
Medium education	1.0012***	1.0046***	1.5804***	1.6268***
	(0.1014)	(0.1024)	(0.2312)	(0.2370)
Household income	2.8554	2.9376	4.0035	5.1506
	(2.0524)	(2.0806)	(4.6775)	(4.8163)
Foreign-born	-0.6778***	-0.6843***	-2.5971***	-2.6877***
	(0.1540)	(0.1563)	(0.3509)	(0.3618)
number of chronic disease	0.0159	0.0126	0.0948	0.0483
	(0.0372)	(0.0396)	(0.0848)	(0.0917)
number of adl limitations	-0.1442	-0.1465	-0.0400	-0.0714
	(0.1445)	(0.1448)	(0.3293)	(0.3353)
number of iadl limitations	-0.4177***	-0.4292***	-0.3921	-0.5517
	(0.1483)	(0.1555)	(0.3379)	(0.3600)
Non-professional activities	0.1779***	0.1759***	0.5911***	0.5627***
	(0.0364)	(0.0374)	(0.0830)	(0.0865)
Physically-demanding Activities	0.1825**	0.1968*	0.7988***	0.9977***
	(0.0882)	(0.1058)	(0.2011)	(0.2449)
Moderate Activities	0.0060	-0.0005	1.1110***	1.0204***
	(0.1399)	(0.1425)	(0.3189)	(0.3298)
Observations	4928	4928	4928	4928
R-squared	0.19	0.19	0.25	0.23

Standard errors in parentheses * significant at 10%; ** significant at 5%; *** significant at 1%
Note: All regressions also include country dummy variables.

Table 9: Age Gradient of the Retirement Effect

	Bad Health		Health Index	
	OLS	IV	OLS	IV
Retired 57-59	0.3016*	-0.4893	0.1979*	0.0053
	(0.0807)	(1.0988)	(0.0562)	(0.7390)
Retired 60-64	0.1762*	-0.8770	0.1588*	-0.8424
	(0.0799)	(0.8002)	(0.0546)	(0.5543)
Retired 65-69	0.0501	-1.5531	0.0209	-1.7640*
	(0.1334)	(1.0225)	(0.0922)	(0.7379)
Obs	4446	4446	4446	4446

Standard errors in parentheses * significant at 10%; ** significant at 5%; *** significant at 1%
Note: All regressions also include age, age squared, sector employment, marital status, educational attainment, number of children, household income, and country dummy variables.

Table 10: Macroeconomic Controls

	Baseline Bad Health	War Controls	Baseline Health Index	War Controls
retired	-0.3545**	-0.3536**	-0.9903***	-1.0385***
	(0.1505)	(0.1636)	(0.3667)	(0.4012)
World War II		0.0253		0.1025
		(0.0270)		(0.0662)
famine		-0.0145		-0.1089
		(0.0539)		(0.1321)
Observations	5282	5282	5282	5282

Table 11: Vignettes

	Without vignettes	With vignettes
Age	0.01152	-0.05456
	(0.1691)	(0.03619)
Age squared	0.0000899	0.0005885
	(0.001425)	(0.0007974)
Probability of retired	-0.7803	-0.6896
	(0.8689)	(0.9552)
Public employment	-0.1429	-0.1503
	(0.1419)	(0.1303)
Self-employed	-0.1610	-0.1477
	(0.2133)	(0.2239)
Married	-0.2424***	-0.2550
	(0.1038)	(0.1822)
High education	-0.3097***	-0.3162***
	(0.1369)	(0.1231)
Medium education	-0.1313	-0.1435
	(0.1004)	(0.09636)
Number of children	0.03012	0.03206
	(0.03211)	(0.0738)
household income	-1.8069	-1.8932
	(3.05765)	(2.5758)
Sweden	0.3491	0.2377
	(0.2367)	(0.3566)
Netherlands	-0.2919	-0.6422***
	(0.1942)	(0.2227)
Spain	-0.4693**	-0.7006***
	(0.1942)	(0.2622)
Italy	-0.1256	-0.01683
	(0.2038)	(0.04929)
France	-0.1629	-0.3367**
	(0.1772)	(0.1636)
Greece	0.2945	-0.04520
	(0.2058)	(0.3340)
Belgium	0.0580	-0.2260
	(0.2155)	(0.1611)
Spring	-0.04421	-0.04814
	(0.1677)	(0.07585)
Summer	0.2214	0.2325
	(0.1739)	(0.1384)
Autumn	0.0548	0.05004
	(0.1548)	(0.10962)
Observations	915	915

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 11: Vignettes (continuation)

threshold1	0.6442	
	(5.2534)	
Germany		-1.5513***
		(0.09573)
Sweden		-1.6722***
		(0.1287)
Netherlands		-1.9049***
		(0.1311)
Spain		-1.7309***
		(0.1188)
Italy		-1.4086***
		(0.0792)
France		-1.6737***
		(0.0988)
Greece		-1.8538***
		(0.1736)
Belgium		-1.8290***
		(0.07586)

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%