

Effects of the onset of an adverse health condition on the retirement decision of European workers

Rosalia Vazquez-Alvarez^{a*}

SIAW, University of St.Gallen, Switzerland

(Preliminary version, first draft)

March 2, 2008

Abstract

This paper estimates the effect of experiencing the onset of an adverse health condition on the retirement decision of European workers. Conditional on institutional characteristics (country and individual dependent) and a selection of socioeconomic and demographic characteristics of the individual, we are able to net out the effect of health conditions that occur before the decision to retire, if retired. This paper compares the effect of adverse health conditions on the decision to retire from individuals aged 50 and over from 11 European countries using the 2004 SHARE survey. The results suggest that suffering an adverse health condition can significantly increase the age of retirement although this may not necessarily imply an increase in total contributory pensions: some specific countries (Greece and Switzerland in particular) may deviate from the overall result. The results are carefully estimated controlling for potential sources of endogeneity and selection bias.

Key words: Retirement decisions, health, Europe, SHARE survey

JEL-Classification: C14, H12

*Email: rosalia.vazquez-alvarez@unisg.ch. Address for correspondence: SIAW, University of St.Gallen, Bodanstrasse, 8, CH-9000. I am particularly grateful to Markus Froelich, Michael Lechner, Taghreed Adam, Raquel Carrasco and Pedro Mira for very helpful comments. I am also grateful for comments from participants at the 2007 COST-23 meeting (The Hague) and participants to the SIAW Internal Labour Seminars. All errors remain my own.

1 Introduction

Workers can choose to retire early, to retire at the official age of retirement or to keep on working after the official age of retirement. The decision to leave the labour market and the age at which this happens, depends on both country dependent institutional aspects and the characteristics of individuals. Institutional aspects are equally applicable to individuals legally working in the same country but may imply different conditions to different individuals because aspects such as gender, length of participation in the labour market or even as result of belonging to a different age cohort. Likewise, specific characteristics such as wealth and household composition may be determinant aspects on the retirement decisions of individuals. One such characteristic that is often thought as an important determinant on the decision to retire is health. Becoming the recipient of some adverse health condition may reduce or altogether eliminate the capacity to remain as participant in the labour market. Thus, early retirement may result from some adverse health condition with the possible consequent loss in accumulated wealth after retirement, including the loss of accumulated pension rights. It may also be the case that the onset of some health condition implies increasing the length of time (delaying retirement) to cover the possible medical expenses associated with such health condition. In the event that individuals do not suffer some adverse health shock, early retirement may be the decision of individuals who prefer to leave the labour market to enjoy longer periods of retirement time. In this latter case retiring is a simple decision to purchase leisure at the expense of lower pension rights. The key issue in this paper is trying to understand the effect of adverse health events on the retirement decision of workers in European countries once we isolate all other possible causal paths that may also determine retirement at an age distinct from the corresponding official age. We use the 2004 SHARE survey (Survey of Health, Ageing and Retirement in Europe) to elicit the effect of health on retirement. The survey is particularly appropriate for this purpose since it contains individuals at the later stages in the labour market and, at the same time, is highly information on the onset of key health conditions (either chronic illness, an acute condition or some mental disorders). At present, the SHARE survey is only available as a cross section (only wave 1 from 2004 is so far available). However, we are able to estimate the dynamic effects of health on retirement decisions because of the self-reported retrospective information in the survey. The results suggest that for countries such as Belgium, France, Greece and the Netherlands, and for sub-populations that are still of working age, the effect of an adverse health condition is to

reduce the probability of early retirement relative to those who belong to the same country but do not suffer any health condition. Once we focus the analysis to those who are already retired we estimate that the onset of an adverse health condition significantly decrease the probability of retiring earlier and furthermore it also increases the age at which individuals decide to retire. At the same time we find that this delayed exit from the labour market is not accompanied by a significant increase in total receipts from occupational pensions.

The paper is organized as follows. Section 2 reviews the literature background on retirement decisions in the labour market with an specific emphasis on exiting work that looks at the link between retirement and health. Section 3 explains the data and the treatment of the data such that our estimates are clean from possible reverse causality, other endogeneity problems and selection bias. Section 4 shows the results based on duration analysis on the variable age of retirement where the full population is taken into account. Section 5 shows the results of estimating the (treatment) effect of health on a selection of outcomes that reflect individual's decision to retire. Section 6 concludes. A set of appendix complement the results.

2 Health and Retirement decisions

It is widely beleived that the health status of labour market participants has a key role when individual workers make retirement decisions. According to the classic predicions from Grossman, health stocks determine healthy time spent on productive activitives with health stocks deteriorating over time so that in some way Grossman predicts that leisure becomes more valued as workers reach the later stages in the labour market. From here, comment on all written literature up to date.

3 The 2004 SHARE survey

The empirical estimates in this paper are based on the 2004 wave of the Survey of Health, Ageing and Retirement in Europe (SHARE 2004, second release). This survey collects and disseminates micro-economic data from 12 distinct countries (11 European economies and Israel) using a computerized assisted personal interviewing system (CAPI). The data provides social, economic and demographic information as well as extensive information on health and labour market indicators from individuals representative of the population aged 50 and over in each of the surveyed countries. The first wave was collected information with reference to the year 2004,

thus eligible respondents are defined as those born before or during 1954. The original sample consists of 31,115 respondents but this drops to 27,353 once the sample is weighted such that all estimates can be used to make inference from the sample to the finite population.¹ The weighted sample represents 11 countries: Austria, Belgium, Denmark, France, Germany, Greece, Italy, The Netherlands, Spain, Sweden and Switzerland.

The core questions in the survey are identical in all countries and for either males or females. In particular, answers by eligible respondents help to classify the contemporaneous labour market status of individuals into one of five possible categories: retired, working (employee or self-employed), unemployed, permanently sick or disabled and houseworker (wife or husband). The main concern in this paper is to understand the causal effect of health on retirement decisions conditional on observed characteristics. This means that our initial sample selection criteria consist on selecting individuals who signal present or past active labour market participation. To this aim, we first eliminate 3,557 individuals whose information set suggests non-participation in paid competitive labour market activities throughout their working-age period: most of these are females who claim housework as main occupation.² Secondly, we drop 539 retirees who claim to be non-citizens of the country where the survey takes place: we need to drop these individuals because later it becomes crucial to condition on institutional characteristics that could affect the retirement decisions of individuals. Since the survey does not directly ask individuals about their nationality we need to eliminate those who may have retired under institutional conditions different than those associated with the country where the survey takes place. Thirdly, we drop 589 German citizens who were inhabitants of the former DRG in or before November 1989 because the retirement conditions of these individuals could be the result of specific re-unification laws not represented by past or present retirement institutional conditions in Germany. Finally, we drop 244 individuals who claim to have become permanently disabled as direct result of work injury in their last place of work: later in the paper it becomes crucial to understand the effect

¹In contrast to the first release for the first wave (2004), the second release of the 2004 SHARE data further includes Israel. However, at the present time the data from Israel remains preliminary with all 2,595 units from such country assigned a zero weight. Furthermore, a zero weight is also assigned to answers from 1,167 non-eligible spouses (i.e., younger than 50 in 2004) who were also surveyed alongside their eligible partners. See www.share-project.org for more details on the designed and possible selection of the different sample weights.

²The cleaning process eliminates individuals who explicitly claim to have never worked for pay (2,677) most of which are females houseworkers. But we also eliminate 880 female houseworkers with ambiguous answers to questions referring to labour market status (e.g., no answer when asked about having worked for pay, respond ‘don’t know’ to ‘year of retirement’ or declare zero contributory pensions)

of health events (first) on the decision to retire (later); individuals who retire with permanent disabled as direct result of work injury describe a situation where the health condition and the retirement decision occurs simultaneously and not sequentially. In total there are 4,929 individuals dropped from the initial 27,353 weighted sample leading to an initial selected sample of 22,424

It is important to select out a benchmark age for retirement; early retirees (e.g., at an age below 40) face different retirement decisions than individuals who may have contributed for longer periods as active labour participants. Using the 2004 SHARE data Meijer, Kapteyn and Andreyeva (2007) study the effect of hand grip measures on the retirement decision of individuals that retired at age 50 or later. In contrast we decide a benchmark retirement age of 40 (and over), the reason being that a contribution of 20 years or more to the labour market is often enough time to affect the individual's future pension's rights (e.g., countries such as Italy or Austria demand a minimum of 20 contributory years in order to access the contributory old age pension). Since labour market participation often starts at the age of 16 (possibly earlier for individuals born in 1954 or before) considering those who may decide to retire at ages 40 and later suggest a sample of (past or present) labour market participants with relatively homogeneous retirement decisions, at least as far as considering institutional characteristics (e.g., the consideration of penalties associated with early retirement). There are 1,932 individuals in the 22,424 set who claim to have retired at the age of 39 or younger: 88.1% of these are females that claim some form of participation but who retired in their middle 20s, possibly as consequence of marriage and family. We eliminate these 1,932 individuals from our sample. Finally, given that our central interest is the causal effect of health on retirement decisions we need to eliminate anyone who may have suffered the consequences of some adverse health condition before entering the labour market. Later (see Section 3.2) we define distinct adverse health conditions and their timing; at this point we make the observation that 760 out of the 20,492 remaining individuals declare the onset of some adverse health condition (chronic, acute or mental disorder) while (or before) completing their start up education.³ The health condition of these 760 individuals could

³The survey provides full information with regards to number of years in education. We use this information to determine an age against which to compare the age of the first significant health event (if one ever happened). Individuals claiming zero years of education are eliminated from the sample if the health event occurred at ages 16 or below. All other individuals claim a number of years in education equal to 7 or greater (up to a maximum of 22). Assuming that education starts at ages 6 or above, the individuals are eliminated from the sample if the health shock occurred during education or within a margin of plus 4 years after having finished their formal

have determined their full labour market history (including education, working position and other important covariates) and this is why we need to eliminate them thus avoiding potential endogeneity effects in our final estimates. Altogether this leads to a final sample selection size of 19,732 from the initial 27,353. The selected units represent a sample of heterogeneous individuals in socio-economic and health characteristics, that have been or still are labour market participants and with homogeneous potential retirement outcomes.

Table 1 shows the distribution of the selected sample according to declared labour market status and compares the population proportions of the 19,732 against the original (27,353) population proportion. The same table shows the distribution by country and gender of the selected sample and again it compares these against the distribution of the 27,353 original sample among the 11 selected countries.

Table 1: Sample distribution by labour market status, gender and country (SHARE 2004 survey on 11 selected countries)

	Selected Sample (n=19,868)	Males (n=11,172)	Females (n=8,696)	Original Sample (n=27,353)
Labour Market Status ⁽¹⁾				
Retired	11,517 (0.59)	6,635 (0.59)	4,882 (0.60)	13,508 (0.50)
Employed	6,873 (0.33)	3,882 (0.34)	2,991 (0.32)	7,532 (0.26)
Unemployed	684 (0.03)	338 (0.03)	346 (0.04)	849 (0.03)
Permanently Sick/Disabled	362 (0.02)	198 (0.02)	164 (0.01)	885 (0.03)
House-worker	308 (0.01)	12 (0.001)	196 (0.02)	4,294 (0.16)
Unclassified ⁽²⁾	224 (0.02)	107 (0.01)	117 (0.02)	285 (0.01)
TOTAL (Sample units)	19,868	11,172 (0.57)	8,696 (0.43)	27,353
TOTAL (Population)	74,309,076	42,432,127	31,876,949	108,010,728
	Selected Sample (n=19,868)	Males (n=11,172)	Females (n=8,696)	Original Sample (n=27,353)
Country of Survey				
Austria	1,470 (0.03)	725 (0.02)	745 (0.03)	1,849 (0.02)
Belgium	2,557 (0.03)	1,503 (0.03)	1,054 (0.03)	3,619 (0.03)
Denmark	1,419 (0.02)	678 (0.02)	741 (0.03)	1,614 (0.02)
France	2,304 (0.20)	1,206 (0.19)	1,098 (0.22)	3,036 (0.18)
Germany	1,775 (0.25)	988 (0.24)	787 (0.27)	2,930 (0.28)
Greece	1,902 (0.04)	1,197 (0.04)	705 (0.03)	2,659 (0.03)
Italy	1,748 (0.21)	1,081 (0.22)	667 (0.19)	2,507 (0.20)
Netherlands	1,955 (0.05)	1,206 (0.05)	749 (0.04)	2,859 (0.05)
Spain	1,351 (0.12)	903 (0.14)	448 (0.08)	2,341 (0.13)
Sweden	2,673 (0.04)	1,300 (0.04)	1,373 (0.05)	2,994 (0.03)
Switzerland	714 (0.02)	385 (0.02)	329 (0.03)	945 (0.02)
TOTAL	19,868	11,172	8,696	27,353

Note: Bracketed values are population probabilities based on the weighted sample using individual calibrated sample weights. The 19,868 selected units represent 74,309,076 individuals in the overall population of the selected countries: the original 27,353 represent 108,010,728 individuals. ⁽¹⁾The original variable that explains the labour market status of individuals in the survey can be found in the employment module (ep005_). ⁽²⁾Unclassified covers those who answer to the variable ep005_ with 'Don't know', 'Refuse to declare', 'Other status not in the list', or do not answer at all (missing values).

education.

Table 1 shows that eliminating non-participants (mostly houseworkers) from the original sample inflates the population proportion of ‘retired’ and ‘working’ in the selected sample. The remaining houseworkers (308) provide sufficient information to identify them as past active labour market participants that have subjectively classified themselves as houseworkers in 2004. The weighted distribution by country using the original sample reflects the contribution of units according to actual country population size.⁴ Relative to the original sample, there are two countries where the population proportion drops; Germany and Spain. Germany’s proportional drop results from the elimination of former DRG inhabitants whereas Spain’s drop may be explained because we have eliminated most houseworkers from the sample thus reflecting the low labour market participation of Spanish females born before the 1950s. The proportional drop from Germany and Spain relative to the original sample slightly inflates the sample representativeness of both France and Italy. Despite this the population representativeness for any given country in the selected sample remains very close to that of the original sample, i.e., our sample selection criteria does not distort the country’s representativeness intended by the survey.

The selected sample of 19,868 become the basis for us to analyse individual’s retirement decisions allowing for the potential effect that the onset of some adverse health condition may have on individual’s productive capacity. The following sub-sections explain in detail how the 2004 SHARE data can be used to elicit information on different retirement outcomes, health indicators and information on important social and economic indicators that may also be causal with regards to individual’s retirement decisions. Institutional indicators are drawn from sources outside the 2004 SHARE data.⁵ To as far as possible, all variables (outcomes and covariates) are carefully constructed to control for all possible endogeneity effects and selection bias that are likely to be a cause of concern in the presence of cross-sectional data.

3.1 Retirement Indicators: The outcome

The 2004 SHARE asks surveyed individuals to classify themselves as either retired, working, unemployed, permanently seek/disabled or as houseworkers: self-reported houseworkers in the 19,868 sample are in fact retirees with a history of past labour market participation. With

⁴For example, Belgium contributes with 3,619 units to the original sample. With a population of approximately 10 million in 2004 the 3,619 units represent only 3% of the overall 108,010,728 represented by the 27,353 units in the survey. Alternatively, Spain contributes with 2,341 units that account for 13% of the population represented by the 27,353; the 13% reflect Spain’s population size in 2004 (approximately 40 million).

⁵See Section 3.4 and Appendix A for further details.

this, we consider the sub-categories retired (11,517), permanently sick or disabled (362) and contemporaneous houseworkers (208) to be ‘permanently retired’. The alternative classification is that of ‘working’ at the time of the survey with a total of 6,874 individuals classified as such. However, 218 of these are part-time workers with ages equal or greater than the ‘official early retirement age’ living in countries where it is possible to combine ‘early retirement’ with part time work (see Section 3.3). These 218 individuals are effectively ‘retired’ and we reclassify them as such. Moreover, we eliminate individuals who classify themselves as ‘working’ but claim disability or unemployment insurance or claim ‘unemployment’ as main status; all these individuals (a total of 853) have ambiguous labour market classification (retired or working) because to be on disability and/or unemployment insurance at ages 50 and above often signals indirect routes of early retirement in European economies (e.g., see Linderboom and Kerkhofs, 2007).⁶ Finally, 223 individuals do not respond to the question on labour market status. Using complementary information (e.g., declared retirement year, pension data or income data) we reclassify 62 of these 223 missing units as retired or working; the remaining 161 remain with unknown labour market classification. In total, the 19,868 are such that 12,342 are classified as (permanently) retired, 6,498 are classified as working in 2004 and 1,028 remain unclassified.

The outcome of interest in this paper is ‘early retirement’, an event that occurs if an individual leaves paid employment before the official age of retirement, i.e., the minimum legal age at which an individual is allowed to draw from the basic statutory old-age pension without any penalties. Two variables are needed to build the ‘early retirement’ indicator; *(i)* the age at which the individual retired (if classified as retired) and *(ii)* the individual statutory minimum official age of retirement. The latter varies by country of residency but it may also depend on gender (Austria, Belgium, Greece, Italy and Switzerland), particular disabilities (e.g. Italy), it can change within country according to year of birth and it may also depend on having lived in the country for a minimum number of years as an official resident (Denmark and Sweden). The 2004 SHARE data provides information on individual’s age, gender, long-term citizenship and residency;⁷ these variables together with external sources of information on the conditions for legal retirement for each of the 11 countries helps us to assign an ‘official retirement age’ to

⁶We define a status of retirement as an absorbing state. Unambiguous labour market classification could in principle violate this assumption. For example, a 57 year old individual self-classified as working but with an income source based on unemployment or disability insurance does not guarantee to be in an absorbing state.

⁷It also provides information on specific disabilities sometimes required to assign specific legal retirement age (e.g., blindness in Italy). See Section 3.3 for more details.

each of the 19,868 individuals.⁸ In our sample this age ranges from as low as 55 (blind female Italian resident) to as high as 67 (Danish resident). At the same time, a question in the 2004 SHARE data elicits the ‘year when last worked’ from all those classified as retired, permanently seek or disabled and houseworkers. Only 11,691 individuals from the 12,342 classified as permanently retired respond with a specific retirement year; the remaining 651 do not provide direct information on ‘year of retirement’. However, using ‘number of years worked in the last place of work’ and ‘number of years in education’ we reduce the non-respondents from 651 to 89 units.⁹ This amounts to a (weighted) nonresponse probability of 0.01 among the retired sample: we consider this probability to be sufficiently low so that we ignore the 89 units from our analysis. For all other individuals we define a binary outcome ‘*REBORA*’ to be 1 if the individual is observed as an early retiree (i.e., if the age at which the individual retired is below his or her official retirement age by at least one year), and zero otherwise (i.e., if the individual retired at the corresponding retirement age or if the individual is still working at the time of the survey and, therefore, by definition is not – at least in 2004 – an early retiree).

In addition to the indicator ‘*REBORA*’ we define the complementary retirement indicator ‘*REBERA*’ with reference to the ‘official *early* retirement age’. That is, most of the 11 countries represented in the sample distinguish between ‘an official age of retirement’ and ‘an official *early* age of retirement’, i.e., an age when individuals can start to draw from the statutory old-age pension despite not having reached the minimum official age of retirement and at the expense of some pre-determined economic penalty. For example, the minimum official retirement age for Austrian males is 65. However, they can retire officially between the ages of 61 and 64 at the expense of losing 4.2% from their old age pensions for each year of anticipated retirement.¹⁰ The ‘official early retirement age’ is country specific and gender dependent but for most countries (Austria, Belgium, Denmark, France, Germany, Greece, Italy, the Netherlands

⁸Section 3.3 provides specific details on the legal requirement for any of the 11 countries, including legal changes to retirement laws since the 1950s. This takes into account that the oldest member in our sample was born in 1900.

⁹The 218 reclassified from ‘working’ to ‘early retired’ are assigned the official ‘early retirement age’ (see Section 3.3). A further 344 individuals claim to have worked in the last place of work for a period that, joint with the number of years in formal education (allowing an starting age of 6) is either equal to or greater than their corresponding minimum official retirement age. These individuals are imputed an age of retirement equal to the age of retirement corresponding to the last year in their last place of work. This leaves 89 unresolved cases.

¹⁰Below the age of 61 an Austrian male can choose to retire but cannot (on average) draw from the old age pension until reaching the age of 61. At that point he pension will be adjusted according to contributory years.

and Spain) it also depends on the number of contributory years. Following the previous example, an Austrian male can retire earlier than at 65 but only if the individual has contributed with 40 years or more towards his pension. Another example is France: this country does not have ‘official *early* retirement age’ but French individuals that have contributed 40 years or more can retire thereafter even if they are not yet 60, the latter being the minimum official retirement age in France. Following this rubric the binary indicator ‘*REBERA*’ equals 1 if an individual retires before reaching the ‘official *early* retirement age’ and zero otherwise. In the case of Spain (if started working after 1967), Sweden and Switzerland the official age of *early* retirement depends only on the age and gender of the individual. For all other countries (and for Spaniards who started working before 1968) we need to know each individual’s total number of years contributing towards their pensions in order to define the outcome ‘*REBERA*’. Unfortunately, the 2004 SHARE data does not ask individuals to declare number of years in the labour market or year of entry into the labour market. One variable that provides some information with regards to contributory years is ‘number of years worked in the last job’. We use such variable to approximate ‘*REBERA*’ for the 7 pertinent countries allowing a 10 year margin with regards to the ‘number of years worked in the last job’. For example, a French male who retired at 57 and claims to have worked 30 years or more in his last work place is classified as ‘early retired’ (*REBORA*=1) because he does so before the age of 60. However, if we assume 10 years of work experience previous to the 30 years in his last work place this individual has retired with 40 contributory years, i.e., with the full benefits implied by the ‘early retirement laws’ in France; in this case *REBERA*=0. Alternatively, a French male that retires at 57 and signals 30 contributory years (allowing for the 10 year margin to the number of years worked in the last place) does so without the full benefits implied by the French ‘early retirement laws’ and, therefore, the individual is assigned a value of *REBERA*=1: whereas *REBORA* indicates ‘retirement before official retirement age’ *REBERA* signals ‘retirement before official early retirement age’. The two indicators differ in that each one indicates difference in monetary compensation as result of different retirement decisions, that is, those with *REBORA*=1 are penalized with a reduced old age pension whereas those such that *REBERA*=1 retire under conditions such that they have to wait for some time (e.g., until reaching the age of 61 in Austria) before they may be able to access their old age pensions.¹¹

¹¹ Austria, Belgium, Denmark, Germany, Greece, Italy and for Spaniard that have entered the labour market in or before 1967 have an ‘official early retirement age’. If individuals have contributed with sufficient number of

The binary indicators *REBORA* and *REBERA* are central to our empirical analysis but we consider two further related outcomes both of which are continuous in nature; ‘retirement age’ (*REAGE*) and ‘total annual received contributory pensions’ (*TOTPEN*) where the word ‘total’ refers to the addition of old age pensions (or old age early retirement pension when it applies), disability and injuries related pensions, private occupational pensions (including early retirement private occupational pensions) as well as any other old age related pensions (e.g., war pensions). Comparison of total pensions received among individuals with different retirement histories shows the monetary effect of retiring at an age distinct from the individual specific official retirement age whereas ‘retirement age’ is well defined for anyone that is retired (i.e., early retired or otherwise).

The four retirement outcomes of interest apply differently to different sub-groups in the represented population. For example, we can contrast the outcomes *REBORA* and *REBERA* among those who are still sufficiently young to be classified as of ‘*working age*’ whereas the outcome *TOTPEN* is only relevant for those classified as ‘*retired*’ and not for those actively working at the time of the survey. Furthermore, relative to individuals that have already reached their retirement year, those who are still of working age in 2004 are homogeneous with respect to age (fewer cohorts represented in such sub-group), more homogeneous with respect to how macro-economic cycles and shocks may have affected their labour market histories, more homogeneous (within countries) with regards to changing social security laws - including laws regarding retirement - and even more homogeneous with respect to cultural changes affecting participation in the labour market (e.g., female participation). But the sub-group who are still of working age in 2004 includes both individuals who are still working and early retired individuals.¹² In fact, the ‘early retired’ individuals in the working age sample are included in the ‘retired’ sample, where the latter includes working age individuals who retired at least one year earlier than their official age of retirement, individuals who have already reached and/or passed their retirement age and

years towards their pensions and retire between the defined ‘official early retirement ages’ (e.g., between 61 and 64 in Austria) they are defined as *REBERA*=0 since they do not retire earlier than the officially designed early retirement years. However, *REBERA*=1 if the individual has not contributed sufficiently and retires before the official age of retirement (e.g., before 65 in Austria) because such individual does not benefit from ‘official early retirement laws’. Likewise, *REBERA*=1 when retirement is one of more years before the ‘official *early* retirement age’.

¹²Individual due to retire officially in 2004 and self-classify themselves as working in 2004 are classified as working. Individuals who are due to retire officially in 2004 and self-classified themselves as retired, are treated as ‘non-early retired’. Early retired must have done so one year or more before their official retirement age.

retired at least one year earlier than their official age of retirement and retired individuals that did so at their assigned expected age of retirement. Throughout our analysis we will distinguish between the ‘working-age’ sample (*WS-sample*) and the ‘retired’ sample (*RS-sample*). Table 2 summarizes the four retirement outcomes of interest for the 19,868 selected sample according to sub-groups (*WS* and *RS*) and country of origin. The number of units in the *WS-sample* is 9,762 whereas the number of units in the *RS-sample* is 12,253: the overlap between the two samples (i.e., of working age but already retired in 2004) equals 3,390, although only 587 of these have correctly retired in 2004 according to their official retirement age; the other 2,812 are classified as ‘early retired’ (i.e., retired in 2003 or before). The numbers 9,762 and 12,253 exclude units that are unclassified, i.e., units with unclassified labour market status in 2004 (1,028), retired individuals with unknown retirement age (89) and 126 individuals who are still working in 2004 despite having passed their minimum age of retirement: these latter group are not relevant to our analysis since they not retired but neither are they of working age. Estimates in Table 2 show that the retired sub-group spent less time in education than the working age sub-group; the result remains when comparing sub-groups within country. The difference in education is likely to reflect cohort differences, i.e., the sample of retired includes a larger number of individuals from older cohorts who would not have benefited from the gradual establishment of universal education during the first two quarters of the 20th century, specially in the case of Mediterranean countries (Greece, Italy and Spain). However, difference in education between the 2,812 working age individuals that have retired early and the 6,372 working age individuals that remain active (on average, 10.5 and 12 years of education, respectively) reveals a difference (1.5 years) that is statistically significant (t-value = -4.1) thus suggesting that early retirement is associated with lower levels of start up human capital. Table 2 also shows a distinction in gender: the data cleaning process has deleted the non-participants so that in countries where female labour market participation has been traditionally lower (i.e., Spain) the percentage of males is statistically higher. Finally, Table 2 shows that average total annual pensions are about €2,000 lower when the average is estimated with a larger percentage of ‘early retired’ (i.e., when based on the retired among the working age sub-group). Finally and as expected, the percentage who retire early, i.e., *REBORA*, is higher than the percentage who retire earlier than at their official *early* retirement age, i.e., *REBERA* (see Section 3.3 and Appendix A for more details).

Table 2: Selected statistics by sample and sub-group definition (11 selected countries)

	Size	Average age	Probability male	Average years in education	Avg. years worked last place ⁽²⁾	Probability REBORA	Probability REBERA	Avg. age of retired ⁽³⁾	Average pensions if retired ⁽⁴⁾	
Complete Sample	19,868 (1.00)	65	0.57	10.0	26.0	0.43	0.34	58.8	€16,859	
Working Age (WS)	9,762 (0.50)	57	0.62	11.5	23.2	0.28	0.25	55.4	€14,654	
Retired (RS)	12,253 (0.67)	71	0.57	9.1	28.1	0.69	0.52	58.8	€16,859	
Overlap (WS & RS)	3,390 (0.24)	60	0.67	10.4	26.3	0.86	0.77	55.4	€14,654	
Unused units ⁽¹⁾	1,243 (0.07)	62	0.58	9.7	28.7	--	--	--	--	
	Working Age Sample (WS), n=9,762					Retired Sample (RS), n=12,253				
	Sample Size	Avg. Age	Percent males	Avg. Education	Percent BO/BE ⁽⁵⁾	Sample Size	Avg. Age	Percent males	Avg. Education	Avg. age retired
Austria	652	57	0.63	12.3	0.45/0.42	1,105	69	0.49	11.3	57
Belgium	1,231	57	0.63	11.6	0.30/0.25	1,614	70	0.59	10.3	58
Denmark	815	58	0.50	13.5	0.31/0.17	788	71	0.43	12.0	61
France	877	55	0.56	11.0	0.14/0.12	1,397	71	0.52	8.1	59
Germany	960	57	0.56	14.2	0.30/0.29	1,041	71	0.52	13.1	60
Greece	976	57	0.72	11.2	0.29/0.18	1,123	70	0.60	7.7	60
Italy	770	57	0.74	9.5	0.42/0.38	1,259	70	0.59	6.6	57
Netherlands	1,144	57	0.62	12.1	0.25/0.09	1,085	70	0.61	10.7	58
Spain	589	56	0.69	7.8	0.24/0.21	821	72	0.75	5.1	61
Sweden	1,395	57	0.53	11.2	0.22/0.17	1,632	71	0.46	9.5	61
Switzerland	353	56	0.60	13.3	0.12/0.11	388	73	0.47	11.5	62

Note: All figures are based on weighted values using individual calibrated sample weights. ⁽¹⁾These include 183 working but claiming disability or unemployment insurance, 684 directly classified as unemployed and 161 with unknown answers to the question 'self-asses your labour market status': these 1,028 cannot be assigned a labour market status. Furthermore, there are 89 retired with unknown retirement age and 126 individuals who are still working but are outside the working age group since their age is greater than their minimum retirement age. ⁽²⁾ The average number of years worked in the last place implies 'current' working place for those currently working or 'last' work place for the retired. Only the respondents are accounted for with 7% nonresponse rate in the full population. ⁽³⁾ Average retirement age takes into account both the early retired and those retired at the corresponding retirement age those retired are accounted for. This explains why the average age of retirement in the WS-sample equals that in the Overlap (WS-RS) sample as well as why the average retirement age in the complete sample equals the average retirement age in the retired sample. ⁽⁴⁾ All money quantities are in Euros and deflated to reflect purchasing power parity between countries, where the PPP deflator is provided in the SHARE survey. Not all individuals classified as retired receive contributory pensions: 6% classified as 'retired' declare zero total contributory pensions. These (679) individuals do in fact signal participation in the labour market with full information on relevant variables, therefore, we choose not to eliminate them from the sample. When analyzing 'total pensions' we assume zero total contributory pensions for them. ⁽⁵⁾ BO stands for 'REBORA' and BE stands for 'REBERA'.

We select three sets of information that could be thought as potentially causal to the retirement outcomes just defined. These sets are 'institutional characteristics' (e.g., financial penalties for early retirement), individual's social and economic characteristics (e.g., gender, education, number of dependent children before retirement, etc.) and indicators that should capture the impact of adverse health events on the productive capacity of individuals. The following subsections explain and motivate the choice of variables for each of the three sets of information. We start by describing how the 2004 SHARE data helps to construct indicators that explain the onset of health conditions among the selected sample.¹³

3.2 Health Indicators

The 2004 SHARE data is very informative on the health status of individuals with two modules designed specifically to pick up detailed information about the physical and mental health of the

¹³From now on, our discussion excludes the 1,243 individuals with unknown values for the outcome variables so that the sample drops from 19,868 to 18,825. The 1,243 account for 7% of the weighted sample but only 250 of these (161 and 89) may be thought as potentially causing sample selection problems (see footnote in Table 2).

underlying population. To a large extent all questions in the two health modules are designed similarly to those found in the Health and Retirement Study (HRS) where the latter is the USA counterfactual survey to the SHARE study.¹⁴ All health indicators in the 2004 SHARE can be classified into two broad categories: subjective and objective health indicators. Subjective health indicators pick up individual's subjective feelings with regards to their health status, for example, classifying their health subjectively as 'very good', 'good', 'bad' or 'very bad'. Answers to subjective variables can indicate true health status but are often biased by effects that may not necessarily relate to health, for example, the effect of *justification bias* by early retirees (*reference*) or the effect of cultural differences that may determine individual's answers (*reference*).¹⁵ In particular for our analysis, the use of subjective health measures from a single cross-section is not adequate because such measures reflect health status of individuals at the time of the survey (2004) and not at key moments over their life-cycle, for example, the health status of retired individuals before they become retired. An adequate alternative for our purpose is the use of self-reported objective health conditions to construct indices of health status. By 'objective health' we mean the diagnoses by medical doctors of particular medical conditions that may be classified as either the onset of some chronic illness (e.g., diabetes), the occurrence of an acute health event (e.g., bone fracture) or the clinical diagnoses of some mental disorder (e.g., clinical depression). The 2004 SHARE data is very informative with respect to objective measures of health, not just on the occurrence of the event (i.e., *Has a doctor ever told you that you have had a given condition?*) but with respect to the timing of the event (i.e., *How old were you when you were first told by a doctor that you had the given condition?*). In particular, the fact that individuals are asked to declare the timing of health events implies that we know the health status of individuals from birth. This implies that we can build indicators of health status with reference to the timing of individual's labour market decisions over their working lives (e.g., before retirement for those who are retired). Using the physical health module in the 2004

¹⁴The HRS is a longitudinal biannual panel study that started in 1992 collecting information on health, retirement and other socioeconomic characteristics of the USA elderly population and is sponsored by the National Institute of Aging at the University of Michigan (USA). See www.hrsonline.isr.umich.edu

¹⁵For example, it has become more common to take measure of health related to measured abilities to perform physical tasks such as hand grip, a measure also available in the 2004 SHARE survey. Although this is an objective measure of health deemed by many authors in the literature (Kaptayn et al., 2007 and reference therein) to be an objective reliable measure, for all we know the measured 'hand grip' may be the result of retirement decisions in the past.

SHARE data we find as many as 17 measures of objective health events, namely ‘heart attack’, ‘stroke or cerebral vascular problems’, ‘high blood pressure’, ‘high blood cholesterol’, ‘diabetes’, ‘chronic lung disease’, ‘asthma’, ‘arthritis’, ‘osteoporosis’, ‘cancer or malignant tumor’, ‘stomach, duodenal or peptic ulcer’, ‘Parkinson’s disease’, ‘cataracts’, ‘hip fracture or femoral fracture’, ‘deafness’ and ‘blindness’: a final category gathers any residual health condition as ‘any other diagnosed conditions not previously mentioned’ and we take answers to this question as another possible health condition.¹⁶ Using the mental health module we define the indicator ‘mental health condition’ to be one if the person declares to have suffered depression and at the same time has been treated for such condition either by a medical psychiatrist, a mental hospital or in a psychiatric ward. In total we end up with 18 distinct health indicators (type and timing) and take all of these to assess if and when an individual has ever suffered the onset of one or more adverse health condition.¹⁷ The events ‘heart attack’, ‘strokes’, ‘stomach ulcers’, ‘cataracts’ and ‘bone fractures’ have an immediate effect on the productivity of working individuals but can often be reversible; we classify these five conditions as ‘acute’ conditions. All other conditions are classified as ‘chronic’ conditions except for ‘mental health problems’ that we leave in a category of its own.¹⁸

¹⁶The SHARE data does not provide further details about these ‘other non-listed’ conditions. However, the HRS survey (any year) does ask individuals to specify what they mean by ‘other’ and this helps us understand that the residual category could in principle include medical problems such as ‘multiple sclerosis’, ‘chronic kidney disorders’, ‘severe allergies’, ‘endocrine disorders’, ‘other cancers’, ‘HIV/AIDS’, etc. All of these ‘other’ conditions listed by respondents in the HRS are conditions that should be classified as non-reversible chronic conditions.

¹⁷The difference in the degree of intensity between conditions implies difficulties when allowing for all of them to be equally indicative of ‘health events’. For example, some diabetes require medication only in the long run whereas when diagnosed they may just require changes in lifestyle. Ideally we would need to have the ‘intensity’ of the health event when it was first diagnosed. The ‘intensity’ variable at the time of diagnoses is not in the data. Alternatively, we know if the person is *currently* taking medication for the condition ‘at least once a week’. For the conditions ‘diabetes’, ‘high blood cholesterol’, ‘high blood pressure’, ‘arthritis’, ‘asthma’ and ‘osteoporosis’ we use the variable ‘taking drugs at present’ to identify the intensity of the condition. Although this is a weak indication on the intensity of the health event(s) when these occurred for the first time, it is clearly the case that not taking drugs at present would indicate no intensity in the past. Likewise, the ‘mental health condition’ is considered as such not only if the person declares to have had a depressive episode at some age, but if such episode required medical intervention.

¹⁸In epidemiology the name ‘acute’ is associated with conditions that last for a short time and can be fully reversed, as opposed to chronic conditions that may start with low effects to progress towards full disability (e.g., diabetes leading to blindness). In the case of heart attack there is a distinction between a sudden heart infarction (acute) and congestive heart failure (chronic). Unfortunately the 2004 SHARE data does not make a distinction

Our primary goal is to understand the effect of an adverse health condition on the retirement outcomes previously defined when such condition is diagnosed before the timing of the outcome.¹⁹ Our sample is such that 7,179 of the 18,625 selected individuals have never suffered the onset of any of the 18 health conditions previously defined. The remaining 11,446 have suffered the onset of one or more health conditions (acute, chronic or a mental disorder) at some point after having entered active labour market participation.²⁰ Unfortunately, 154 of the 11,446 do not provide information on the age of onset of at least one of the suffered health problems. Since no other variables in the survey helps us to know the ‘age of onset’ of the health conditions for these 154 individuals, we cannot assess if they suffered the health problem before or after the timing of the retirement outcome. This is why we eliminate the 154 individuals from our sample, leaving a final sample size of 18,471 where 7,179 have never suffered an adverse health condition and 11,292 (or 62% of the weighted sample) have suffered the onset of one or more of the 18 health problems considered. At this point we define the key treatment variable *hs_bora* that equals 1 if an individual has suffered the onset of at least one health condition before the timing of the outcome (i.e., *before retirement* if retired or *before 2004* if not retired), an 0 otherwise. Only 7,930 of the 11,292 (i.e., 43% of the weighted sample) suffered their first health problem at least one year before the timing of the retirement outcome so that for these individuals *hs_bora*=1: the other 3,362 that are also defined as having suffered a health condition suffered such event after retirement, simultaneously with retirement or in 2004 if not yet retired by then: for these 3,362 *hs_bora*=0 since the treatment ‘onset of the health condition before the timing of the retirement outcome’ does not apply. Likewise, *hs_bora*=0 for anyone that has never suffered the onset of an adverse health condition. Table 3 shows the distribution of the sample according to number and type of health status (acute, chronic or mental disorder)

between these two (whereas in the HRS the distinction is clearly made). We take all who answer to having had a ‘heart attack’ to be of the acute nature. It is also difficult to classify ‘mental health problems’ as acute or chronic and this explains why we have isolated this condition in a unique category.

¹⁹That is, individuals may or may not suffer an adverse health condition and may or may not retired as consequence of it. In what follows, the sentence ‘retirement outcome’ refers to any of the four outcomes defined in Section 3.1, while ‘timing of the retirement outcome’ means the time of retirement (if retired by 2004) or the time just before 2004 if the individuals is still working at the time of the survey.

²⁰At this point we recall that the onset of a health condition happens 5 year or after having finished their formal education. The choice of a 5 year benchmark aims at ensuring that we capture health events that happen after entering the labour market where the variable ‘year of entry into the labour market’ is missing from the data. With this, 352 individuals have been eliminated as result of premature health shocks (see Section 3.1).

and compares the retirement status of individuals by health condition. The same table compares the health condition of the sample by country of residence and retirement status. The estimates in Table 3 show that suffering the onset of a health condition in the working age sample leads to a 29% early retirement rate, 8% more than the probability of early retirement among those of working age with $hs_bora=0$. Despite this large difference in the probability of early retirement the difference on average retirement age between the two sub-groups is very similar (i.e., 55 and 56, respectively). Based on those who suffer only one type of health condition (either chronic type, acute type of mental disorder), the onset of a mental health disorder happens earlier (on average at age 37) than the onset of chronic or acute conditions, and although the probability of early retirement is similar among the three types (38%, 41% and 40% respectively), those who experience mental health problems retire on average 2 to 3 years earlier than sufferers of chronic or acute conditions. Comparing countries we see that except for Austria and Switzerland, on average there are similarities with respect to the probability that their populations suffer a health condition before retirement, similarities with regards to the age at the onset of the first condition and similarities in the probability distributions among the three types of health conditions. We notice that in Austria the probability of suffering a health condition before retirement is 35%, that is, 10% less likely than the average probability in the population of 11 countries. However, once Austrians suffer a health condition the probability of early retirement is 60%, i.e., 16% more likely than those who suffer some health condition before retirement among all the 11 countries.

Table 3: Distribution of health events by type of event, first onset and country (11 selected countries)

	Size	Average number of health conditions	Average Age at onset of first health condition	Probability of Early Retirement by 2004	Avg. retirement age if retired by 2004
Distribution of sample:					
Never suffered a health event ⁽¹⁾	7,179 (0.39)	0.00	--	0.34	59
WS-Sample & Never	4,960 (0.67)	0.00	--	0.21	55
RS-Sample & Never	3,498 (0.49)	0.00	--	0.69	59
Health event after retirement	3,362 (0.18)	1.86	65	0.74	58
HS_BORA=1 ⁽¹⁾	7,930 (0.43)	1.87	44	0.44	59
WS-Sample & HS_BORA=1	4,272 (0.50)	1.55	43	0.29	56
RS-Sample & HS_BORA=1	5,310 (0.67)	2.10	45	0.66	59
Distribution by type (if HS_BORA=1):					
Only suffered CHRONIC conditions	4,444 (0.53)	1.35	45	0.38	59
Only suffered ACUTE conditions	935 (0.09)	1.15	44	0.41	60
Only suffered MENTAL conditions	159 (0.03)	1.00	37	0.40	57
Suffered CHRONIC & ACUTE	1,899 (0.25)	3.10	43	0.56	60
Suffered CHRONIC & MENTAL	272 (0.05)	2.60	37	0.49	56
Suffered ACUTE & MENTAL	51 (0.01)	2.10	37	0.54	59
Suffered CHRONIC, ACUTE & MENTAL	170 (0.04)	4.50	41	0.71	57

	Size HS_BORA= 1	In-Country ⁽²⁾ Probability HS_BORA= 1	Percent 1st Cond Chron ⁽³⁾	Percent 1st Cond Acute ⁽³⁾	Percent 1st Cond Mental ⁽³⁾	Avg. Age, onset ⁽³⁾	Percent retired/ AGE ⁽⁴⁾
Austria	493	0.35	0.63	0.33	0.04	44	0.60 / 57
Belgium	1,037	0.44	0.52	0.38	0.10	43	0.52 / 58
Denmark	586	0.44	0.58	0.34	0.08	45	0.49 / 61
France	948	0.45	0.59	0.30	0.11	52	0.31 / 59
Germany	709	0.42	0.57	0.37	0.06	46	0.53 / 60
Greece	753	0.42	0.55	0.43	0.02	47	0.35 / 62
Italy	692	0.42	0.63	0.32	0.05	43	0.53 / 58
Netherlands	724	0.39	0.58	0.36	0.07	42	0.44 / 59
Spain	554	0.45	0.62	0.31	0.07	45	0.36 / 60
Sweden	1,201	0.47	0.53	0.40	0.07	43	0.39 / 61
Switzerland	233	0.34	0.65	0.28	0.07	47	0.29 / 63
TOTAL	7,930	0.43	0.59	0.34	0.07	44	0.44 / 59

Note: All figures are based on weighted values using individual calibrated sample weights. HS_BORA is an indicator that equals 1 (HS_BORA=1) if the individual has suffered a health condition before retirement. Adding 'never suffered a health condition' (0.39), suffered health condition after retirement (0.18) and suffered health condition before retirement, i.e., HS_BORA=1 (0.43) adds up to 100% of the sample. All those that suffer 'only' chronic or 'only' acute conditions may have suffered one or more chronic or one or more acute; however, a 'mental' condition can only average 1 because this category includes only one condition. ⁽¹⁾ The overlap between the WS and RS samples such that 'never suffered a health condition' is 1,279 whereas the analogous overlap in the case of HS_BORA=1 is 1,652 units. ⁽²⁾ In-country means that the probabilities are based on within country sample. For example, Austria contributes with 1,399 units (3%) to the 18,471 in the sample; 493 of the 1,399 are treated by a health condition before retirement (or 2004 if not retired) and these 493 account for 35% of all Austrians in the sample. ⁽³⁾ All percents based on the 1st Health Condition ever suffered (Chron(ic), acute or mental disorder). ⁽⁴⁾ The average age refers to the retired only.

Finally, we notice from Table 2 that the probability of early retirement in the population is 43%, while the probability of early retirement for those who suffered some health condition is 44% (Table 3). Therefore, unconditionally we would conclude that suffering the onset of a health condition is uncorrelated with the decision to retire early. In what follows we introduce a set of covariates (institutional and socio-economic characteristics) that will help us to clear up potential confounding effects when estimating the causal effect of the onset of some adverse health condition on the decision to retire. We start by describing variables that define the institutional characteristics that can underline the individual's decision with respect to labour market participation.

3.3 Institutional characteristics

Institutional characteristics are those that define the retirement conditions (laws and legislations) in each of the 11 countries. Our aim is to establish these characteristics for each of the individuals

in the sample. To do so we need to identify the legislation that defines the retirement rules for individuals living in different countries. Furthermore, we need to identify differences in legislation as result of within country changes over time to these retirement rules because such dynamic changes implies differences in the retirement legislation for individuals with equal nationality but who belong to a different age cohort. For example, the oldest person in our sample of 18,471 individuals is a Danish citizen born in 1900; for this particular individual it is likely that the appropriate Danish legislation governing his retirement decision would have been the one in force in the 1950s-1960s, whereas the legislation governing the retirement decision of the youngest Danish in the sample (born in 1954) would probably be the one in force in the late 1990s or early 2000s. In fact, since all individuals in our sample are born between 1900 and 1954, we need to know the retirement legislation for the 11 countries in the sample from the the 1940s (since we are taking individuals who retire at the age of 40 or more) up to our present days. Information with regards to historical retirement laws and legislation can be found in the OECD publications (J. Martin, Eds, 2007) as well as in the European Union information website on employment and social security.²¹ Much of the legislation governing the legal requirements for retirement refers to the official age of retirement or the minimum required number of contributory years that a person needs to have before he or she is able to draw from the national old age pension. The two outcome variables *REBORA* and *REBERA* (see Sections 3.1) already use all the information regarding official age of retirement, an age that changes for each individual according to country of residency, age cohort within country, gender, number of contributory years as active paid worker and in some cases (e.g., Denmark) number of years as legal resident in the surveyed country. Table A1 in Appendix A summarizes all the retirement legal changes that each country in the sample has experienced over the 20th century as well as more recent legislation that would have had an impact when assigning a legal age of retirement to each of the individuals in the sample.²² Besides all the institutional information that defines the

²¹See for example, http://ec.europa.eu/employment_social/missoc

²²Needless to say that we can only assume that the country where individuals are surveyed is the country that has governed each individual's retirement laws when they became retired (or will govern the future retirement rules of those in the sample who are still currently working in 2004). The survey does not ask individuals about their birth nationality or if they have always lived in the country where the survey takes place. We do eliminate individuals who claim not to be nationals of the country where the interview takes place. Our assumption of linking the legislation to the the country where the survey takes place is more likely to be undermined for respondents in Mediterranean countries (e.g., Spain and Italy) where some surveyed elderly workers would have been economic immigrants to Northern European economies in the 1950s-1960s that have returned to their countries of origin

individual’s legal age of retirement, there are other institutional legal characteristics that can be turned into variables and that may potentially affect the individual’s participation decision in the labour market. In particular we take the variables *(i)* ‘time of entry into the European Union’, *(ii)* ‘the introduction of supplementary contributory pension plans’, *(iii)* ‘the introduction of a National Health System’, *(iv)* ‘the possibility to combine part-time work while drawing from the national old age pension’, *(v)* ‘type of benchmark to adjust annual old age pensions’, *(vi)* ‘Annual percentage penalty to old age pensions if retiring early’ and *(vii)* ‘Annual percentage benefit to old age pensions if retiring late’. Clearly, the European Union changes the social security system for any country that becomes a member: for example, this would have been the case in Spain after entry into the European Union in 1986 and, therefore, Spanish individuals retiring before 1986 would have faced a different set of incentives than Spanish retiring after 1986. In some countries (Denmark from 1986, The Netherlands since the first introduction of the Social Security system, Sweden since 2001 and Switzerland since 1982) individuals can choose to contribute towards a supplementary pension plan to top up the basic old age pension: the introduction of a supplementary plan could have altered the retirement incentives that individuals would have faced compared to the incentives faced by similar citizens making retirement decision before the introduction of such supplementary plans. The introduction of a National Health System can also affect the retirement decisions for those who thus far had to face the possibility of having to pay for existing or future unexpected adverse health conditions. Some countries (e.g., Austria) use ‘wage inflation’ to adjust old age pensions annually whereas other countries (e.g., France) use the consumer price index to adjust annual old age pensions: the chosen benchmark (wage rate or CPI) is often linked to macro-economic performance that differs over time in ways that may affect individual’s retirement decisions. Finally, most countries have established penalties and benefits that aim at modifying the incentives of workers with respect to early retirement. For example, Germany penalizes early retiree with a drop in old age pensions of 3.6% for each year of early retirement, whereas Germans who decide to retire late (i.e., after their official age of retirement) see their old age pensions increase by 6% for each extra year worked beyond their official retirement age: each of the 11 countries has a different set of penalties and benefits as described in Table A1, Appendix A: unquestionably, economic penalties and benefits are designed to directly affect (conditional on number of years worked) the retirement decision of individuals. In general, we believe that it is important to construct variables that capture

since then.

the effect of items (i) to (vii) on individual's retirement decisions. Items (i), (ii) and (iii) are time related (e.g., *entry year* into the European Community); this means that we should introduce these variables so that they capture the impact they may have had over the length of the individual's working life. For example, Spanish individuals retiring in 1987 (one year after European Union entry) would have been affected by European Union membership differently than Spanish retiring in later years when the membership effects on the social security system would have been more established. Likewise, individuals that experience longer exposure to supplementary pension plans or to a national health system should react differently to these institutional variables than those with shorter or no exposure to them. In order to capture such 'exposure' to these time-based institutional variables we introduce these as the *distance* between the introduction of the institutional characteristic (e.g., entry into the European Union) and the individual specific 'officially assigned year of retirement'. For those with assigned retirement year before (or at exactly) the arrival of the institutional characteristic the variable 'distance between the establishment of the institutional characteristic and official retirement year' is zero. For example, a Spanish male born in 1920 is likely to be assigned 1985 as the official retirement year. This individual will have enjoyed zero years of European Union membership over his working life-cycle, whereas a Spanish male born in 1950 is likely to be assigned 2015 as his official retirement year and therefore can enjoy the impact of European Union membership for 29 years of his working life-cycle.²³ In total we construct three distance measures that refer to institutional characteristics, namely, 'distance between EU membership to official retirement year' to capture the definition of item (i), 'distance between year of introducing supplementary pension plans and official retirement year' for item (ii) and 'distance between year of introducing the public National Health System and official retirement year' for item (iii). Some countries have never implemented supplementary pension plans or public National Health Systems: in

²³An alternative is to measure the distance using the actual timing of the retirement outcome (e.g. if the Spanish born in 1950 due to retire in 2015 in fact retired in the year 2000, then the distance measuring the effect of EU membership would become 14 years). As we will see from Section 3.4 onwards, the same 'distance' trick is used to incorporate individual's socio-economic characteristics where distances between the actual characteristic and the retirement decision could vary endogenously in unobservable ways. To avoid measuring the distance endogenously we use the benchmark 'official retirement year' as opposed to actual timing of the retirement outcome (i.e., retirement year of 2004 if not retired), since 'official retirement year' is unique to the individual but cannot be influenced directly by the individual because it is determined by legislation. Since 'official retirement year' is used for the define socio-economic characteristics we prefer to be consistent throughout the paper thus using the same benchmark for all time-based institutional characteristics.

these cases their citizens are assigned a distance of zero years for the corresponding related distance variable, and we interpret these zeros as ‘zero years of the corresponding institutional effect’ over the individual’s working life-cycle. With reference to the variable that measures distance to entry into the European Union, we assume that all countries have membership with regards to labour markets and social security conditions.²⁴ Finally, the institutional items *(iv)*, *(v)*, *(vi)* and *(vii)* are not defined in terms of distance. In the case of item *(iv)* - combination of part-time work with old age pension - the details in Appendix A show that four countries have allowed for this characterisit since they established a system of social security. In the case of items *(v)* - type of monetary adjustment to pensions - and indicator equals one if the adjustment is based on changes in real wage rate, and zero if the consumer price index is the adjustment instrument. Items *(vi)* and *(vii)* - penalties and benefits - are based on ‘percentage reduction or increases’ and we assign these percentages to each individual according to country of survey. We allow the percentage ‘penalty’ to enter as a negative number, the reason being that for some countries the penalty is zero, thus, zero is a meaningful number that implies a zero penalty relative to countries where a penalty is imposed. Following the same reasoning, the percentage ‘benefit’ for late retirement is introduced as a positive percentage.²⁵ Finally, we draw from Section 3.1 to define an institutional variable that explains the ‘distance between official age or retirement and official age of *early* retirement’ for each individual in the sample, a distance that may capture incentives with regards to labour force participation in the late stages of an individual’s working life. For example, in the Netherlands the official retirement age is 65 with an statutory old age pension that reflects a 70% replacement rate. However, Dutch workers can also retire at the earlier age of 55 but at the expense of a lower replacement rate (of 53% for most workers). Introducing the distance between 55 and 65 (i.e., an average distance of 10) implies that workers who make a decision to retire early take into account the 10

²⁴ Appendix A shows the year of entry to the European Community for each of the 11 countries. Among these, Denmark does not belong to the monetary Euro-zone but belongs in full to the European Community for issues related to social security and labour market conditions. Switzerland, on the other hand, is not a member of the European Community, although in 2004 this country signed a treaty with the European Union that implies a ‘relatively’ free movement of labour between Swiss citizens and citizens in the 15-Eurozone as from January 2005. In order to capture possible EU effects we let 2005 as ‘year of entry’ for Swiss citizens.

²⁵ We have to emphasised that having revised all possible sources of retirement legislation, we have not found specific information on possible over-time changes for the actual magnitudes of the percentage penalty and percentage benefits. To the best of our knowledge, we assume that all the information relating to this percentages apply to each of the individuals in the sample.

year distance between the two official retirement ages. Table 4 shows all institutional variables just defined summarizing these according to the nature with which these are introduced in the analysis. Appendix A provides further details to complete the information required to construct all these institutional indicators.

Table 4: Summary of institutional characteristics (11 selected countries)

	Sample Size	Distance between year EU and ORAGE	Distance between year of SUPPLEMENT and ORAGE	Distance between year of NHS and ORAGE	Distance between ORAGE and OERAGE	Possibility to combine PT work Old Age PENSION	Adjustment based on REAL WAGES	Percentage penalty for each year of early retirement	Percentage benefit for each year of late retirement
Complete Sample	18,471 (1.00)	41.2	4.3	40.8	3.2	41%	41%	-3.48	3.63
Working Age (WS) ⁽¹⁾	9,706 (0.50)	49.2	5.8	46.0	3.1	45%	45%	-3.51	3.70
Retired (RS) ⁽¹⁾	12,127 (0.67)	36.5	3.3	37.8	3.3	39%	38%	-3.36	3.60
By Country									
Austria	1,399	4.7	No SP	32.8	3.1	Yes	Yes	-4.2%	+4.0%
Belgium	2,349	47.1	No SP	33.0	4.2	No	No	No penalty	No benefit
Denmark	1,319	26.0	12.8	No NHS	6.6	Yes	Yes	-10.0%	No benefit
France	2,087	41.9	No SP	47.6	3.7	No	No	-5.0%	+2.5%
Germany	1,658	47.0	No SP	65	0.5	Yes	Yes	-3.6%	+6.0%
Greece	1,792	15.8	No SP	13.8	5.8	No	Yes	-4.5%	+3.0%
Italy	1,621	46.2	No SP	21.9	4.0	No	No	No penalty	+3.5%
Netherlands	1,844	47.9	65	No NHS	10	No	Yes	-3.4	No benefit
Spain	1,212	11.4	No SP	30.5	4.0	Yes	No	-6.0%	+2.0%
Sweden	2,519	4.7	1.7	42.3	4.0	No	No	-6.0%	+8.4%
Switzerland	671	0.5	13.2	No NHS	2.0	No	Yes	-6.8%	+5.2%

Note: All figures are based on weighted values using individual calibrated sample weights. EU stands for European Union, SUPPLEMENT stands for 2nd tier supplementary pension plans, NHS stands for public National Health System, ORAGE stands for Official Retirement Age and OERAGE stands for Official Early Retirement Age. All distances refer to 'averages' in the corresponding sub-group defined in Column 1, whereas 'Yes' or 'No' describes if a particular institutional item is applicable or not in a given country.⁽¹⁾The overlap between the Working age and the Retired samples is 3,362 and these are all working age individuals who have retired early.

Besides the institutional variables explained in Table 4, the social, economic and demographic characteristics of individuals over their working lives are also important determinants at the time of making labour market decisions. In the following section we explain the use of the 2004 SHARE data to elicit socio-economic and demographic information over the working life of individuals.

3.4 General Socio-economic characteristics

The 2004 SHARE data is very informative on the social, economic and demographic characteristics of individuals with questions answered both at individual and at household level. We recall that our analysis aims at understanding the effect of the treatment 'onset of some health condition' (*hs_bora*) on the retirement outcomes described in Section 3.1. Therefore, It is important

to control for the characteristics that may affect the individual’s labour market decisions and/or the health status of individuals: conditioning on such characteristics should clear the effect of all observable confounders so that we can isolate the effect of health on retirement. However, it is crucial to our analysis that all observed characteristics refer to a point in time before the occurrence of the outcome of interest. Moreover, with reference to those that suffered a health condition, any observed characteristic correlated to health should be defined at a point in time before the onset of the first health condition.²⁶ Some important causal variables can be thought as happening previously (and therefore exogenously) to both the timing of the retirement outcomes and the onset of some health condition, for example, birth year (age), gender, years of education, skill class, the existence of siblings, etc. However, many other covariates also important in our analysis (e.g., financial variables, household composition, etc.) are time varying. We cannot take the contemporaneous value of these time varying variables because answers that refer to 2004 could be the result of labour market decisions (e.g., retirement) and therefore potentially endogenous to the outcome. The correct interpretation for these time varying variables is to know their value at a point *before* the timing of the retirement outcome. Once more we rely on self-reported retrospective information in the 2004 SHARE data to define the characteristics of individuals at a time previous to the determination of the retirement related outcomes. In particular we pick up indicators to construct (a) information for the financial standing of individuals before the timing of the retirement outcomes, (b) some behavioural habits of individuals before the timing of the retirement outcomes, (c) demographic characteristics prevailing in the household of individuals before the timing of the retirement outcomes and (d) the health status of other household members before the timing of the retirement outcomes. With reference to (a) we have the variables ‘first year of receiving a substantial inheritance’ and ‘year of household ownership’; these two variables assess the wealth standing of individuals before the timing of the retirement outcomes.²⁷ With reference to group (b) the survey provides information to construct

²⁶This simply guarantees the exogeneity of all the covariates with regards to the treatment (health) and the outcome (retirement decisions). Thus, our strategy implies that all covariates in some set X_t at time t precede the treatment HS and the outcome of interest as defined in Section 3.1, (i.e., *REBORA*, *REBERA*, *REAGE* or *TOTPEN*).

²⁷All other wealth variables in the survey are based on answers in 2004 (e.g., accumulated savings at the time of the survey) and as such we do not know in what way does the causal link run between these wealth variables and the individual’s labour market status. In order to avoid endogeneity resulting from reverse causality between accumulated financial wealth (savings, incomes, declared ownership of households *at the time* of the survey) we avoid the use of wealth variables that take as reference the day of the survey.

the variable ‘year when individuals stopped smoking if they ever smoked’. With reference to group (c) we have the variables ‘birth year of youngest dependent (daughter or son)’, ‘year of marriage’ and ‘year of lost-partnership (because of widowship or divorce)’. Finally, with reference to group (d) we use the answers to the health modules by other household members - even if these are not eligible respondents - to pick up ‘year of onset of some health condition suffered by other household members’.²⁸ A problem with *temporal variables* is that not all individuals are associated with a reference year for each of them (e.g., not all have received an inheritance, have bought a house, have had children, etc.). To avoid this problem we use the same ‘distance’ trick as that used when defining the institutional variables in Section 3.3, i.e., all *temporal variables* are used to construct ‘distance between the reference year and the assigned official age of retirement’ if and only if the ‘reference’ year happens such that we can establish the causal link between the corresponding temporal variable, the treatment and the retirement outcome. For example, knowing the year when an individual first received a substantial inheritance can be used to construct the distance between this event and the individual’s assigned official retirement year as long as the inheritance was received (and is therefore potentially causal) before retirement (or before 2004 if not retired by then). Moreover, in case that individuals are affected by health conditions, receiving an inheritance is only causal to retirement and not endogenous to the treatment ‘*hs_bora*’ if such inheritance happens before the onset of the health condition where, by definition, the health condition is accounted for as treatment if this has happened before the timing of the retirement outcome. Following the rubric just described we can construct a ‘distance’ variable that controls for the effect that ‘inheritance’ may have when this is received at a particular stage over the individual’s working life-cycle. Those who never received an inheritance are assigned a distance equal to zero: they have had zero years of enjoyment of an inheritance over their working life-cycles. Likewise, we consider ‘distance between household ownership and official retirement year’, ‘distance between marriage and official retirement year’, ‘distance between lost-partnership and official retirement year’, ‘distance between stop smoking

²⁸Picking up this information is subject to the possibility that all pertinent household members who were living with the individual at the time of retirement (or before 2004 if not retired) also respond to the questions in the survey at the time of the survey. In general, questions from ‘other household members’ are either from partners of the individuals or children of the individuals who are still living with the eligible respondent at the time of the survey. The survey does not explicitly ask the eligible respondent to provide information on the health of other household members.

and official retirement year'²⁹ and 'distance between the onset of the health conditions suffered by *others* in the household and *own* official retirement year': this latter distance variable is complemented with dummies that explain the type of condition suffered by the family member (chronic, acute or mental disorder). The variable 'birth year of dependents - daughter or son -' cannot be used to construct a 'distance' (e.g., 'distance between year of birth of the youngest child and official retirement year') because there is no reference year for those without children and in these cases a 'zero' would have the meaning that the child was born just before the official retirement year. As an alternative we construct a variable that counts the number of dependent children (i.e., ages below 18) that individuals face at the official retirement year. For example, let's assume that a male born in 1954 has 2019 as assigned official retirement year and that he has become a father in three occasions: in 1979, in 2001 and 2003. By 2019 this individual still has 2 dependent children below the age of 18, thus defining his value for the variable 'number of dependent children at the official retirement year'.³⁰ A final variable that takes into account time variability is the measured distance between the onset of a health event and the official retirement year: such distance should pick up the position in the individual's working life-cycle when the onset of the first health condition arrived, relative to his or her assigned official retirement year. This latter variable complements the treatment variable '*hs_bora*' as well as complementing all other temporal variables when 'health conditions' are accounted for. For those who have never suffered a health condition the 'distance between the onset of a health

²⁹Most individuals in the sample that ever started smoking did so at an early age (during their teenage years or early 20s). Therefore there is little heterogeneity with regards to starting point. Instead, the 'stopping' point is more heterogeneous. The distance between 'stop smoking' and official retirement age for those who never smoked is their full working life. Otherwise, for those who ever started smoking the distance is zero if they are still smoking and the time of the survey, or greater than zero (but smaller than their full working life) if they stopped at some point in the past.

³⁰An alternative would be to define the number of 'dependence years left at the year of official retirement' with reference the youngest child in the household. In the example that would be 2 years since the youngest child would be 16 in 2019, while those that have no children have zero years left of dependence at the time of their official retirement so that for them the zero is meaningful. We have tried this alternative 'years of dependency left' and the analytical outcome is invariant to that obtained when including 'number of dependents left'. At this point we make a reference with regards to both 'the death of offsprings' and 'disabled' offsprings. The 2004 SHARE data does not ask individuals if the recorded child is still alive at the time of the survey. Thus we need to assume that all child counted as 'dependents' are still alive in 2004. Finally, parents are asked to declare if their child is disabled and if the disabled child lives with them (same house or building). If individuals declare to have a handicapped child living with them we assume this child to be a dependent irrespective of his or her birth year.

problem and their official retirement year’ is clearly zero: they have experience zero years with health problems over their working lives.³¹ Table 5 shows the distribution of all these ‘*distance defined*’ variables by population sub-groups and according to treatment outcome (i.e., according to the binary outcome *hs_bora*).

Table 5: Summary statistics for distances of temporal variables

	Working Age Sample (WS-Sample, 9,706)			Retired Sample (RS-Sample, 12,127)		
	Percent ZERO value ⁽¹⁾	Average Distance if HS_BORA=0 ⁽²⁾	Average Distance if HS_BORA=1 ⁽²⁾	Percent ZERO value ⁽¹⁾	Average Distance if HS_BORA=0 ⁽²⁾	Average Distance if HS_BORA=1 ⁽²⁾
Distance: INHERITANCE	0.83	21.4	26.1	0.89	18.3	23.0
Distance: HOUSE-OWNER	0.30	28.8	31.3	0.40	25.1	28.7
Distance: MARRIAGE	0.51	36.1	37.3	0.62	36.0	36.3
Distance: LOST PARTNERSHIP	0.90	20.7	22.6	0.90	15.2	19.22
Distance: STOP SMOKING	0.40	26.3	30.0	0.33	21.4	25.0
Number of DEPENDENTS	0.98	1.2	1.2	0.97	1.3	1.2
Distance: Onset OWN HEALTH	0.57	--	21.3	0.57	--	18.8
Distance: Onset FAMILY HEALTH	0.92	11.5	12.7	0.94	8.1	8.4
Sample Size, <i>HS_BORA</i>	--	5,434	4,272	--	6,819	5,310

Note: All figures are based on weighted values using individual calibrated sample weights. Each of the distances (rows) refer to the distance between the item and the individual’s specific official retirement year (e.g., the distance in years between receiving the inheritance and the assigned official retirement year). The ‘number of dependent children’ refers to the number of under-age children in the household when individual reaches the official retirement year, including disabled children even if not minor. ⁽¹⁾ Percent of ‘ZERO’ value means percentage in the representative population without the item (e.g., 0.83 in the first row implies that among the working-age sample, 83% have never received an inheritance). ⁽²⁾ All distances take into account ‘positive’ valued items (e.g., the average number of years between inheriting and the official retirement year is 21.4 for those in the working-age sample that have received an inheritance before retirement and are not observed as treated by any of the health conditions).

Besides variables that are truly exogenous (e.g., gender and age) and the constructed ‘*distance*’ based temporal variables, there are other variables important for our analysis but for which we have only limited information to determine if these are or are not defined before the timing of the outcome and/or the (health) treatment. Generally speaking these variables refer to labour market characteristics in the last working place (for those who are retired) or current working place (for those who are still working in 2004), for example, size of firm of last (or present) working place. These are important variables with a potential causal effect on both the treatment ‘health’ and outcomes defined in Section 3.1, and therefore include these in our analysis.³²

³¹I am particularly grateful to Pedro Mira for pointing out the importance of allowing for the variable ‘distance between onset of health and official retirement year’ into the analysis.

³²For example, large firms offer their employees retirement conditions that are often not found in smaller firms. We only know the size of the firm in the last work place. When including this variable we are assuming that on average individual’s working lives have passed by under similar conditions as in their last recorded working job. This also applies to variables that explain ‘type of employee’ or ‘type of pension plans’ offered in last (or current

Table B1 in Appendix B classifies all variables described in Section 3 (outcomes, covariates or the treatment indicator) and Table B2 complements the analysis with summary statistics according to sub-sample (*WS-sample* or *RS-sample*) and health condition (i.e., treatment). In the following sections we use all the variables described so far to empirically analyse the conditional effect of the onset of an adverse health event (chronic, acute or mental disorder) on the retirement decision of European workers. The empirical strategy follows closely the formal causal framework proposed by Neyman (1923) and further expanded by Roy (1952) and Rubin (1974). This approach leads to nonparametric estimates of conditional causal effects that aim at eliminating both selection and specification bias as well as allow for complete heterogeneity effects of the onset of health conditions among individuals in the population.

4 Empirical strategy: Matching on the Propensity Score

The key issue in the field of treatment evaluation is that of identification of the effect under evaluation clean from selection and specification bias. In our case the treatment under evaluation is ‘the onset of a health condition’ and the key aim is to identify the impact of such treatment on the all the retirement outcomes defined in Section 3.1.

5 Conclusions

The paper aims at understanding the effect of the onset of some adverse health condition on the decision to retire, the age of retirement and the received annual amount of contributory pensions (in Euros). We consider health conditions of the three types (chronic, acute and mental conditions). In order to isolate the effect of health on retirement decisions we need to control for all confounding variation that may affect both the health outcome of individuals and the decision to retire. To this aim we take into account both institutional characteristics (individual and country specific) and socioeconomic characteristics that define the individual’s observed heterogeneity. Our estimates are based on the 2004 SHARE data (release 2), a cross-sectional survey that contains information from individuals age 50 and over from 11 European countries and Israel. We eliminate Israel from the data since for this particular country the

if still working) work place. To account for the potential effect of such an assumption the analysis introduces a set of binary indicators to explaining the age when individuals enter their last work place. See Appendix B.

information remains preliminary. Thus, our estimates are based on 11 European countries (Austria, Belgium, Denmark, France, Germany, Greece, Italy, The Netherlands, Spain, Sweden and Switzerland). Despite the fact that the 2004 SHARE data remains a cross-section at the present time, it contains sufficient retrospective information such that one is able to build the dynamic effects of health events on retirement decision without incurring in much lost data as result of lost retrospective information. Our estimates suggest that becoming subject to adverse health conditions, and specifically to those associated with chronic types, may induce a delay in the decision to retire. This is specially true in Belgium, Greece, The Netherlands and Spain. The results also suggest that compared to those who do not receive a health shock, those affected by adverse health conditions see their pensions significantly reduced in the case of Greek and Swiss citizens. The opposite is true if individuals are citizens of Denmark where average pensions are significantly higher for those who have suffered a health conditions (relative to those who have not). Our estimates are based on a selected sample such that endogeneity and selection bias problems should be minimized. Results and paper remain preliminary at the time of print.

References

- [1] Baker, M. J. Gruber and K. Milligan, (2003), The retirement incentive effects of Canada's Income security programs, *The Canadian Journal of Economics*, 36, 261-290
- [2] Clark, L.L, (2003), School and work: Technical and vocational education in France since the Third Republic, *Journal of Social History*, V 36(4), 1099-1102
- [3] Conde-Ruiz, I and V. Galosso, (2003), Early Retirement, *Review of Economic Dynamics*, 6, 12-36
- [4] Disney, R., C. Emmerson and M. Wakefield, Ill Health and retirement in Britain: A panel data-based analysis, *Journal of Health Economics*, 25, 621-649
- [5] Gustman, A. and T. Steinmeier, (2005), The social security early entitlement age in a structural model of retirement and wealth, *Journal of Public Economics*, 89, 441-463
- [6] Kingston, G.H., (1999), Efficient timing of retirement, *The Review of Economic Dynamics*, 3, 831-840
- [7] Linderboom, M. and M. Kerkhofs, (2002), Health and work of the elderly, IZA WP457, IZA-Bonn.

- [8] Manski, C.F., (1989), Anatomy of the selection problem, *Journal of Human Resources*, 24, 343-360
- [9] Meijer, E., A. Kapteyn and T. Andreyeva, (2007), Health indexes and retirement modeling in international comparisons, RAND Corporation Discussion Papers, RAND-USA.
- [10] Moller, A, M. Ejrnaes and L. Huster, (2005), Do single women value early retirement more than single men? *Labour Economics*, 12, 47-71
- [11] McGarry, K., (2004), Health and Retirement: Do changes in Health affect retirement expectations?, *The Journal of Human Resources*, 39, 624-648
- [12] Rubin, D., (1987), Multiple imputation for nonresponse in surveys, New York, John Wiley & Sons, Inc.
- [13] Stewart, J.M., (2001), The impact of health status on the duration of unemployment spells and the implication for studies of the impact of unemployment on health status, *Journal of Health Economics*, 20, 781-796
- [14] Schellhorn, M., A. Stuck, C.E. Minder and J. Beck, (2000) Health services utilization of elderly Swiss: Evidence from panel data
- [15] Spareto, L., (2005), Social Security incentives and retirement decisions in Italy: An empirical insight, *Research in Economics*, 59, 223-256
- [16] Wu, S., (2003), The effects of health on the Economic status of married couples, *Journal of Human Resources*, 38, 219-230

Appendix A Summary of laws and institutional characteristics for the 11 selected countries
in the 2004 SHARE data