



Università
Ca' Foscari
Venezia

Corso di Laurea Magistrale
in Economia e Finanza

"ordinamento ex D.M. 270/2004"

Tesi di Laurea

**An econometric analysis of the
transition to retirement in Europe:
results from SHARE data**

Relatore

Ch. Prof.ssa Irene Mammi

Laureando

Elena Donà

Matricola 864288

Anno Accademico

2017/2018

Abstract

Pensions and, in general, social security are one of the most significant expenditure items of the government budget, especially in Europe. The constant ageing process and the decreasing fertility rate in developed countries in the recent years urge policy makers to implement strategies to avoid an explosive rise of government spending. Policies, in turn, may influence individuals' decisions about consumption, saving but also their choice about retirement. The aim of this research is to find which factors are significant for the transition to retirement, and how they differ across European countries, namely in Western Europe. The study uses data from the 5th and 6th Wave of the SHARE survey. The work will develop an econometric analysis of the transition between labor force and retirement.

Contents

1	Introduction	7
2	Literature Review	9
2.1	Definition of retirement	9
2.2	Pension policies in Europe	10
2.3	Previous findings	13
3	The dataset	17
3.1	The SHARE Project	17
3.2	The baseline dataset	18
4	Descriptive analysis	21
4.1	The initial sample	21
4.2	Transition to retirement	35
4.3	The estimating sample	37
5	The model	45
5.1	Logistic regression	45
5.2	Multinomial logit regression	46
6	Results	49
7	Concluding remarks	62
A	Appendices	69
B		72

List of Tables

1	Country of origin: full sample	21
2	Geographical area: full sample	21
3	Age: full sample	22
4	Age by country	23
5	Level of education: full sample	24
6	Education by country	25
7	Gender: full sample	26
8	Employment: full sample	27
9	Employment by country	28
10	Relationship status: full sample	29
11	Number of children: full sample	29
12	Health condition: full sample	30
13	Monthly household income: full sample	30
14	Monthly household income by country	32
15	Household's economic position: full sample	33
16	Expectations about future reduction of pensions: full sample .	33
17	Expectation to work after age 63: full sample	33
18	Expectations about future raise in retirement age: full sample .	34
19	Retirement status	35
20	Transition to retirement	36
21	Transition to retirement by country	36
22	Summary statistics	37
23	Country of origin: estimating sample	38
24	Geographical area: estimating sample	38
25	Age: estimating sample	39
26	Level of education: estimating sample	40
27	Gender: estimating sample	40
28	Employment: estimating sample	41
29	Relationship status: estimating sample	41
30	Number of children: estimating sample	42
31	Health condition: estimating sample	42
32	Monthly household income: estimating sample	43
33	Summary statistics - MHI	44
34	Old age pension eligibility: estimating sample	44
35	Estimation logit model (country)	51
36	Estimation logit model (areas)	52

37	Indicators of goodness - Logit model (country)	53
38	Indicators of goodness - Logit model (area)	53
39	Estimation multinomial logit model (country)	55
40	Estimation multinomial logit model, rrr (country)	56
41	Estimation multinomial logit model, ame (country)	57
42	Estimation multinomial logit model (area)	58
43	Estimation multinomial logit model, rrr (area)	59
44	Estimation multinomial logit model, ame (area)	60
45	Indicators of goodness - Multinomial logit model (country) . .	61
46	Indicators of goodness - Multinomial logit model (area)	61
47	Expectations about future reduction of pensions by country . .	72
48	Expectation to work after age 63 by country	73
49	Expectations about future raise in retirement age by country .	74

1 Introduction

Pensions and, in general, social security are one of the most significant expenditure items of the government budget, especially in Europe. In 2014, public expenditure in pensions represented more than 13% of the GDP of the European Union [18]. The 2018 Ageing Report issued by the European Commission gives projections on the main demographic and macroeconomic aspects of the European population in the period 2016-2070 [16]. It states that the working-age population is expected to drop in that period with an associated decrease of labor supply and of the old-age dependency ratio; besides, it points out that this phenomenon has already been appreciated since 2011, when the baby-boom generation started to enter retirement. The constant ageing process and the decreasing fertility rate in developed countries in the recent years urge policy makers to implement strategies to avoid an explosive rise of government spending. Policies, in turn, may influence individuals' decisions about consumption, saving but also their choice about retirement. Several variables play a role in determining when an individual decides to receive his pension benefits. The aim of this research is to find which factors are significant and how they differ across European countries, namely in Western Europe. Countries will be classified in three main regions: Northern Europe, Continental Europe and Southern Europe. Two countries are taken into consideration for each region: Sweden and Denmark for Northern Europe; France and Germany for Continental Europe; Italy and Spain for Southern Europe. The purpose is also to check if living in a particular country or area of Europe is significant to their retirement decisions. In fact, southern and continental countries rely on public pension system more than countries in Northern Europe, even if their welfare state is more developed in terms of public spending. The study is conducted using data from the SHARE survey. The dataset provides cross-sectional data about the quality of life, economic aspects and personal information regarding people aged more than 50 years old. In particular, the focus is on the 5th and on the 6th Wave of the survey conducted respectively in 2013 and in 2015. The analysis is carried out by estimating econometric models, specifically a logistic regression and a multinomial logit model drawing from a research by Achdut, Sinai and Troitsky on older Israelis [1]. In their model, the dependent variable captures the transition between labor force and retirement. It is a discrete variable, which assumes different values depending on whether the individual moves from a working status to retirement (or vice versa) or if the status is unchanged. The

explicative variables are chosen among the questions provided by the SHARE survey. Following their research, we focus on transition to retirement by either considering a binary indicator or a multi-category dependent variable. The explanatory variables are decided among the topics of the survey and they consist of 9 variables, which are categorized in 4 main groups: demographics, health, household composition, and economic condition. This last, in particular, specifies also if a person has reached the retirement age in his country and so, if he is eligible of pension benefits.

The paper has the following structure: Section two is devoted to a literature review about previous findings on retirement choice (2.3), as well as an overview on European pension policies (2.2); Section three presents the SHARE project (3.1) and introduces the dataset (3.2); Section four reports the descriptive analysis of the sample (4); Section five explains the models used in the analysis (5); Section six presents the results (6); the last section is dedicated to the conclusions (7).

2 Literature Review

2.1 Definition of retirement

At the beginning of the 20th century, the concept of leaving the labor force and retire, once reached a certain age, started to pose questions among many scholars from different research fields [13]. Nevertheless, a decline in the labor force of men aged more than 65 has already been recorded since the end of the 19th century. There are many ways to define a "retired" individual, which depend also on whether a person is fully or partially retired. Many researchers have tried to define this latter condition using a quantitative approach, so that it was possible to take it into account in their analysis. Gustman and Steinmeier [24] describe the condition of partial retirement as a decrease of more than 40% in either workers' annual wages or working hours of their main full-time job, defined as the job held at age 55. A further study of Honig and Hanoch [25] questions Gustman and Steinmeier's method arguing that the cause of a reduction of wage or working hours may be different from a decision to partially retire, so the authors suggest a different approach. They consider the ratio between current and potential earnings to measure this decline. In particular, the potential earnings are the maximum earnings of an individual during the largest part of his working life. According to this ratio, they define an individual to be fully retired, if he earns no wages, partially retired if he receives 1%-50% of his maximum annual earnings, and not retired if the percentage is above 50%. Burtless and Moffitt [10] use a similar approach, which takes into account discontinuous drops in working hours together with Social Security receipts and self-reported status. In fact, regarding this last aspect, one of the issues that may arise is that the answer given by respondents about their retirement status may not match with their economic/working condition. According to Honig and Hanoch's research [25], 86.4% of individuals, who reported themselves as fully retired, are actually fully retired using the Earning-Ratio, and almost 40% of people who answered to be fully retired are partially retired. Ruhm [34] uses self-reported status to classify individuals, unless there is a contradiction between their statement and their earnings using income thresholds: if the respondent states to be retired but his annual earnings are greater than 500\$, he is defined as partially retired; if they are less than 500\$, he is classified as retired without considering his reported status; if his earnings are between 500\$ and 2000\$ and his self-reported status is not retired, he is considered as partially retired. As far as the SHARE survey is concerned, when an individual reports his current job situation as retired,

both full and partial retirement are included. Nevertheless, this does not cause substantial problems, as the analysis is more focused on why rather than how people decide to reduce (or leave) their labor force participation. However, it is important to keep in mind that a transition from a working condition to retirement does not necessarily mean a complete exit from the labor market.

2.2 Pension policies in Europe

The 2018 Ageing Report issued by Eurostat [16] provides one method to classify public pension systems, according to the source of funding and the type of risk that it covers. Pensions can be financed through contributions, which are earnings-related, or taxation. With regard to the specific risk covered, pensions are labeled as old-age and early pension, disability, survivor, minimum/basic or other. As part of the EU framework, entitlement for public pensions can follow three different schemes: defined-benefit (DB), notional defined contribution (NDC), and point systems (PS). A define-benefit scheme computes pension benefits as a proportion of earnings over a given number of years and it is currently used in Denmark, Spain and partly in France. A notional defined contribution scheme is an auxiliary mandatory pension plan based on contribution effective in Italy and Sweden. In a point system, used in Germany and partially in France, benefits are calculated on points given in proportion to earnings and years of contribution [15]. Most of the pensions provided by the member states are based on a pay-as-you-go (PAYG) system, in which current working generations' contributions finance the pension benefits of current retirees, and they provide a «(quasi-)minimum guaranteed or basic pension to those who do not qualify for the earnings-related scheme or have accrued only a small earnings-related pension» [16].

The economic literature provides several approaches to group pension provisions. A research by Grech [22] gives a categorization of European pension plans using three variables: the level of public spending, the replacement rate, and the condition of poverty of retirees. The countries are grouped by looking whether their result is above or below EU average. According to the analysis, Southern Europe countries are in the same group, characterized by a higher level of public expenditure (except Spain and Cyprus), a lower replacement rate and a higher poverty level. On the contrary, France and Germany have a higher replacement rate and a lower poverty rate, even though they share a higher level of public spending. A third group, in which Scandinavian countries are included, is defined to have a lower replacement rate, as well as a lower

level of poverty and state spending. Another approach is based on the entitlement of the benefits and on the purpose of the policy. A research by Bonoli [4] reports this classification, which was common in the French literature, between Bismarckian and Beveridgean welfare states. Bismarckian countries prefer to cover different groups with different social protection plans, which provide benefits based on individual earnings and contributions. The aim of this approach is to maintain a given level of income to workers after retirement. Countries that rely on this system come prevalently from Continental Europe. Differently, Beveridgean countries take a universal approach, where benefits are flat-rate and financed by general taxation, whose entitlement depends on residency and economic need. The purpose of this type of social policy is to prevent poverty and it is adopted in Northern European countries.

The European Welfare State concerning pensions has faced several reforms, especially since the 1990's. A research by Grech explains that, at the beginning of that decade, the prevalent model for retirement schemes was based on the PAYG system and an «earnings-related defined benefit determination structure» [23]. During the 90's, there was a wave of reforms devoted to change this paradigm due to the unsustainability of the previous system, which was too costly for the national accounts. For instance, Sweden and Italy switched from a defined benefit system to a defined contribution one [22]. This model allows the pension benefits to be calculated through the amount of contributions given during the working life and some individual/economic factors at the time of retirement; the aggregate is, then, converted into an annuity [23]. Also in France and Germany, there were some changes, respectively, in the number of contribution years and in the normal retirement age. A paper by Bonoli and Palier [5] shows the existence of a reform process that facilitates the evolution of the pension system, which encourages incremental reforms. The authors explain that two main elements play an important role in the process: a long period before the implementation of a reform and a step-by-step approach. The first one allows to accommodate the most politically influential cohorts and the second permits the process not to be too economically unsustainable for the population. Their analysis on the main pension reforms in France, Germany, and Italy in the 1990's and early 2000's reveal that their full implementation will occur on average about 15 years after their adoption for Germany, 25 years for France and 40 years for Italy. Bonoli and Palier also describe this particular reform process, detecting 4 specific stages: the first step implies an increase in the contribution rate, which happened around the

late '80s and in the 90's; the second step is the first wave of reforms approved in the 1990's, which are more intended to put concern about pensions into the public debate than to be very effective; the third step provides stricter policies in terms of contribution rate, retirement age and replacement rate; the last step opens the way to further cuts in benefits and encourages the use of private funds. Other papers suggest a tendency for different countries to reach similar social security schemes. An analysis by Overbye [32], which focuses on the pension system in Scandinavia and Australia, partly agrees with the so-called Convergence Theory, for which «roughly similar groups of citizens end up with roughly comparable pension security 'packages' across countries» thanks to the increasing concern and participation of the population in the retirement issue. In the beginning, all countries start with a minimum level of benefits based on income, while those countries are poor. As the economic condition improves, there is an increase in the demand for flat-rate benefits and for second-tier schemes to maintain their level of income. The convergence process leads to a dual pension system, where the working population is provided with different kinds of second-tier pension schemes and a limited group receives a minimum level of benefits based on income and tax-financed. Nevertheless, the author points out that there are still differences in the pension structure among countries, namely in the measure of the means-tested in minimum schemes and other organizational plans, which are explained by the role of politics in those countries.

The constant rise of public expenditure in pensions has brought concern to the public opinion about the future of retirement benefits provided by the Welfare State. Many countries, such as Sweden and some Eastern European countries, have shifted part of their public pension scheme into mandatory or quasi-mandatory private schemes [16]. In 1998, Germany opened the possibility to have a tax-free private pension fund, known as the *Riester Rente*, where employees could voluntarily contribute a part of their earnings [5]. France introduced two arrangements for voluntary savings in 2003: an individual pension plan and a scheme to be regulated within the firm. Italy enacted a legislation, which allowed the transfer of the saving system used to finance severance payments (TFR - *trattamento di fine rapporto*) into a pension fund. The financial crisis in 2008 has hindered retirement savings and severely reduced pensions, as private funds are based on individual accounts and benefits are computed only at the moment of retirement [28]. In fact, private savings for retirement face several risks, from financial risks due to the uncertain returns of the as-

sets to longevity risks, for which an individual lives longer than what he can afford with his pension means, and to behavioral risks, which are caused by the lack of knowledge on portfolio management [15]. After the financial crisis, some countries tried to reform their pension reserve funds to protect them from fluctuations of the financial market. In Sweden, whose pension system is a NDC, a reform in 2009 changed the adjustment system used to value the funds. Previously, the financial stability of the fund was computed with a balance ratio between the asset, composed of contributions and the market value of the system's reserve, and its liabilities. The market value of the system's reserve was substituted with a three-year average in order to reduce the impact of the financial instability on pension benefits [39]. In a report by OECD on pensions shows the growing size of funded pensions, which affected mainly Northern European countries [30]; in Denmark the share of private pension was almost 90% of its GDP in 2001, whereas it exceeded 200% in 2012, in Sweden the size increased from 30.4% in 2001 to 76% in 2014. In other European countries, the share of funded pensions is much lower than in Northern European countries, even though they also recorded a rise in the same period.

2.3 Previous findings

Numerous studies have been conducted to investigate the retirement decision of individuals both in a macro-funded and micro-funded framework. Laun and Wallenius' research [27] underlines the fact that the employment rate of older workers varies significantly across European countries, comparing it with the US level. While Sweden and the UK have a similar level as the United States, Germany's and France's employment rate is less than the two thirds of the American one. In some European countries, in fact, besides an accommodating social security system, many employers are reluctant to hire or keep older workers, as a research by Van Dalen et al. points out [43]. In [27], the authors apply a life-cycle model to make predictions about when people will ask for pension benefits according to the social security programs of each country. They find out that most of these countries have a generous pension scheme, which encourages early retirement. A previous work written by Börsch-Supan in the early '00s [6] focuses on the German public pension system, which he found out to largely favour retirement at the earliest legal retirement age. Staubli and Zweimüller's research [38] analyzes the dynamics of early retirement considering a reform brought into force in Austria in the period 2000-2003, which consists of a gradual increase in the early retirement age for both men and women. The results show a delay in the claims

of pension benefits, and thus a rise in the employment rate, especially among healthy and high-wage workers. Nevertheless, it has been proved not to be a one-to-one relationship and the reform has caused large spillover effects to unemployment programs. Other works study the effect of different policies on retirement choice. Fields and al. make some simulations using four policy scenarios to measure their impact on average retirement age in the United States [21]. These reforms include the raise of normal retirement age, the decrease of benefits for early retirement, the increase of late retirement benefits, and the delay of adjustments in the Consumer Price Index. The analysis has been carried out using a logit model, and the results show that cutting benefits for early ages while giving rewards for continued work gives the largest response in terms of retirement delay, even though the amount is quite small. After the implementation of this policy, workers are likely to postpone retirement only by three months and the reaction is even lower considering the other reforms.

At a more microeconomic level, there are some interesting findings about technological change and retirement. An analysis developed by Bartel and Sicherman [3] illustrates that people, who work in industries with frequent (expected) technological change, are more likely to have longer careers and retire later. An explanation for this result is that some of these industries have a higher depreciation rate of human capital, so they need more on-the-job training. At the same time, the positive effect of the introduction of new technologies on the profitability of training must be larger than the negative effect on training given by the depreciation rate.

Other works are more focused on the individual and try to identify the critical elements, which determine when a person decides to retire. Several scholars have, for instance, examined if the marital status can influence retirement choice and how this decision is affected. According to a research by Lalive and Perrotta on Swiss married couples in the period 1999-2000 [26], men are more than twice likely to retire than women after reaching their full retirement age. Nevertheless, women are 5% more likely to exit the labor market when their spouse is eligible for pension benefits than if he is not. The corresponding effect, instead, has been found not to be significantly different from zero. Moreover, both men and women with a low level of education have a stronger reaction to pension eligibility. Pozzebon and Mitchell's analysis looks more closely to women's retirement choice [33]. In particular, they analyze how their retirement behavior is affected by economic and family status.

The results demonstrate that unlike male workers, married female workers do not significantly evaluate their economic opportunities for retirement decisions. On the contrary, they value leisure highly, especially if their spouse is much older than them, suggesting a complementarity of leisure time when age difference is large; however, delayed retirement is preferred when husband's income is high, probably due to, as the authors say, «a marital selection in tastes toward work». Later retirement is also favored, if their husband is in poor health conditions for their need to cover increasing health insurance costs.

Part of the literature concerning retirement patterns has treated SHARE data. Wahrendorf, Morten et al. [44] have carried out a study on older European workers using SHARE data since the 1st Wave until 2012 to detect the conditions, for which some individuals retire later than others. From the analysis it appears that 30% of them is self-employed or they hold a high-wage occupation. People, who keep working, have also a better health condition compared to retired. Some interesting results are found likewise in a research on older Israelis by Achdut, Tur-Sinai et al. [1], who apply a multinomial logit model to measure the probability to move from working life to retirement and the factors that facilitate this transition. They find that there is a positive relationship between age and probability to leave the labor market from one Wave to another, which increases every year by almost 1 percentage points with respect to the probability to stay in the labor force. Moreover, the data show that education has a negative effect on the retirement process, even though the impact is quite small, as well as gender (considered as a dummy variable, in which male = 1). On the contrary, there is a positive correlation between the number of children and the probability to retire, and so is having a spouse in the labor force, and being wealthy.

The analysis in this paper will cover this last aspect, focusing on some of the individual characteristics provided by the SHARE survey. The model refers to the econometric representation applied by Achdut, Tur-Sinai et al. [1], which will be used in a European context.

3 The dataset

3.1 The SHARE Project

As population ageing increasingly becomes part of the agenda of European policy makers, there is also demand for scientific evidence to study the evolution of this process. In 2002 the SHARE project was developed to understand ageing and its impact on individuals coming from different cultural backgrounds in Europe [9]. The project emerged as the first European Research Infrastructure Consortium with a long-term perspective, which was set to 2024 [8]. It provides cross-country comparable microdata, which are collected at the individual and household level through a survey. The respondents are 50 or older, or there must be at least one eligible person in the household. The survey has been conducted every two years since 2004 and it covers several attributes of the life of an individual including, besides demographic variables, health, economic condition, and social support [7]. The dataset has been developed and harmonized following the characteristics of the Health and Retirement Study (HRS) in the USA and the English Longitudinal Study of Ageing (ELSA) in the UK. The first wave was run in 2004 and it involved more than 20000 people across 11 European countries from different geographical areas [9]. In 2008-09 a special release, called SHARELIFE, was carried out as a retrospective study to review the history of respondents' life [8]. The latest wave (6th) was conducted in 2015 and 18 countries participated in the survey, whereas a new wave (7th) will take place in 28 countries, with a full coverage of the EU, and it is expected to be released in spring 2019 [35].

The data collection of the SHARE survey has mainly been funded by the European Commission and resources have been allocated to each country such that they all would be able to interview 1500 households. This phase follows a procedure so that the sampling design represents a probability sample in each country and, thus, the dataset is representative of the population. A probability sample is «a sample drawn from a universe [...] such that every elementary unit of the universe has a nonzero probability of being selected» [9]. Another important aspect is that the samplings of each country are completely independent of each other. To reach this goal, researchers make also use of administrative data, at least for basic demographic variables, as the information given by the respondents might be incomplete or inaccurate [8]. During the third wave of the survey, SHARE has collaborated with the German Pension Fund, which has provided them with longitudinal data on socio-demographic aspects and work history of Germans, and the list of collaborations has been

expanded during the following waves. For other countries, they had to cooperate with the national statistical offices to have access to registers, such as in France and Spain.

Some weaknesses of the SHARE survey is that sometimes the response rate is rather low and the attrition rate (defined as the amount of sample lost throughout the waves) is moderate, even though the response rate is higher than in other European surveys and some US studies [8]. To cope with this issues, SHARE has adopted a system of ex-post calibrated weights, so as to reduce the potential selectivity bias. Another feature, which causes some concerns in the analysis of the SHARE survey, is the complexity of the dataset. The SHARE team tries to manage this problem by broad data cleaning, supplying generated variables and a wide documentation for users such as release guides and the complete questionnaire in all languages used. SHARE provides, in fact, a standardization of some variables, such as education, in order to simplify the interpretation of the answers. A further assistance is given by a set of technical variables, which help identify the type of respondent (family, household, or financial), state if the interview is a single or couple interview, or the relationship status of the individual. SHARE also updates its dataset with new countries and respondents over the waves, providing a large variety of information in terms of individual and household characteristics, and their context, in which they live.

3.2 The baseline dataset

The analysis is run using the data from the 5th and the 6th wave of the survey, carried out respectively in 2013 and in 2015. Every answer is recorded by its identification number, which identifies the respondent and remains the same if the same person is interviewed in more than one wave. SHARE provides many variables that are grouped in different modules according to their topic. As a consequence, the different subsets have to be merged to perform the analysis. The sample is composed of individuals, who are aged 50 or more, and live in one of these 6 European countries belonging to 3 different European areas: Sweden and Denmark for Northern Europe, Germany and France for Central Europe, Italy and Spain for Southern Europe. The countries are grouped in these 3 areas, based on the hypothesis that each area has peculiar characteristics in terms of quality of life, economic condition, and pension policy. The variables are chosen from the questions of the survey and they are classified in:

- Demographic variables: beside the **country** of origin, they include the

year of birth, gender, and level of education;

- Economic variables: they include **current job situation, monthly household income, general economic condition** expressed as ease to make ends meet, and **old age pension eligibility;**
- Health variables: it is composed of one variable, which explains the **general health condition** of the individual;
- Household composition variables: they specify how the household is designed for each individual. They include **relationship status** and **number of children;**
- Expectation variables: they are three variables, which reveal the expectation about their future retirement period. Respondents are asked to say a number from 0 to 100 about how likely they think these 3 situations can happen: **future raise of retirement age, future reduction in pensions, work after age 63.**

In case of answers as "Don't know" or if the respondent explicitly refuses to answer, the values are dropped from the sample. However, missing values are reported for each variable, since a high non-response rate has been recorded for some questions. For further details about the variables chosen, see Appendix A.

4 Descriptive analysis

4.1 The initial sample

The dataset includes 50,038 individuals, of which 18,802 individuals are part of the longitudinal sample, 7,650 have been interviewed only in the 5th Wave and 4,784 only in the 6th Wave. These specific countries have been preferred because they have the highest GDP in the European Union with respect to their belonging geographical area [17]. The macro-areas are structured as follows: Denmark and Sweden for Northern Europe, Germany and France for Central Europe, Italy and Spain for Southern Europe. Countries have been chosen in pair not to have overrepresented macro-areas. Therefore, no other countries have been added, as Denmark and Sweden were the only Nordic countries, which participated in the SHARE survey in the 5th Wave. The distribution of the sample according to the country of origin is displayed in Table 1. There is no migration from one country to another throughout the waves.

Table 1: Country of origin: full sample

Country	Freq.	Percent
Germany	9,012	18.01
Sweden	7,758	15.50
Spain	9,719	19.42
Italy	8,907	17.80
France	7,398	14.78
Denmark	7,244	14.48
Total	50,038	100.00

Spain and Germany are the most represented samples with respectively 9,719 and 9,012 individuals. If we aggregate countries according to the geographical area, the sample is divided as shown in Table 2.

Table 2: Geographical area: full sample

Geographical area	Freq.	Percent
Northern Europe	15,002	29.98
Central Europe	16,410	32.80
Southern Europe	18,626	37.22
Total	50,038	100.00

The subsample from Northern Europe is the least represented with less than 30% of the overall sample, whereas more than 35% of the sample comes from Southern Europe.

As for the age, Table 3 shows the overall sample grouped by year of birth, divided in 3 age classes.

Table 3: Age: full sample

Age class	Freq.	Percent	Cum.
1909 - 1934	7,263	14.51	14.51
1935 - 1954	30,525	61.00	75.52
1955 - 1965	12,250	24.48	100.00
Total	50,038	100.00	

The first age class is composed of individuals, who are more than 79 or 81 at the time of the interview¹. The second age class includes people, whose age is in the interval 80/78-61/59. The third class groups individuals, who are less than 61/58². More than a half of the sample belongs to the middle age class, while less than one quarter is younger than 60 years old. In Table 4, it is shown the composition of age classes by each country.

¹The age depends on whether they answered in the 5th or in the 6th Wave.

²For individuals of the 6th Wave, the age class stops at 1963.

Table 4: Age by country

(a) Age Denmark				(b) Age Sweden			
Age class	Freq.	Percent	Cum.	Age class	Freq.	Percent	Cum.
1909 - 1934	791	10.92	10.92	1909 - 1934	1,158	14.93	15.33
1935 - 1954	4,027	55.59	66.51	1935 - 1954	5,507	70.98	85.91
1955 - 1965	2,426	33.49	100.00	1955 - 1965	1,093	14.09	100.00
Total	7,244	100.00		Total	7,758	100.00	
(c) Age Germany				(d) Age France			
Age class	Freq.	Percent	Cum.	Age class	Freq.	Percent	Cum.
1909 - 1934	890	9.88	9.88	1909 - 1934	1,243	16.80	16.80
1935 - 1954	5,390	59.81	69.68	1935 - 1954	4,359	58.92	75.72
1955 - 1965	2,732	30.32	100.00	1955 - 1965	1,796	24.28	100.00
Total	9,012	100.00		Total	7,398	100.00	
(e) Age Italy				(f) Age Spain			
Age class	Freq.	Percent	Cum.	Age class	Freq.	Percent	Cum.
1909 - 1934	1,131	12.70	12.70	1909 - 1934	2,050	21.09	21.09
1935 - 1954	5,563	62.46	75.15	1935 - 1954	5,679	58.43	79.52
1955 - 1965	2,213	24.85	100.00	1955 - 1965	1,990	20.48	100.00
Total	8,907	100.00		Total	9,719	100.00	

It is possible to notice that the highest number of very aged people both in absolute value and relative to the country subsample is in Spain, followed by France with respectively 21.09% and 16.80%. Nevertheless, the highest proportion of individuals over 60 within the country is in Sweden (85.91%). Denmark has, instead, the highest percentage of younger individuals within the subsample (33.49%), whereas Germany has the most younger people in absolute value (2,732).

Table 5 reports information about the level of education of the overall sample.

Table 5: Level of education: full sample

Level of education	Freq.	Percent	Cum.
None	3,811	7.62	7.62
Primary education	10,794	21.57	29.19
Lower secondary	7,457	14.90	44.09
Upper secondary	14,459	28.90	72.99
Post-secondary	1,418	2.83	75.82
1st stage of tertiary	10,772	21.53	97.35
2nd stage of tertiary	479	0.96	98.31
Still in school	8	0.02	98.32
Other	168	0.34	98.66
No response	672	1.34	100.00
Total	50,038	100.00	

The largest proportion of the sample has an upper secondary degree or high school degree (28.90%) and more than 20% of the sample have either a primary education or a university degree as highest level of education. It is important to point out that more than three-quarters of the sample do not have a tertiary education (75.82%). In Table 6, it is shown the level of education for each country. If labels are missing, it means that the frequency is 0. The widest gap among countries is the level of tertiary education, which is the highest in Denmark (42.21%) and the lowest in Italy (7.06%). A significant difference can be noticed also in the number of people, who did not get any education, in Spain (2,269) corresponding to the 23.23% of the country subsample. If we compare it to the value of the overall sample, it appears that almost the 60% of people with no education comes from Spain. A possible explanation is that it may be due to the high number of very aged people in the Spanish sample. In both countries of Northern Europe, the tertiary education is the most represented level, whereas the upper secondary education is the most spread in Central Europe, and so is the primary education in Southern Europe.

Table 6: Education by country

(a) Level of education - Denmark

Level of education	Freq.	Percent	Cum.
None	9	0.12	0.12
Primary education	659	9.10	9.22
Lower secondary	657	9.07	18.29
(Upper) secondary	2,779	38.36	56.65
Post secondary	.	.	.
1st stage of tertiary	3,058	42.21	98.87
2nd stage of tertiary	23	0.32	99.19
Other	10	0.14	99.32
No response	49	0.68	100.00
Total	7,244	100.00	

(c) Level of education - Germany

Level of education	Freq.	Percent	Cum.
None	56	0.62	0.62
Primary education	72	0.80	1.42
Lower secondary	996	11.05	12.47
(Upper) secondary	4,755	52.76	65.24
Post-secondary	342	3.79	69.03
1st stage of tertiary	2,614	29.01	98.04
2nd stage of tertiary	90	1.00	99.03
Still in school	2	0.02	99.06
Other	4	0.04	99.10
No response	81	0.90	100.00
Total	9,012	100.00	

(e) Level of education - Italy

Level of education	Freq.	Percent	Cum.
None	716	8.04	8.04
Primary education	3,306	37.12	45.16
Lower secondary	2,202	24.72	69.88
(Upper) secondary	1,649	18.51	88.39
Post-secondary	254	2.85	91.24
1st stage of tertiary	629	7.06	98.30
2nd stage of tertiary	65	0.73	99.03
Still in school	1	0.01	99.05
Other	7	0.08	99.12
No response	78	0.88	100.00
Total	8,907	100.00	

(b) Level of education - Sweden

Level of education	Freq.	Percent	Cum.
None	56	0.72	0.72
Primary education	1,623	20.92	21.64
Lower secondary	1,052	13.56	35.20
(Upper) secondary	1,844	23.77	58.97
Post-secondary	684	8.82	67.79
1st stage of tertiary	2,288	29.49	97.28
2nd stage of tertiary	46	0.59	97.87
Other	59	0.76	98.63
No response	106	1.37	100.00
Total	7,758	100.00	

(d) Level of education - France

Level of education	Freq.	Percent	Cum.
None	705	9.53	9.53
Primary education	1,699	22.97	32.50
Lower secondary	603	8.15	40.65
(Upper) secondary	2,639	35.67	76.32
Post-secondary	3	0.04	77.36
1st stage of tertiary	1,333	18.02	94.38
2nd stage of tertiary	236	3.19	97.57
Still in school	1	0.01	97.58
Other	59	0.80	98.38
No response	120	1.62	100.00
Total	7,398	100.00	

(f) Level of education - Spain

Level of education	Freq.	Percent	Cum.
None	2,269	23.35	23.35
Primary education	3,435	35.34	58.69
Lower secondary	1,947	20.03	78.72
(Upper) secondary	793	8.16	86.88
Post-secondary	135	1.39	88.27
1st stage of tertiary	850	8.75	97.02
2nd stage of tertiary	19	0.20	97.21
Still in school	4	0.04	97.25
Other	29	0.30	97.55
No response	238	2.45	100.00
Total	9,719	100.00	

One of reasons to explain this large discrepancy may be the difference in public spending on tertiary education and so the cost to pursue high-level studies for

households. A research by Castles has found, in fact, a positive correlation between public expenditures in education and the number of enrolled student in tertiary education [12]. While primary and secondary education have become almost universal, especially in Europe, after the two world wars and it does not vary much among countries, the level of spending on higher education, therefore, determines the policy on public spending [11]. A report issued by OECD on education shows that in 2014, Denmark and Sweden spent close to 4% or more of its total government spending in tertiary education, whereas Germany spent 3%, France and Spain 2.2%, and Italy 1.6% [31]. Nordic countries are marked not only by a high public expenditure, but also high subsidies to students in terms of subsidised loans [2]. OECD reports that only 5% of source of funds for tertiary education comes from private sources in Denmark, whereas it is 14% in Germany and 32% in Spain [31]. Another reason may be the different economic situation, which is reflected in the diverse demand for skilled labor, especially in less recent times [11].

Table 7 shows the distribution of gender in the sample. Females are slightly more represented than males. The proportions are similar at the country level, as well, and so they are not reported.

Table 7: Gender: full sample

	Freq.	Percent	Cum.
Male	23,504	46.97	46.97
Female	26,534	53.03	100.00
Total	50,038	100.00	

In Table 8, it is reported the job condition of the overall sample.

Table 8: Employment: full sample

Current job situation	Freq.	Percent.
Retired	27,440	54.84
Employed or self-employed	13,722	27.42
Unemployed	1,497	2.99
Permanently sick or disabled	1,507	3.01
Homemaker	4,453	8.90
Other	698	1.39
No response	721	1.44
Total	50,038	100.00

More than half of the sample is composed of retirees (54.84%), followed by employed people (27.42%).

In Table 9, they are displayed the distributions of every job condition for all countries considered. In all subsamples, the retirement condition exceeds or is close to 50% of the respondents. Sweden has the highest share of retirees in its subsample (65.36%) and the most retired individuals of all subsamples (5,071), followed by France, whose subsample is 64.79% composed of retirees. The highest number of people, who are employed, are in Denmark both in absolute value and in percentage (3,015 and 41.62%, respectively), even though the most represented condition is "retired" (48.83%). A significant part of individuals in Southern Europe are homemaker, especially with respect to the other countries (20.53% in Spain and 18.29% in Italy, compared to 5.34% in Germany and 0.47% in Denmark).

Table 9: Employment by country

(a) Employment Denmark			(b) Employment Sweden		
Current job situation	Freq.	Percent	Current job situation	Freq.	Percent
Retired	3,537	48.83	Retired	5,071	65.36
Employed or self-employed	3,015	41.62	Employed or self-employed	2,290	29.52
Unemployed	170	2.35	Unemployed	87	1.12
Permanently sick or disabled	286	3.95	Permanently sick or disabled	139	1.79
Homemaker	34	0.47	Homemaker	9	0.12
Other	92	1.27	Other	45	0.58
No response	110	1.52	No response	117	1.51
Total	7,244	100.00	Total	7,758	100.00
(c) Employment Germany			(d) Employment France		
Current job situation	Freq.	Percent	Current job situation	Freq.	Percent
Retired	4,742	52.62	Retired	4,793	64.79
Employed or self-employed	3,001	33.30	Employed or self-employed	1,725	23.32
Unemployed	304	3.37	Unemployed	201	2.72
Permanently sick or disabled	282	3.13	Permanently sick or disabled	191	2.58
Homemaker	481	5.34	Homemaker	305	4.12
Other	83	0.92	Other	57	0.77
No response	119	1.32	No response	126	1.70
Total	9,012	100.00	Total	7,398	100.00
(e) Employment Italy			(f) Employment Spain		
Current job situation	Freq.	Percent	Current job situation	Freq.	Percent
Retired	4,700	52.77	Retired	4,597	47.30
Employed or self-employed	1,855	20.83	Employed or self-employed	1,836	18.89
Unemployed	237	2.66	Unemployed	498	5.12
Permanently sick or disabled	198	2.22	Permanently sick or disabled	411	4.23
Homemaker	1,629	18.29	Homemaker	1,995	20.53
Other	190	2.13	Other	231	2.38
No response	98	1.10	No response	151	1.55
Total	8,907	100.00	Total	9,719	100.00

Table 10 shows that the large majority of the sample is in a couple. The percentage is similar looking at the country subsamples.

Table 10: Relationship status: full sample

In a couple	Freq.	Percent
No	12,154	24.29
Yes	37,884	75.71
Total	50,038	100.00

Table 11 shows the number of children that the respondents reported. Almost 40% of the sample has two children and the large majority does not have more than 3 children.

Table 11: Number of children: full sample

Number of children	Freq.	Percent	Cum.
0	4,804	9.60	9.60
1	8,632	17.25	26.85
2	19,951	39.87	66.72
3	10,143	20.27	86.99
4	3,737	7.47	94.46
5	1,395	2.79	97.25
6	611	1.22	98.47
7	266	0.53	99.00
8	97	0.19	99.19
9	46	0.09	99.28
>10	43	0.09	99.37
No response	313	0.63	100.00
Total	50,038	100.00	

In Table 12, it is reported the general health condition of the overall sample. More than 60% of individuals state that they are in good health, while almost 10% of the sample declares to live in a poor health condition.

Table 12: Health condition: full sample

Health condition	Freq.	Percent	Cum.
Excellent	4,776	9.54	9.54
Very good	9,401	18.79	28.33
Good	17,816	35.60	63.94
Fair	12,957	25.89	89.83
Poor	4,983	9.96	99.79
No response	105	0.21	100.00
Total	50,038	100.00	

Table 13 reports the household income per month of the sample divided in classes (expressed in Euros).

Table 13: Monthly household income: full sample

Income class	Freq.	Percent	Cum.
<1000	5,048	10.09	10.09
1000-2000	13,834	27.65	37.74
2000-3000	10,552	21.09	58.82
3000-4000	5,659	11.31	70.13
4000-5000	3,601	7.20	77.33
>5000	6,952	13.89	91.22
No response	4,392	8.78	100.00
Total	50,038	100.00	

Almost 60% of people live in a household with less than 3000€ per month. The largest share of households belongs to the class 1000-2000 Euros (27.65%). There is also a moderate non-response rate with almost 9% of the sample. In Table 14, the results from country subsamples are reported. Denmark has the highest share of people in the top income class (28.64%) both within the subsample and with respect to the other subsamples, while Spain has the highest share of people in the lowest class compared to other countries (23.13%). The most represented classes are 2000-3000€ for Sweden and Germany, 1000-2000€ for France, Italy, and Spain. The data from the subsamples reflect the statistics from Eurostat on median household net income [19]. Denmark and Sweden have a monthly median income above 2000€, whereas France and Germany share a monthly median income close to 1700€, Spain and Italy register a monthly median income of less than 1500€. Eurostat also sets a threshold

to state if individuals are at risk of poverty, which is defined as 60% of the «national median equivalised disposable income (after social transfers)» [20]. According to this classification, Spain and Italy have a higher share of people, who are at risk of poverty, despite a lower median income. In addition, there is a significant part of the Spanish subsample that did not answer (18.11%).

Table 14: Monthly household income by country

(a) Monthly household income - Denmark

Income class	Freq.	Percent	Cum.
<1000	88	1.21	1.21
1000-2000	1,096	15.13	16.34
2000-3000	1,614	22.28	38.63
3000-4000	956	13.20	51.82
4000-5000	1,111	15.34	67.16
>5000	2,075	28.64	95.80
No response	304	4.20	100.00
Total	7,244	100.00	

(b) Monthly household income - Sweden

Income class	Freq.	Percent	Cum.
<1000	181	2.33	2.33
1000-2000	1,477	19.04	21.37
2000-3000	2,152	27.74	49.11
3000-4000	1,308	16.86	65.97
4000-5000	1,020	13.15	79.12
>5000	1,216	15.67	94.79
No response	404	5.21	100.00
Total	7,758	100.00	

(c) Monthly household income - Germany

Income class	Freq.	Percent	Cum.
<1000	589	6.54	6.54
1000-2000	2,199	24.40	30.94
2000-3000	2,309	25.62	56.56
3000-4000	1,295	14.37	70.93
4000-5000	671	7.45	78.37
>5000	1,344	14.91	93.29
No response	605	6.71	100.00
Total	9,012	100.00	

(d) Monthly household income - France

Income class	Freq.	Percent	Cum.
<1000	723	9.77	9.77
1000-2000	2,007	27.13	36.90
2000-3000	1,659	22.42	59.33
3000-4000	993	13.42	72.75
4000-5000	468	6.33	79.08
>5000	984	13.30	92.38
No response	564	7.62	100.00
Total	7,398	100.00	

(e) Monthly household income - Italy

Income class	Freq.	Percent	Cum.
<1000	1,219	13.69	13.69
1000-2000	3,661	41.10	54.79
2000-3000	1,668	18.73	73.52
3000-4000	771	8.66	82.17
4000-5000	215	2.41	84.59
>5000	618	6.94	91.52
No response	755	8.48	100.00
Total	8,907	100.00	

(f) Monthly household income - Spain

Income class	Freq.	Percent	Cum.
<1000	2,248	23.13	23.13
1000-2000	3,394	34.92	58.05
2000-3000	1,150	11.83	69.88
3000-4000	336	3.46	73.34
4000-5000	116	1.19	74.53
>5000	715	7.36	81.89
No response	1,760	18.11	100.00
Total	9,719	100.00	

The following tables include the last two groups of variables (General economic condition in Table 15 and Expectations in Table 16, Table 17 and Table 18) that will not be part of the model due to the high non-response rate, but they

are reported as they may be useful for the interpretation of the results. It is important to remind that questions about expectations are asked, only if the individual answers "Employed" as job condition.

Table 15: Household's economic position: full sample

Is household able to make ends meet	Freq.	Percent	Cum.
With great difficulty	2,797	5.59	5.59
With some difficulty	6,586	13.16	18.75
Fairly easily	8,287	16.56	35.31
Easily	13,122	26.22	61.54
No response	19,246	38.46	100.00
Total	50,038	100.00	

18.75% of the sample does not have sufficient economic resources to live decently, which becomes 30.47%, if we do not consider people who gave no response. The participation rate is rather low (61.54% of respondents).

Table 16: Expectations about future reduction of pensions: full sample

Expectation	Freq.	Percent	Cum.
Don't know	210	2.50	2.50
0-24	1,960	23.31	25.81
25-49	512	6.09	31.89
50-74	2,406	28.61	60.51
>75	3,321	39.49	100.00
Total	8,409/13,722	100.00	

Table 17: Expectation to work after age 63: full sample

Expectation	Freq.	Percent	Cum.
Don't know	104	1.55	1.55
0-24	1,922	28.56	30.11
25-49	393	5.84	35.95
50-74	1,311	19.48	55.43
>75	2,999	44.57	100.00
Total	6,729/13,722	100.00	

Table 18: Expectations about future raise in retirement age: full sample

Expectation	Freq.	Percent	Cum.
Don't know	190	2.26	2.26
0-24	2,412	28.69	30.95
25-49	497	5.91	36.87
50-74	1,976	23.51	60.37
>75	3,331	39.63	100.00
Total	9,647/13,722	100.00	

For expectation variables, the answer "Don't know" is kept, as it is plausible that individuals are not certain about the future. Looking at the tables, respondents do not seem so optimistic about their future retirement. The highest share of answer belongs to the interval 75-100 for all three questions, i.e. people highly expect changes in pension policy, and so they feel high uncertainty about the future. In Table 47, Table 48, Table 49 in Appendix B, it is possible to appreciate the differences among countries. Northern countries and Germany tend to be more confident about the future and expect fewer changes in pension policy, in particular concerning a possible raise in the retirement age. Instead, Southern countries strongly believe that their retirement period will be negatively affected by future policies. Nevertheless, it is useful to notice that the response rate is relatively much lower in Northern countries than in the others. A publication of the European Commission describes the evolution of pensions with respect to changes in retirement age and pension benefits. On the one hand, between 2008 and 2013 large increases in the legal retirement age were put into force in France and Sweden (even though Sweden was the only country, which did not further increase the retirement age in 2015) [14]. On the other hand, Italy and Denmark projected the largest change for the period 2008-2060. As for the results in Table 17, a high share of people, who reported a number above 75, may be due to the fact that in most of these countries, the retirement age is already set above 63, and the incentives to early retirement have been closed to new entrants or strictly limited.

The choice to analyze European countries in three areas comes from the fact that they are characterized by specific attributes, which make them differ from each other. The results of the descriptive analysis, in fact, show that Nordic countries have a larger number of individuals, who pursued high-level studies,

on average have a higher income and expect fewer changes in pension policy. Besides, countries in Northern Europe are marked by a welfare state that is very present both with respect to public expenditure and because it has a universal approach in terms of coverage, as discussed in Section 2. Southern countries, by contrast, have on average a lower education level, lower income and are more pessimistic about their future retirement. In addition, their welfare state is notable for a lower public spending and a policy coverage, which depends on contributions. Countries from Central Europe are in the middle, with a high public expenditure and a welfare state based on contribution records. The analysis aims to verify if these peculiarities are significant for retirement decisions, mainly for people who are near the statutory retirement age.

4.2 Transition to retirement

This section provides a descriptive analysis of the sample, focusing on the transition to retirement, i.e. when individuals move from a working status to retirement in the next period. For simplicity, a person is considered retired if he answers "Retired" as job condition in the survey, otherwise he is considered as not retired, without making any other assumption.

Table 19 shows the retirement status of the sample, divided in "Retired" and "Not retired". Most of the sample is defined as retired (54.84%).

Table 19: Retirement status

Retired	Freq.	Percent
No	21,877	43.72
Yes	27,440	54.84
No response	721	1.44
Total	50,038	100.00

In the next tables, it is shown the results of the transition probabilities to retirement from Wave 5 to Wave 6 for the whole sample (Table 20) and by country (Table 21).

Table 20: Transition to retirement

Retired	No	Yes	Total
No	83.45	16.55	100.00
Yes	4.41	95.59	100.00
Total	40.19	59.81	100.00

Most of individuals do not change their status in the next period; 83.45% of people, who are not retired, maintains their condition after two years, and the share is even higher, when retirees are considered (95.59%). However, there is a significant part of the sample, which leaves a non-retired status to become retired and it involves more people than its opposite movement (16.55% against 4.41%). The total results show that it is more likely that people are retired in the next period than non-retired (59.81% and 40.19%, respectively).

Table 21: Transition to retirement by country

(a) Denmark				(b) Sweden			
Retired	No	Yes	Total	Retired	No	Yes	Total
No	85.29	14.71	100.00	No	81.53	18.47	100.00
Yes	2.24	97.76	100.00	Yes	2.56	97.44	100.00
Total	46.22	53.78	100.00	Total	30.80	69.20	100.00
(c) Germany				(d) France			
Retired	No	Yes	Total	Retired	No	Yes	Total
No	84.01	15.99	100.00	No	75.23	24.77	100.00
Yes	3.80	96.20	100.00	Yes	1.84	98.16	100.00
Total	42.73	57.27	100.00	Total	26.70	73.30	100.00
(e) Italy				(f) Spain			
Retired	No	Yes	Total	Retired	No	Yes	Total
No	88.24	11.76	100.00	No	82.91	17.09	100.00
Yes	6.24	93.76	100.00	Yes	10.18	89.82	100.00
Total	42.21	57.79	100.00	Total	49.14	50.86	100.00

Italy has the highest percentage of people, who maintains a non-retired status in the next period (88.24%), although it is worth recalling that Italy has a relatively high share of homemakers, which are included in the non-retired category. The largest share of people, who becomes retired in the second period, is in France with almost one quarter of individuals, who were in a non-retired status in the previous period (24.77%). France has also the highest frequency of being retired in the following period, with any previous status (73.30%). There is also a significant part of the Spanish subsample of people, who were retired in period one and return to a non-retired condition (10.18%).

Table 22: Summary statistics

Variable	Obs	Resp. Rate	Mean	Std. Dev.	Min	Max
Year of birth	50038	100%	1946	10	1910	1965
Education	49366	98,66%	Upper secondary	N/A	N/A	N/A
Gender	50038	100%	N/A	N/A	N/A	N/A
Relationship status	50038	100%	N/A	N/A	N/A	N/A
Number of children	49725	99.37%	2.174	1.345	0	17
Physical health	49933	99,79%	Good	N/A	Excellent	Poor
Job condition	49317	98,56%	N/A	N/A	N/A	N/A
Household income	45646	91,22%	6.65e+08	8.11e+10	0	1.00e+13
General economic condition	30792	61,53%	Fairly easily	N/A	With great difficulty	Easily
Exp. pension reduction	8426	61,40%	55,48	35,07	0	100
Exp. work after 63	6736	49,09%	55,52	38,86	0	100
Exp. raise retirement age	8426	61,40%	52,99	37,48	0	100

4.3 The estimating sample

The sample that is used to estimate the model is composed only of the individuals, who have been interviewed in both waves, and so they are part of the longitudinal study. The main reason for this choice is that the analysis is focused on the transition from one Wave to the following one, and so we need people, who responded at least twice. As for the explicative variables, only the values given in the 5th Wave have been kept. This is a reasonable choice, as these characteristics show a limited variability over time. After dropping missing observations for "Job condition", as it is necessary for the formulation of the dependent variable, the estimating subsample consists of 18,505 individuals. For each variable used in the model, they are reported also the statistics by geographical area, instead of by country as previously shown, since it is also interesting to analyze differences across macro-areas.

Table 23: Country of origin: estimating sample

Country	Freq.	Percent
Germany	3,583	19.36
Sweden	3,042	16.44
Spain	3,364	18.18
Italy	3,025	16.35
France	2,551	13.79
Denmark	2,940	15.89
Total	18,505	100.00

In Table 23, it is shown the frequency of each country in the sample. Germany has the most populated sample followed by Spain (19.36% and 18.18%, respectively). The distributions are quite similar to the values in the starting sample. Table 24 reports the frequency of each geographical area. In this case, the frequencies becomes more homogeneous with all areas close to one third of the sample.

Table 24: Geographical area: estimating sample

Geographical area	Freq.	Percent
Northern Europe	5,982	32,32%
Central Europe	6,134	33,15%
Southern Europe	6,389	34,53%
Total	18,505	100.00

Table 25 shows the distribution of age of the estimating sample and for the three areas.

Table 25: Age: estimating sample

(a) Estimating sample				(b) Northern Europe			
Age class	Freq.	Percent	Cum.	Age class	Freq.	Percent	Cum.
1909-1934	2,422	13.09	13.09	1909-1934	681	11.38	11.38
1935-1954	11,837	63.97	77.05	1935-1954	3,941	65.88	77.27
1955-1965	4,246	22.95	100.00	1955-1965	1,360	22.73	100.00
Total	18,505	100.00		Total	5,982	100.00	
(c) Central Europe				(d) Southern Europe			
Age class	Freq.	Percent	Cum.	Age class	Freq.	Percent	Cum.
1909-1934	684	11.15	11.15	1909-1934	1,057	16.54	16.54
1935-1954	3,842	62.63	73.79	1935-1954	4,054	63.45	80.00
1955-1965	1,608	26.21	100.00	1955-1965	1,278	20.00	100.00
Total	6,134	100.00		Total	6,389	100.00	

All subsamples are more concentrated in the middle age class compared to the initial sample. Central Europe has the most people in the youngest age class (1,608 individuals), whereas Southern Europe in the oldest age class (1,057).

Table 26: Level of education: estimating sample

(a) Estimating sample				(b) Northern Europe			
Level of education	Freq.	Percent	Cum.	Level of education	Freq.	Percent	Cum.
None	1,271	6.96	6.96	None	19	0.32	0.32
Primary education	3,892	21.32	28.28	Primary education	838	14.15	14.47
Lower secondary	2,619	14.34	42.62	Lower secondary	665	11.23	25.71
Upper secondary	5,460	29.90	72.52	Upper secondary	1,881	31.77	57.47
Post-secondary	529	2.90	75.42	Post-secondary	268	4.53	62.00
1st stage of tertiary	4,258	23.32	98.74	1st stage of tertiary	2,203	37.21	99.21
2nd stage of tertiary	173	0.95	99.69	2nd stage of tertiary	26	0.44	99.65
Still in school	1	0.01	99.69	Still in school	.	.	.
Other	56	0.31	100.00	Other	21	0.35	100.00
Total	18,259	100.00		Total	5,921	100.00	
(c) Central Europe				(d) Southern Europe			
Level of education	Freq.	Percent	Cum.	Level of education	Freq.	Percent	Cum.
None	236	3.89	3.89	None	1,016	16.19	16.19
Primary education	591	9.75	13.64	Primary education	2,463	39.26	55.45
Lower secondary	583	9.61	23.25	Lower secondary	1,371	21.85	77.30
Upper secondary	2,794	46.08	69.33	Upper secondary	785	12.51	89.82
Post-secondary	136	2.24	71.57	Post-secondary	125	1.99	91.81
1st stage of tertiary	1,579	26.04	97.61	1st stage of tertiary	476	7.59	99.39
2nd stage of tertiary	123	2.03	99.64	2nd stage of tertiary	24	0.38	99.78
Still in school	.	.	.	Still in school	1	0.02	99.79
Other	22	0.36	100.00	Other	13	0.21	100.00
Total	6,064	100.00		Total	6,274	100.00	

The distribution of the level of education of the estimating sample is rather similar to the distribution of the initial sample. The most frequent education level is the upper secondary with almost 30% of the estimating sample. The geographical areas maintain their difference in the partition of the degree of education.

Table 27 displays the frequency of gender. Also in the estimating sample, there is a slightly higher share of females than males.

Table 27: Gender: estimating sample

Gender	Freq.	Percent
Male	8,807	47.59
Female	9,698	52.41
Total	18,505	100.00

In Table 28, it is reported the job condition for the estimating sample and the three geographical areas. Northern Europe is the area, where there are the most employed people (33.17% of its subsample). The highest share of retired within its subsample are in Central Europe, while there is a significant percentage of people, who are homemakers, in Southern Europe (19.74%). If we look more in detail, it appears that almost all homemakers in this area are women.

Table 28: Employment: estimating sample

(a) Estimating sample			(b) Northern Europe		
Current job situation	Freq.	Percent	Current job situation	Freq.	Percent
Retired	11,067	59.81	Retired	3,686	61.62
Employed	4,634	25.04	Employed	1,984	33.17
Unemployed	445	2.40	Unemployed	88	1.47
Permanently sick or disabled	511	2.76	Permanently sick or disabled	138	2.31
Homemaker	1,54	8.32	Homemaker	15	0.25
Other	308	1.66	Other	71	1.19
Total	18,505	100.00	Total	5,982	100.00

(c) Central Europe			(d) Southern Europe		
Current job situation	Freq.	Percent	Current job situation	Freq.	Percent
Retired	3,922	63.94	Retired	3,459	54.14
Employed	1,577	25.71	Employed	1,073	16.79
Unemployed	157	2.56	Unemployed	200	3.13
Permanently sick or disabled	171	2.79	Permanently sick or disabled	202	3.16
Homemaker	264	4.30	Homemaker	1,261	19.74
Other	43	0.70	Other	194	3.04
Total	6,134	100.00	Total	6,389	100.00

Table 29 reports the relationship status of the sample. Almost 76% of the estimating sample is in a couple likewise the share in the initial sample.

Table 29: Relationship status: estimating sample

In a couple	Freq.	Percent
No	4,443	24.01
Yes	14,062	75.99
Total	18,505	100.00

Table 30 reports the frequencies of children per individual. Also in the estimating sample, having two children is the most represented category.

Table 30: Number of children: estimating sample

Number of children	Freq.	Percent	Cum.
0	1,665	8.99	8.99
1	3,15	17.02	26.01
2	7,451	40.26	66.27
3	3,837	20.73	87.00
4	1,375	7.43	94.43
5	513	2.77	97.20
6	213	1.15	98.35
7	103	0.55	98.90
8	32	0.17	99.07
9	17	0.09	99.16
>10	13	0.07	99.23
No response	136	0.77	100.00
Total	18,505	100.00	

Table 31 represents the distribution of health condition for the estimating sample. The values are quite similar to the initial sample. More than half of the sample declares to have a good health status.

Table 31: Health condition: estimating sample

Health condition	Freq.	Percent	Cum.
Excellent	1,676	9.06	9.06
Very good	3,616	19.54	28.60
Good	6,718	36.30	64.90
Fair	4,874	26.34	91.24
Poor	1,621	8.76	100.00
Total	18,505	100.00	

Table 32 reports the income classes for the estimating sample and for each area.

Table 32: Monthly household income: estimating sample

(a) Estimating sample				(b) Northern Europe			
Income class	Freq.	Percent	Cum.	Income class	Freq.	Percent	Cum.
<1000	1,880	10.90	10.90	<1000	116	2.00	2.00
1000-2000	5,269	30.55	41.45	1000-2000	1,069	18.48	20.48
2000-3000	4,243	24.60	66.05	2000-3000	1,61	27.83	48.31
3000-4000	2,201	12.76	78.81	3000-4000	951	16.44	64.74
4000-5000	1,403	8.13	86.94	4000-5000	813	14.05	78.79
>5000	2,252	13.06	100.00	>5000	1,227	21.21	100.00
Total	17,248	100.00		Total	5,786	100.00	

(c) Central Europe				(d) Southern Europe			
Income class	Freq.	Percent	Cum.	Income class	Freq.	Percent	Cum.
<1000	495	8.52	8.52	<1000	1,269	22.45	22.45
1000-2000	1,621	27.90	36.43	1000-2000	2,579	45.62	68.07
2000-3000	1,585	27.29	63.71	2000-3000	1,048	18.54	86.61
3000-4000	898	15.46	79.17	3000-4000	352	6.23	92.84
4000-5000	465	8.00	87.18	4000-5000	125	2.21	95.05
>5000	745	12.82	100.00	>5000	280	4.95	100.00
Total	5,809	100.00		Total	5,653	100.00	

The distribution of the estimating sample follows a similar pattern as the initial sample. In this case, the class 1000-2000€ becomes slightly more populated in percentage (30.55%). In Southern Europe almost 70% of the subsample has a household income lower than 2000€, while in the other regions it is less than 40%. Table 33 shows the summary statistics of monthly household income. It appears that there is a very high variance in the distribution of income, and looking at the percentiles we can say that there is a small number of respondents who stated to receive a very high income.

Table 33: Summary statistics - MHI

Percentiles	Value	
1%	400	Obs 17248
5%	705	
10%	926	Mean 4636.742
25%	1,400	Std. Dev. 49297.22
		Variance 2.43e+09
50%	2.200	Skewness 107.9691
		Kurtosis 13004.32
75%	3.500	
90%	5,559.7	
95%	9,000	
99%	50,000	

We define a new variable named **Old age pension eligibility**, which will be used in the model. It states if the individual is eligible of old age pension benefits, according to the retirement age of his country in 2015, i.e. during the 6th Wave. The retirement age for each country is provided by the 2015 OECD issue about pensions [29], and it is set to 65 in Denmark, Sweden, Germany and Spain, 63 in France and Italy (women), 66 in Italy (men). Table 34 shows how many individuals are eligible for old age pensions. Almost 65% of the sample has reached the retirement age for their country.

Table 34: Old age pension eligibility: estimating sample

Eligible	Freq.	Percent
No	6,600	35.67
Yes	11,905	64.33
Total	18,505	100.00

5 The model

Previous studies on retirement choice have used logistic regressions to carry out their estimations. In the paper by Fields [21], it is used an ordered logit model to measure the effect of different policies on retirement delay. Wahrendorf, Morten et al. have used a series of multinomial logistic regressions on SHARE data to forecast the likelihood of working for individuals [44]. They have also provided results in terms of odd ratios and average marginal effects, so that coefficients are more comparable across the models. Achdut, Aviad et al. have run a multinomial logit model to analyze the transition to retirement on older Israeli, where the dependent variable assumes 4 different values and the explicative variables are chosen among the questions provided by the SHARE survey. In this research the analysis is, firstly, conducted with a logistic regression using a binary dependent variable. Afterwards, the estimation is made using a multinomial logit regression. All the estimations are run using Stata 13.

5.1 Logistic regression

The first model applied to the estimating sample is a logistic regression. The dependent variable is a binary variable that captures the transition from a non-retirement status and retirement. It assumes value equal to 1 if this transition occurs, 0 if the condition remains unchanged or if he returns to a non-retirement state after being retired in the previous period. A logistic model estimates the variation of probability of an outcome to occur and use a logit transformation to link the explanatory variables and the dependent one [40]. The probability is given by:

$$\Pr(y = 1|X) = p_i = \frac{\exp(\beta_j X)}{1 + \exp(\beta_j X)}$$

and the model is described as:

$$\text{logit}(p_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_N x_N$$

where X is the vector of regressors, $\text{logit}(p) = \ln\left(\frac{p_i}{1-p_i}\right)$ and β_N are the coefficients estimated with the maximum likelihood method.

The explicative variables are transformed into dummies as follows:

- Gender assumes value 1 if the individuals is male, 0 if female;
- Relationship status assumes value 1 if a person is in a couple, 0 if he is single;

- Health condition assumes value 1 if the respondent answers "Good", "Very good", or "Excellent";
- Education assumes value 1, if the individuals has a high degree of education, i.e. "Post secondary", "1st or 2nd stage of tertiary" education, 0 otherwise;
- Pension eligibility assumes value 1 if the person is eligible for an old age pension, 0 if he is not;

As for "Age", "Monthly household income", and "Number of children", they are kept as non-binary variables. Household income has been logarithmic transformed to minimize the effect of very large values. In the analysis, countries are first taken individually to detect any country effect and then, they will be aggregated into macro-areas. The results will be presented also in terms of odds ratio and average marginal effects. The former indicate the ratio between the probability of an event to occur and the probability that it does not, and in the model it is described as:

$$\frac{p_i}{1 - p_i} = \exp(\beta_j X)$$

The latter give the average change in probability when the explanatory variables change by one unit, *ceteris paribus*.

5.2 Multinomial logit regression

The second model is a multinomial logit model (MNL) and it is based on the application made in the research by Achdut, Aviad et al [1]. The dependent variable has the same object as in the logit model, but now it can assume four different values: $j = 0$, if an individual is in a non-retired status in Wave 5 and maintains his condition in Wave 6; $j = 1$, if a person moves from a non-retired status to retirement; $j = 2$, if he moves towards retirement after being in a non-retired status in the previous period; $j = 3$, if he is retired both in Wave 5 and Wave 6. The probability that an individuals is in one of the four outcomes depends on a set of N explicative variables x and it is given by the conditional probability:

$$\Pr_{ij}(y = j|x_i) = \frac{\exp(x_i \beta_j)}{1 + \sum_{k=1}^j (\beta_k x_i)}$$

The base outcome for the analysis is $j = 0$. The multinomial logistic model is, then, given by:

$$\log \left[\frac{\pi_j(x_i)}{\pi_0(x_i)} \right] = \beta_0 + \beta_{1j} x_{1i} + \beta_{2j} x_{2i} + \dots + \beta_{Nj} x_{Ni}$$

where $\pi_j(x_i) = \Pr_{ij}(y = j|x_i)$. The explicative variables are maintained as in the logit model. Coefficients of the estimation are also expressed as relative risk ratio, i.e. «how the risk of an outcome falling in the comparison group compared to the risk of the outcome falling in the referent group changes with the variable in question» [41].

6 Results

The first results are derived from the logit model, in which we take into account the countries individually (Table 35). Values are written as coefficients, odds ratios, and average marginal effects. The p-values for OR and AME are not reported, as they are the same as for the coefficients. Looking at the signs of coefficients, there is positive correlation between being eligible for pension benefits and move to a retired status, as well as the number of children. The result is consistent with previous findings from Achdut et al [1], who have also found a positive relationship between these two variables and a transition to retirement. Their values of AME show that passing the legal retirement age increases the probability to retire by 74.5 percentage points, and the number of children increments the probability to move to retirement by 0.43 percentage points. In this representation, it appears that there is a positive relationship between year of birth and probability to retire, which is counterintuitive, since it would mean that it is more likely to retire for younger people than for older ones. Being in good health, instead, has a negative impact on retirement by decreasing the probability by 12.57 percentage points. Individuals with a higher household income also have less probability to retire by 0.45 percentage points. These results are in line with the literature, which states that people, who have a higher income, have a high-wage occupation, and so they have more incentive not to leave their job [44]. Nevertheless, in this analysis the control variable for job condition does not specify the type of work of individuals. If we consider the countries of origin, values are to be interpreted with respect to Germany. According to the results, people from Denmark, France and Spain are more likely to retire than Germany. On the contrary, individuals from Sweden and Italy are less likely to retire than German people. In particular, coefficients for Italy and Spain are highly significant in this model. The lower probability to retire for the former may be caused by the fact that the legal retirement age for men is the highest among the chosen countries. However, it is worth recalling that being in a non-retired condition does not necessarily imply an employment status as discussed in Section 4.2.

Table 36 reports the values of the logit model with aggregated countries. The results are similar to the values of the first model and both the signs of the coefficients and the signs of AME do not change. If we look at the geographical macro-areas, Northern Europe and Southern Europe both show a negative sign. This means that it is less likely that people from that areas retire with respect to Central Europe, even though the values are not statistically signif-

icant. Tables 37 and 38 show the indicators of goodness of their respective model.

Table 35: Estimation logit model (country)

VARIABLES	Coefficients	Odds ratio	AME
<i>Demographic variables</i>			
Year of birth	0.0776*** (0.00585)	1.0807	0.0051
Gender (male=1)	0.00818 (0.0608)	1.0082	0.0005
High education	0.0433 (0.0709)	1.0443	0.0029
<i>Household composition</i>			
Relationship status	-0.0925 (0.0736)	0.9116	-0.0061
Number of children	0.0649*** (0.0224)	1.0671	0.0043
<i>Health status</i>			
Good health	-0.190*** (0.0657)	0.8269	-0.1257
<i>Economic variables</i>			
Household income	-0.0686** (0.0296)	0.9337	-0.0045
Eligibility	1.126*** (0.100)	3.0839	0.7449
<i>Country of origin</i>			
Denmark	0.0574 (0.0987)	1.0591	0.0038
Sweden	-0.0531 (0.103)	0.9483	-0.0035
France	0.102 (0.101)	1.1074	0.0067
Italy	-0.400*** (0.112)	0.6705	-0.2644
Spain	0.259*** (0.0988)	1.2955	0.0171
Constant	-153.8*** (11.42)	1.59e-67	

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Table 36: Estimation logit model (areas)

VARIABLES	Coefficients	Odds ratio	AME
<i>Demographic variables</i>			
Year of birth	0.0764*** (0.00583)	1.0794	0.0051
Gender (male=1)	0.0163 (0.0607)	1.0164	0.0011
High education	0.0426 (0.0706)	1.0435	0.0028
<i>Household composition</i>			
Relationship status	-0.0992 (0.0733)	0.9056	-0.0066
Number of children	0.0776*** (0.0223)	1.0807	0.0051
<i>Health status</i>			
Good health	-0.189*** (0.0654)	0.8278	-0.0125
<i>Economic variables</i>			
Household income	-0.0794*** (0.0293)	0.9237	-0.0053
Eligibility	1.096*** (0.0995)	2.9929	0.0727
<i>Geographical area</i>			
Northern Europe	-0.0397 (0.0734)	0.9610	-0.0026
Southern Europe	-0.0867 (0.0758)	0.9170	-0.0057
Constant	-151.3*** (11.39)	1.91e-66	

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 37: Indicators of goodness - Logit model (country)

Indicators of goodness	
Log likelihood	-4,289.7895
Number of observations	16,994
LR chi2(13)	248.75
Prob > chi2	0.0000
Pseudo R2	0.0282
Correctly classified	92,76%

Table 38: Indicators of goodness - Logit model (area)

Indicators of goodness	
Log likelihood	-4,308.785
Number of observations	16,994
LR chi2(13)	210.76
Prob > chi2	0.0000
Pseudo R2	0.0239
Correctly classified	92,76%

Hereinafter, they are presented the results of the multinomial logit model. Tables 39, 40 and 41 show the values respectively of the coefficients, relative risk ratios, and average marginal effects. We remind that the base outcome is remaining in the labor force in both periods, whereas the first outcome defines the transition to retirement. In this case, there is a negative correlation between year of birth and probability to retire. In fact, the higher the year of birth, the younger the individuals is, and so it is reasonable that he is less likely to retire. Also being a male increases the probability to retire with respect to the base outcome, and this reflects previous findings in the paper by Lalive and Perrotta [26], where men are more than twice likely to retire, once they reach their eligibility age. A good health condition and a high income still have a negative impact on the probability to retire as in the logit model; the former reduces the probability to retire by 0.12 percentage points, the latter makes it decrease by 0.34 percentage points. A high degree of education, instead, has no significant impact in any model; this may be due to the fact that people with a higher education represent quite a small part of the sample, as shown in Table 26. In this representation, France becomes

significant for a retirement transition, together with Italy and Spain. In the descriptive analysis about transition to retirement, France is the country with the highest share of people, who moves to retirement, as shown in Table 21. As for the results in the second outcome, i.e. individuals who are retired in the first period and move to a non-retired status in the second period, the fact of being in a couple have a positive impact on the return to the labor force. However, it is hard to draw conclusion on this outcome, since it is a quite rare event, as displayed in Tables 20 and 21. As for the third outcome, in which individuals are retired in both periods, household income is not significant. This is a plausible result, as it is possible that income does not influence retirement decisions, once individuals are retired.

At a macro-area level (Table 42), the model gives similar results as the previous estimation. There is a high significance of year of birth, gender and pension eligibility. Differently from the logistic regression, geographical areas are statistically significant and they both reduce the probability to retire with respect to living in Central Europe. If we look at AME, being in Northern Europe and Southern Europe respectively reduce the probability by 0.36 and 0.84 percentage points. Yet, the presence of the same negative sign may be due to different reasons. From the results of the descriptive analysis (Section 4), countries from Northern Europe may have incentive not to retire because individuals, who are working, belong to the top income class. Instead, countries from Southern Europe register a high share of people, who are homemakers, a category that is considered as a non-retired condition in this model. So, it is possible that they have less incentive to change their status because they may have other income resources (namely partner's income). However, control variables are not so specified to state a precise explanation.

Table 39: Estimation multinomial logit model (country)

VARIABLES	Coefficients			
	0 (base outcome)	1	2	3
<i>Demographic variables</i>				
Year of birth		-0.0258*** (0.00747)	-0.146*** (0.00968)	-0.149*** (0.00573)
Gender (male=1)		0.699*** (0.0703)	0.0597 (0.122)	1.286*** (0.0585)
High education		0.0198 (0.0795)	-0.0208 (0.151)	-0.0696 (0.0657)
<i>Household composition</i>				
Relationship status		0.149* (0.0832)	0.417*** (0.132)	0.356*** (0.0657)
Number of children		-0.0461* (0.0255)	-0.174*** (0.0413)	-0.175*** (0.0204)
<i>Health status</i>				
Good health		-0.112 (0.0747)	-0.250** (0.117)	0.144** (0.0596)
<i>Economic variables</i>				
Household income		-0.0685** (0.0328)	-0.0562 (0.0550)	0.00229 (0.0267)
Eligibility		2.098*** (0.112)	0.858*** (0.187)	3.290*** (0.0914)
<i>Country of origin</i>				
Denmark		-0.0765 (0.110)	-0.646*** (0.232)	-0.195** (0.0931)
Sweden		-0.0942 (0.117)	-0.187 (0.204)	-0.0263 (0.0941)
France		0.576*** (0.117)	0.0819 (0.236)	0.818*** (0.101)
Italy		-0.918*** (0.124)	-0.106 (0.179)	-0.818*** (0.0916)
Spain		-0.757*** (0.114)	-0.428** (0.177)	-1.793*** (0.0944)
Constant		48.39*** (14.59)	282.8*** (18.87)	287.3*** (11.18)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 40: Estimation multinomial logit model, rrr (country)

VARIABLES	Relative risk ratio			
	0 (base outcome)	1	2	3
<i>Demographic variables</i>				
Year of birth		0.9745***	0.8639***	0.8620***
Gender (male=1)		2.0108***	1.0615	3.6188***
High education		1.0200	0.9794	0.9327
<i>Household composition</i>				
Relationship status		1.1603*	1.5167***	1.4280***
Number of children		0.9549*	0.8402***	0.8391***
<i>Health status</i>				
Good health		0.8941	0.7789**	1.1548**
<i>Economic variables</i>				
Household income		0.9337**	0.9453	1.0023
Eligibility		8.1485***	2.3583***	26.8495***
<i>Country of origin</i>				
Denmark		0.9264	0.5242***	0.8229**
Sweden		0.9101	0.8297	0.9740
France		1.7785***	1.0854	2.2649***
Italy		0.3994***	0.8996	0.4412***
Spain		0.4689***	0.6520**	0.1665***
Constant		1.04e+21***	6.6e+122***	5.7+e124***

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 41: Estimation multinomial logit model, ame (country)

VARIABLES	Average marginal effects			
	0 (base outcome)	1	2	3
<i>Demographic variables</i>				
Year of birth		0.00380***	-0.0012***	-0.0136***
Gender		-0.000305	-0.0169***	0.1157***
High education		0.00382	0.0004	-0.0082
<i>Household composition</i>				
Relationship status		-0.00356	0.0040	0.0290***
Number of children		0.00346**	-0.00141	-0.0155***
<i>Health status</i>				
Good health		-0.0122***	-0.0075***	0.0235***
<i>Economic variables</i>				
Household income		-0.00446**	-0.0012	0.0034
Eligibility		0.0517***	-0.0372***	0.4476***
<i>Country of origin</i>				
Denmark		0.00296	-0.0080**	-0.0116
Sweden		-0.00488	-0.0029	0.0027
France		0.00896	-0.0072**	0.0607***
Italy		-0.0255***	0.0115***	-0.0606***
Spain		0.0155**	0.0202***	-0.1832***

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 42: Estimation multinomial logit model (area)

VARIABLES	Coefficients			
	0 (base outcome)	1	2	3
<i>Demographic variables</i>				
Year of birth		-0.0239*** (0.00738)	-0.144*** (0.00951)	-0.140*** (0.00560)
Gender (male=1)		0.697*** (0.0702)	0.0429 (0.121)	1.246*** (0.0578)
High education		-0.00178 (0.0791)	-0.0239 (0.150)	-0.105 (0.0650)
<i>Household composition</i>				
Relationship status		0.123 (0.0829)	0.417*** (0.132)	0.313*** (0.0649)
Number of children		-0.0355 (0.0252)	-0.175*** (0.0409)	-0.181*** (0.0200)
<i>Health status</i>				
Good health		-0.0839 (0.0743)	-0.236** (0.116)	0.173*** (0.0589)
<i>Economic variables</i>				
Household income		-0.0700** (0.0325)	-0.0531 (0.0545)	0.0247 (0.0264)
Eligibility		2.126*** (0.112)	0.881*** (0.186)	3.386*** (0.0911)
<i>Geographical area</i>				
Northern Europe		-0.304*** (0.0836)	-0.428*** (0.163)	-0.441*** (0.0693)
Southern Europe		-1.024*** (0.0887)	-0.283** (0.141)	-1.604*** (0.0712)
Constant		44.87*** (14.42)	278.5*** (18.55)	270.4*** (10.92)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 43: Estimation multinomial logit model, rrr (area)

VARIABLES	Relative risk ratio			
	0 (base outcome)	1	2	3
<i>Demographic variables</i>				
Year of birth		0.9764***	0.8658***	0.8695***
Gender (male=1)		2.0080***	1.0438	3.4747***
High education		0.9982	0.9764	0.9000
<i>Household composition</i>				
Relationship status		1.1308	1.5179***	1.3671***
Number of children		0.9651	0.8391***	0.8345***
<i>Health status</i>				
Good health		0.9195	0.7896**	1.1891***
<i>Economic variables</i>				
Household income		0.9324**	0.9482	1.0250
Eligibility		8.3811***	2.4139***	29.5529***
<i>Geographical area</i>				
Northern Europe		0.7377***	0.6521***	0.6434**
Southern Europe		0.3590***	0.7536**	0.2011***
Constant		3.05e+19***	1.7e+122***	2.8+e117***

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Table 44: Estimation multinomial logit model, ame (area)

VARIABLES	Average marginal effects			
	0 (base outcome)	1	2	3
<i>Demographic variables</i>				
Year of birth		0.0037***	-0.0012***	-0.0130***
Gender		-0.0005	-0.0169***	0.1140***
High education		0.0037	0.0009	-0.0116*
<i>Household composition</i>				
Relationship status		-0.0038	0.0045*	0.0255***
Number of children		0.0044***	-0.0014	-0.0168***
<i>Health status</i>				
Good health		-0.0115**	-0.0076***	0.0260***
<i>Economic variables</i>				
Household income		-0.0054***	-0.0014	0.0060*
Eligibility		0.0494***	-0.0384***	0.4736***
<i>Geographical area</i>				
Northern Europe		-0.0036	-0.0029	-0.0338***
Southern Europe		-0.0084*	0.0186***	0.1471***

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 45: Indicators of goodness - Multinomial logit model (country)

Indicators of goodness	
Log likelihood	-10096.883
Number of obs	16994
LR chi2(39)	13239.82
Prob > chi2	0.0000
Pseudo R2	0.3960

Table 46: Indicators of goodness - Multinomial logit model (area)

Indicators of goodness	
Log likelihood	-10220.565
Number of obs	16994
LR chi2(30)	12992.46
Prob > chi2	0.0000
Pseudo R2	0.3886

Tables 45 and 46 describe the properties of the model in terms of goodness. The values of the R squared for the multinomial logit is largely higher than in the logistic case, specifically it is equal to 0.40 compared to the value of its corresponding logistic representation (0.03). Also in the multinomial logit model, the representation with countries as single variables has a slightly higher value of the R squared. Nevertheless, the value of the R squared in the multinomial logit model has not the same meaning as the one in the OLS estimation.

7 Concluding remarks

The aim of the study was to identify the factors, which encourage individuals to retire. The SHARE survey allows to narrow the search to people, who are 50 or older, and so they are closer to the retirement age. In particular, research was focused on some European countries, where there is much concern about future public spending on pensions. One objective of the analysis was also to recognize if specific countries or geographical areas have some particular conditions, which make transition to retirement more favorable. The analysis was carried out considering a longitudinal sample in two periods, in order to detect if there were any transition from one period to the other. The first estimation was conducted using a logistic regression, which identifies the transition to retirement as a binary dependent variable. The second model was a multinomial logit model with a multi-category outcome that could distinguish transition or not in both directions. The explanatory variables were chosen among the questions of the SHARE survey and they included different indicators about the quality of life of the respondents.

The main findings of the analysis are the following:

- Individuals, who passed the pension eligibility age are much more likely to be retired in the following period;
- Having a good health condition and a high household income reduces the probability to move from a non-retired status to a retired condition;
- In the logit model, the number of children increases the probability to retire, even though the control variable does not specify their age;
- In the multinomial case, there is a negative correlation between year of birth and probability to retire, so it is more likely to retire, as the age increases;
- If we aggregate countries in macro-areas, Northern and Southern Europe have less probability to retire with respect to Central Europe. However, the same negative correlation does not necessarily derive from the same causes.

Some weaknesses of these representations are due to the fact that they are trying to estimate an uncommon event in this sample, and that the independent variables are not able to completely define some attributes. As for

future research, a further specification of the model using more detailed control variables could provide more significant results. The SHARE survey is a large source of information, which provides many other interesting indicators to specify the variables already used in these models. Moreover, individual variables could be complemented by additional country attributes. At the same time, this type of dataset is quite complex, where the more detailed the information is, the higher is the number of missing values, which makes the analysis more challenging. Another suggestion for further developments is to include more countries in order to expand the sample of the macro-areas, as from the next wave the SHARE survey will cover all members of the European Union. It is important to pursue analyses of this nature to help policymakers find measures to indirectly induce people to retire later. The challenges that modern economies are currently facing to cope with future public spending in pensions, compel them to implement policies, which may become economically and politically unsustainable in the long run.

References

- [1] Achdut, Leah, Aviad Tur-Sinai, and Rita Troitsky. "Transitions between states of labor-force participation among older Israelis." *European journal of ageing* 12.1, 2015: 39-49.
- [2] Agasisti, Tommaso. "Performances and spending efficiency in higher education: a European comparison through non-parametric approaches." *Education Economics* 19.2, 2011: 199-224.
- [3] Bartel, Ann P., and Nachum Sicherman. "Technological change and retirement decisions of older workers." *Journal of Labor Economics* 11.1, Part 1, 1993: 162-183.
- [4] Bonoli, Giuliano. "Classifying welfare states: a two-dimension approach." *Journal of social policy* 26.3, 1997: 351-372.
- [5] Bonoli, Giuliano, and Bruno Palier. "When past reforms open new opportunities: Comparing old-age insurance reforms in Bismarckian welfare systems." *Social Policy & Administration* 41.6, 2007: 555-573.
- [6] Börsch-Supan, Axel. "Incentive effects of social security on labor force participation: evidence in Germany and across Europe." *Journal of public economics* 78.1-2, 2000: 25-49.
- [7] Börsch-Supan, Axel, Karsten Hank, and Hendrik Jürges. "A new comprehensive and international view on ageing: introducing the 'Survey of Health, Ageing and Retirement in Europe'." *European Journal of Ageing* 2.4, 2005: 245-253.
- [8] Börsch-Supan, Axel, et al. "Data resource profile: the Survey of Health, Ageing and Retirement in Europe (SHARE)." *International journal of epidemiology* 42.4, 2013: 992-1001.
- [9] Börsch-Supan, Axel, and Hendrik Jürges. "The survey of health, aging, and retirement in Europe." *Methodology*, 2005.
- [10] Burtless, Gary, and Robert A. Moffitt. "The joint choice of retirement age and postretirement hours of work." *Journal of Labor Economics* 3.2, 1985: 209-236.
- [11] Busemeyer, Marius R. "Determinants of public education spending in 21 OECD democracies, 1980–2001." *Journal of European Public Policy* 14.4, 2007: 582-610.

- [12] Castles, Francis G. "Explaining public education expenditure in OECD nations." *European Journal of Political Research* 17.4, 1989: 431-448.
- [13] Costa, Dora L. "The evolution of retirement." *The Evolution of Retirement: An American Economic History, 1880-1990*. University of Chicago Press, 1998. 6-31.
- [14] European Commission, Directorate-General for Economic and Financial Affairs, "Pension reforms in the EU since the early 2000's: Achievements and challenges ahead", 2016.
- [15] European Parliament, Directorate General for Internal Policies, "Pension Schemes" *Study for the EMPL Committee*, 2014.
- [16] Eurostat, "The 2018 Ageing Report: Economic and Budgetary Projections for the EU Member States (2016-2070)", 2018.
- [17] Eurostat, "Gross domestic product (GDP) at market prices - annual data", dataset,
<https://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tipsau10&plugin=1>
- [18] Eurostat, "Expenditure on pensions", dataset,
<https://ec.europa.eu/eurostat/web/products-datasets/-/tps00103>
- [19] Eurostat, "Mean and median income by household type - EU-SILC survey", dataset, 2018,
https://ec.europa.eu/eurostat/web/products-datasets/-/ilc_di04
- [20] Eurostat, "People at risk of poverty or social exclusion", dataset, 2018,
https://ec.europa.eu/eurostat/web/products-datasets/-/t2020_50
- [21] Fields, Gary S., and Olivia S. Mitchell. "The Effects of Social Security Reforms on Retirement Ages and Retirement Incomes.", Working paper, 1984.
- [22] Grech, Aaron George. "Assessing the sustainability of pension reforms in Europe." *Journal of International and Comparative Social Policy* 29.2, 2013: 143-162.
- [23] Grech, Aaron George. "Pension reforms in the 1990s and during the financial crisis: More of the same?.", 2015
- [24] Gustman, Alan L., and Thomas L. Steinmeier. "Partial retirement and the analysis of retirement behavior." *ILR Review* 37.3, 1984: 403-415.

- [25] Honig, Marjorie, and Giora Hanoch. "Partial retirement as a separate mode of retirement behavior." *Journal of Human Resources*, 1985: 21-46.
- [26] Lalive, Rafael, and Pierpaolo Parrotta. "How does pension eligibility affect labor supply in couples?." *Labour Economics* 46, 2017: 177-188.
- [27] Laun, Tobias, and Johanna Wallenius. "Social insurance and retirement: a cross-country perspective." *Review of economic dynamics* 22, 2016: 72-92.
- [28] OECD, "OECD Private Pensions Outlook 2008", 2009
- [29] OECD, "Pensions at a Glance 2015 - OECD and G20 Indicators", 2015
- [30] OECD, "OECD Pensions Outlook 2016", 2016
- [31] OECD, "Education at a Glance 2017 - OECD Indicators", 2017
- [32] Overbye, Einar. "Convergence theory reconsidered: The politics of pensions in Scandinavia and Australia." *Australian Journal of Political Science* 32.1, 1997: 7-28.
- [33] Pozzebon, Silvana, and Olivia S. Mitchell. "Married women's retirement behavior." *Journal of Population Economics* 2.1, 1989: 39-53.
- [34] Ruhm, Christopher J. "Bridge jobs and partial retirement." *Journal of Labor Economics* 8.4, 1990: 482-501
- [35] SHARE, <http://www.share-project.org/en/data-documentation/waves-overview/wave-7.html>
- [36] SHARE, Questionnaire Wave 5,
http://www.share-project.org/fileadmin/pdf_questionnaire_wave_5/SHARE_paperversion_5_4_10_en_GB.pdf
- [37] SHARE, "Release Guide 6.1.0", 2018
- [38] Staubli, Stefan, and Josef Zweimüller. "Does raising the early retirement age increase employment of older workers?." *Journal of public economics* 108, 2013: 17-32.
- [39] Sundén, Annika. "The Swedish pension system and the economic crisis." *Center for Retirement Research, Boston College, Chestnut Hill, MA* 2009.
- [40] Tranmer, Mark, and Mark Elliot. "Binary logistic regression." *Cathie Marsh for census and survey research* paper 20, 2008.

- [41] UCLA: Statistical Consulting Group, Multinomial logistic regression - Stata annotated output,
<https://stats.idre.ucla.edu/stata/output/multinomial-logistic-regression>
- [42] UNESCO, International Standard Classification of Education ISCED 1997, http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm
- [43] Van Dalen, Hendrik P., Kène Henkens, and Joop Schippers. "Dealing with older workers in Europe: a comparative survey of employers' attitudes and actions." *Journal of european social policy* 19.1, 2009: 47-60
- [44] Wahrendorf, Morten, et al. "Who in Europe works beyond the state pension age and under which conditions? Results from SHARE." *Journal of population ageing* 10.3, 2017: 269-285.

A Appendices

- Current job situation: it is found in the employment and pensions module (EP) of the dataset and the options are:
 1. Retired: it includes «semi-retired, partially retired, early retired, pre-retired; [r]etired refers to retired from own work only, so recipients of survivor pensions are not considered as retired» [36].
 2. Employed or self-employed (including working for family business)
 3. Unemployed
 4. Permanently sick or disabled
 5. Homemaker
 6. Other (Rentier, Living off own property, Student, Doing voluntary work)
- Country: it is available in all modules of the dataset and it is the country, where the respondents have their domicile.
- Year of birth: it is found in the module about demographic variables (DN). They have been kept all respondents, who were born before 1963 for Wave 5 and before 1965 for Wave 6.
- Education: it is in a specific module about education (GV_ISCED), where the variables have been generated by SHARE. The chosen variable registers the values according to the 1997 International Standard Classification of Education maintained by UNESCO [42]. The module includes also the 2011 ISCED classification, but the former has been preferred for a higher response rate. The codification is as follows (supported by an example):
 1. None
 2. Primary education or primary school
 3. Lower secondary or middle school
 4. (Upper) secondary or high school
 5. Post-secondary
 6. 1st stage of tertiary or college degree (either Bachelor's or Master's)
 7. 2nd stage of tertiary or PhD
 8. Still in school

9. Other

- Gender: it is included in the set of demographic variables.
- Relationship status: it is a dichotomous generated variable in the set about imputations (GV_IMPUTATIONS). It states if an individual is in a couple or not. This variable has been preferred to answers on marital status due to a lack of data for the latter. The variable for relationship status has been chosen after a cross-analysis with other similar variables such as marital status (DN) and the imputation sample 3 (GV_IMPUTATIONS), which includes all individuals in a couple. Besides, we are interested more in the presence of another individual, who might influence the decision of the partner, than if they effectively have a registered relationship.
- Number of children: It is a discrete variable that indicates how many children an individual has. It is in the module about children (CH). In case of couples, sometimes the answer was given only by one member of the couple. In this case, the couple has been identified through the couple identification number, and the number of children has been imputed also to the partner.
- General health condition: it is in the set of variables about physical health (PH). The respondents are asked to describe their health by choosing one of the options of the survey.
- Household income: it is in the module about household income (HH), and it is intended as the overall income, after taxes and contributions, that the household has in an average month. In case of couples, sometimes the answer was given only by one member of the couple as for the number of children. The couple has been, then, identified through the couple identification number, and the household income has been imputed also to the partner. Values are all expressed in Euros and in the module GV_EXRATES, there are the exchange rates used in each Wave to convert local currencies for non-Euro countries.
- General economic condition: it is included in the consumption (CO) module. The respondents are asked to state if the household is able to make ends meet in comparison with its total monthly income.
- Expectations: the three variables are in the expectations set (EX). Here are the questions reported, which have been asked only if the individual

answered "Employed" as current situation job:

1. «What are the chances that before you retire the government will reduce the pension which you are entitled to?»
 2. «Thinking about your work generally and not just your present job, what are the chances that you will be working full-time after you reach age 63?»
 3. «What are the chances that before you retire the government will raise your retirement age?»
- Old age pension eligibility: it is a constructed variable, which defines if an individual is eligible for pension benefits according to the statutory retirement age of his country of origin.

B

Table 47: Expectations about future reduction of pensions by country

(a) Denmark				(b) Sweden		
Expectation	Freq.	Percent	Cum.	Freq.	Percent	Cum.
Don't know	26	1.45	1.45	18	1.31	1.31
0-24	426	23.76	25.21	369	26.86	28.17
25-49	132	7.36	32.57	93	6.77	34.93
50-74	584	32.57	65.14	425	30.93	65.87
>75	625	34.86	100.00	469	34.13	100.00
Total	1,793/3,015	100.00		1,374/2,290	100.00	
(c) Germany				(d) France		
Expectation	Freq.	Percent	Cum.	Freq.	Percent	Cum.
Don't know	29	1.65	1.65	24	2.09	2.09
0-24	759	43.08	44.72	158	13.78	15.87
25-49	113	6.41	51.14	48	4.18	20.05
50-74	379	21.51	72.64	305	26.59	46.64
>75	482	27.36	100.00	612	53.36	100.00
Total	1,762/3,001	100.00		1,147/1,725	100.00	
(e) Italy				(f) Spain		
Expectation	Freq.	Percent	Cum.	Freq.	Percent	Cum.
Don't know	41	3.37	3.37	72	6.45	6.45
0-24	170	13.98	17.35	78	6.98	13.43
25-49	83	6.83	24.18	43	3.85	17.28
50-74	414	34.05	58.22	299	26.77	44.05
>75	508	41.78	100.00	625	55.95	100.00
Total	1,216/1,855	100.00		1,117/1,836	100.00	

Table 48: Expectation to work after age 63 by country

(a) Denmark				(b) Sweden		
Expectation	Freq.	Percent	Cum.	Freq.	Percent	Cum.
Don't know	7	0.48	0.48	5	0.59	0.59
0-24	304	21.01	21.49	198	23.52	24.11
25-49	69	4.77	26.26	50	5.94	30.05
50-74	252	17.42	43.68	125	14.85	44.89
>75	815	56.32	100.00	464	55.11	100.00
Total	1,447/3,015	100.00		842/2,290	100.00	

(c) Germany				(d) France		
Expectation	Freq.	Percent	Cum.	Freq.	Percent	Cum.
Don't know	8	0.54	0.54	10	0.96	0.96
0-24	517	35.17	35.71	549	52.44	53.39
25-49	93	6.33	42.04	74	7.07	60.46
50-74	281	19.12	61.16	186	17.77	78.22
>75	571	38.84	100.00	228	21.78	100.00
Total	1,470/3,001	100.00		1,047/1,725	100.00	

(e) Italy				(f) Spain		
Expectation	Freq.	Percent	Cum.	Freq.	Percent	Cum.
Don't know	25	2.39	2.39	49	5.58	5.58
0-24	219	20.96	23.35	135	15.38	20.96
25-49	68	6.51	29.86	39	4.44	25.40
50-74	267	25.55	55.41	5 200	22.78	48.18
>75	466	44.59	100.00	455	51.82	100.00
Total	1,045/1,855	100.00		878/1,836	100.00	

Table 49: Expectations about future raise in retirement age by country

(a) Denmark				(b) Sweden		
Expectation	Freq.	Percent	Cum.	Freq.	Percent	Cum.
Don't know	20	1.12	1.12	15	1.09	1.09
0-24	759	42.33	43.45	434	31.68	32.77
25-49	136	7.59	51.03	96	7.01	39.78
50-74	399	22.25	73.28	281	20.51	60.29
>75	479	26.72	100.00	544	39.71	100.00
Total	1,793/3,015	100.00		1,370/2,290	100.00	
(c) Germany				(d) France		
Expectation	Freq.	Percent	Cum.	Freq.	Percent	Cum.
Don't know	27	1.53	1.53	15	1.31	1.31
0-24	608	34.53	36.06	235	20.47	21.78
25-49	129	7.33	43.38	51	4.44	26.22
50-74	462	26.24	69.62	268	23.34	49.56
>75	535	30.38	100.00	579	50.44	100.00
Total	1,761/3,001	100.00		1,148/1,725	100.00	
(e) Italy				(f) Spain		
Expectation	Freq.	Percent	Cum.	Freq.	Percent	Cum.
Don't know	39	3.21	3.21	74	6.62	6.62
0-24	269	22.19	25.33	107	9.57	16.19
25-49	57	4.69	30.02	28	2.50	18.69
50-74	342	28.13	58.14	224	20.04	38.73
>75	509	41.86	100.00	685	61.27	100.00
Total	1,216/1,855	100.00		1,118/1,836	100.00	