Employment Duration and Shifts into Retirement in the EU[1]

Ted Aranki

European Central Bank
Kaiserstrasse 29, D-60311, Frankfurt am Main.

Corrado Macchiarelli

European Institute, London School of Economics and Political Science
Houghton Street, London WC2A 2AE

Abstract

The decision to cease working is traditionally influenced by a wide set of socio-economic and environmental variables. In this paper, we study transitions out of work for 26 EU countries over the period 2004-2009 in order to investigate the determinants of retirement based on the Eurostat Survey on Income and Living Conditions (EU-SILC). Applying standard survivor analysis tools to describe exits into retirement, we do not find any significant differences in the patterns into retirement between the average euro area and EU non-euro area countries. Moreover, we find that shifts into retirement have increased during the onset of the 2009 economic and financial crisis. Income, together with flexible working arrangements, is found to be important as regards early retirement decisions, compared to retiring beyond the legal retirement age. Finally, we show that institutional measures (such as, state/health benefits, minimum retirement age) could not be sufficient alone if individuals withdraw earlier from the labour market due to a weakening of their health. Especially, these latter results are of importance for structural and macroeconomic policy, for instance, in increasing the employment of both people and hours worked against the background of population ageing.

JEL Classification: J14, J26, C41

Keywords: Retirement, ageing population, hazard model, duration analysis, EU countries.

[1] The authors are grateful to Julian Morgan, Bob Anderton and other participants at an internal seminar organized by the Directorate Economic Developments of the European Central Bank. The authors also acknowledge comments from an anonymous referee. The paper was drafted while Corrado was an economist at the ECB's Directorate General Economics.

1. Introduction

Population ageing is expected to result in a slowdown of labour force growth and, later, into its contraction and change in composition, as projected by the "2012 Ageing Report" by the European Commission (2011). While it is accepted that the demographic shift will add to pressure on the sustainability of public finances in many European countries, the implications for the long-term growth of the labour force are still open issues. If, on the one hand, labour demand is expected to be lower, owing to the shrinkage of the working age population, on the other hand, participation rates for certain age cohorts could increase as well given that working lives will be longer, foreseeing pension reforms. Both aspects prompt several policy questions on labour market developments; such as how to promote longer working lives or how to improve choices for those workers forced to continue to work late in their lives.

Policy makers have been promoting the expansion of working lives finding measures to postpone the labour market activity. However, reflecting retirement patterns, the decision to enter retirement will no longer be a discrete choice: with some workers remaining fully in employment and/or others reducing the number of hours worked as they age (see also Hurd, 1993). Therefore, understanding in greater details the motivations for retirement is key as it could assist the formulation of policies encouraging the return of retirees to employment or decreasing the incentives of withdrawing earlier from the labour market.

Workers are often assumed to dace the choice of leaving the labour market based on their own preferences (Fengler, 1975; Hayward, Grady and McLaughlin, 1988) and/or based on the trade-off between market work versus home production or leisure (for an overview see Duggan, 1984; Bazzoli 1985; Blöndal and Scarpetta 1999; Duval 2003; Gruber and Wise 2002; Meghir and Whitehouse 1997). This latter specification has been particularly supported in modern micro-founded models (e.g., of the New Keynesian type) for macroeconomic analysis. In practice, however, different constraints can influence the labour force participation decisions of the elderly.

In this paper a wide set of socio-economic and environmental variables is employed to study exits into retirement in the EU. Based on longitudinal data from the Eurostat Survey on Income and Living Conditions (EU-SILC) – covering the period 2004-2009 – we analyse the probability of retiring at a given age, given that the person has not retired yet. In particular, we study transitions from employment into (early) retirement by using a hazard based duration model framework (see also Diamond and Hausman, 1984; Hagan, Jones and Rice, 2009; Jones, Rice and Roberts, 2010).

The contribution of this paper can be gauged under two perspectives. First, we provide, for the first time, results for a large set of 26 EU countries, by providing a systematic, conditional approach to estimate labour market shifts into retirement. Secondly, we exploit cross country differences, including measures of between-country heterogeneity, in quantifying the size and the speed with which employment-to-retirement changes took place.

Since it would be natural to hypothesize upfront that retirement dynamics has changed over time – especially during 2009, reflecting the extent to which the global economic and financial crisis hit in most countries – and differs across euro area versus non-euro area countries – reflecting region-specific dynamics and institutional set ups – we model this explicitly. Nonetheless, the results in this paper do not support any significant differences in the patterns into retirement between the average euro area and EU non-euro area countries. However, shifts into retirement seem to have increased during the onset of the 2009 crisis, when controlling for income.

Turning to personal and household-level characteristics, income and benefits (also temporary in nature, e.g., sickness benefits) are found to be important as regards early retirement decisions – when accumulated income/wealth is presumably lower – compared to retiring beyond the legal retirement age. In the same vein, flexible working arrangements are found to be important in order to keep people at work beyond the legal retirement age, thus suggesting that making use of partial working schemes could modify retirement patterns towards postponing the labour market withdrawal.

Finally, this analysis shows that institutional measures (such as, state/health benefits, minimum retirement age) could not be sufficient alone if individuals withdraw earlier from the labour market due to a weakening of their health. Particularly, for early retirees, policies aimed at improving the health of the workforce and at keeping people who experience health problems active may be crucial.

Particularly, these latter results have implications for the effectiveness of active labour market policies, by getting retired people back into work or helping the prolongation of long term employment spells. Moreover, the findings are of importance for structural and macroeconomic policy, for example, in increasing the employment of both people and hours worked against the background of population ageing.

The reminder of the paper is organized as follows: Section 2 describes the data and provides a brief descriptive analysis. Section 3 presents the econometric strategy. Section 4 outlines the main results. Section 5 concludes.

2. Data

In this paper we use the Eurostat Survey on Income and Living Conditions (EU-SILC) which consist of a database available in yearly frequencies, based on a rotating panel of longitudinal data for 4 sub-samples. The EU-SILC provides the longest time series of comparable and consistently defined individual level data for income and living conditions available for the EU, and our sample consists of 26 countries covering the period 2004-2009. The sample excludes Germany owing to data unavailability.[2] Compared to other surveys, the EU-SILC provides not only details on each individual's personal characteristics (i.e. gender, age, marital status, education, family composition, etc.), but also information on the level of (household) income prior to retiring and measures of the individual's wealth status. This represents an advantage compared to other analyses, given that income and wealth can be important determinants of retirement decisions (see for instance, Hanoch and Honig, 1983; Mitchel and Fields, 1984; Dugan, 1984; Ruhm, 1990).[3]

An individual's transition from employment into retirement is the event of interest in this study. From the EU-SILC, we construct transitions from employment to retirement, or of remaining in employment, based on each respondent's current and past activity status.

Moving from employment to retirement, or retiring in the next period, is typically referred to as a 'failure' event which can occur at any point in time after an 'onset of risk' period is defined. Here, the 'onset of risk' period is defined as each individual's first entry into the labour market.

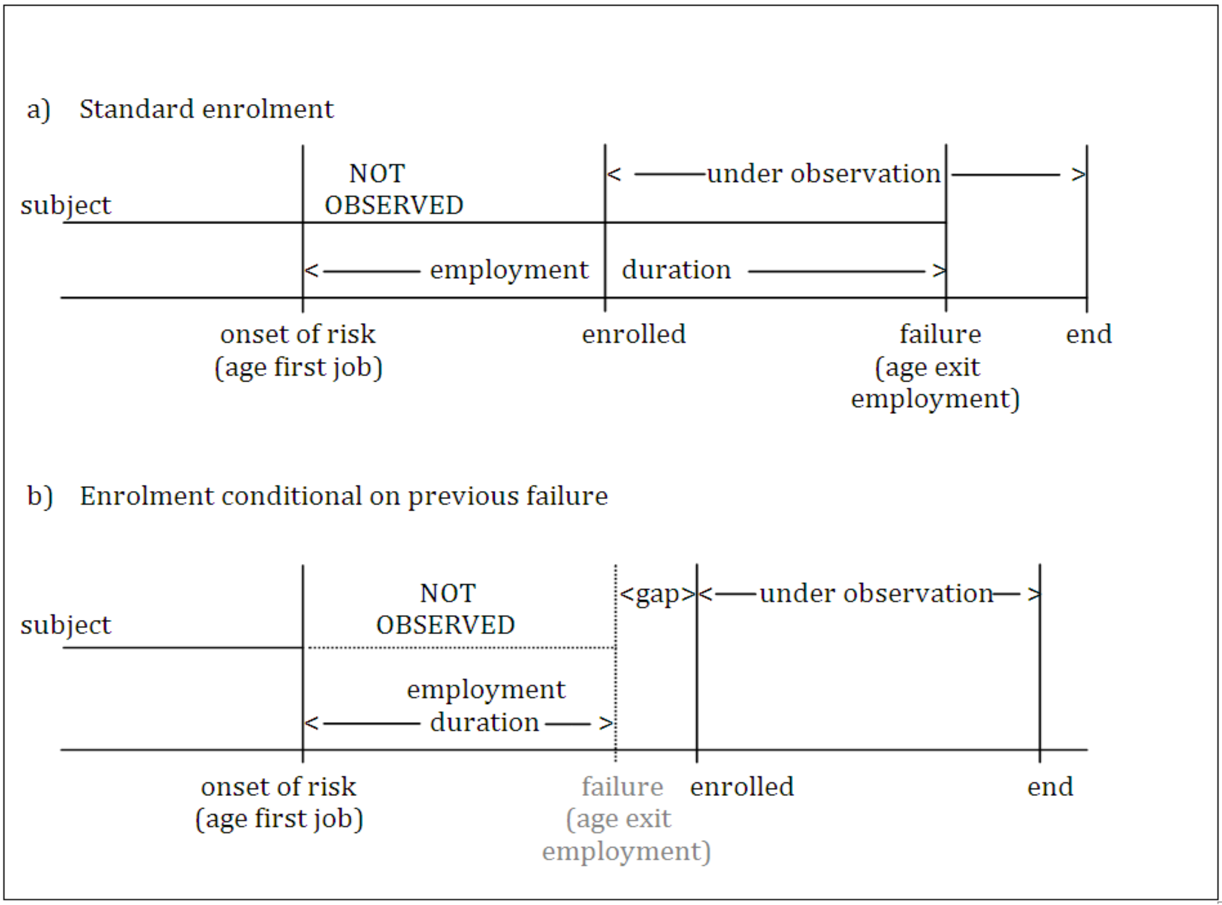
[2] Germany is covered by EU-SILC but their longitudinal microdata are not disseminated according to the EC Regulation no. 223/2009.

[3] Overall, however, the EU-SILC is not designed to distinguish between job transitions and short retirement spells during the working life. Yet, although there exists an observed retirement framework based around state pension eligibility in each country (see OECD, 2011), people make transitions into and out of work until advanced ages, making observed employment histories rather complex.

To analyse transitions from employment into retirement, a hazard based duration model framework is employed (see, Diamond and Hausman, 1984; Hagan, Jones and Rice, 2009; Jones, Rice and Roberts, 2010). This allows modelling the length of time spent at work before moving into retirement. The dependent variable is thus the amount of time that an individual spends in employment before entering retirement (i.e. employment duration).[4]

One statistical motivation for employing a duration analysis framework includes the presence of censored and left-truncated data. In practice, not all observations' full history is observed until the 'failure' event. This naturally classifies the EU-SILC data as right-censored. Instead, the left truncation problem refers to the fact that individuals become at risk or even fail before we can enrol them in the study (see Figure 1).

Figure 1 – Examples of left truncation in the EU-SILC data



[4] In Table A3 in the Appendix we detail the variables used in the estimation.

Left truncation is a natural feature of our data and involves the impossibility of tracking individuals over the whole working life. Taking into account this particular data structure, an individual's full employment history is inferred based on retrospective information about the age at which he / she first started to work and the years spent on paid work. This formalisation does not clearly take into account the possibility of multiple 'failure' events within the same employment history, but rather assumes that each individual's working history is continuous until retirement.

In the same vein, individuals can enter the observation period / being enrolled in the study upon having already retired. Here, an important difference compared to standard duration analyses is that the failure event does not represent a rationale for an individual to drop out of sample (e.g., as death). Whenever enrolment occurs conditional on a previous retirement event (see Figure 1(b)), there may exist a positive difference between employment duration and the year of enrolment in the study, representing a *gap* of information about each individual's activity between the period he / she ceased working and the period he / she became under observation. Only those with a *gap* < |G|, where G is arbitrarily chosen, as well as those reporting to have most recently changed their activity status 'from employment to retirement' are considered in this analysis. The gap variable is however not restricted to be exactly zero, i.e. G = 3, allowing for reporting errors in (*i*) age, (*ii*) age of the first job and (*iii*) the number of years spent on paid work. Importantly, this is not found to significantly affect our results, as the inclusion of a whole set of covariates in the regressions will anyway require censoring on many of these individuals with *0* < *gap* < |G|,[5]

As individuals' working histories are inferred based on retrospective questions, only the last spells are considered, for individuals employed at all times. Conversely, only the spell of retirement is considered for individuals retired in the sample. This allows data tractability in such a duration analysis framework.

The dataset employed consists of 209.183 individuals. Out of this number, 6756 individuals, that is just over 3% of the sample, are observed retiring. As Table 1 suggests, the majority of these retirees ceased working at the age of 55 or later. Women represent nearly 40% of the sample.

---

[5] The analysis in the empirical section requires individuals to be observed at least for two consecutive periods (t-1 and t). For instance, an individual retired at time t, should provide information on his previous (t-1) employment status (be it part time or full time employment) or the occupation sector in which he / she most recently worked prior to retiring. Thus, when individuals are enrolled upon having retired, information on previous employment status is clearly missing, making those individuals not eligible for the empirical analysis in Section 4.

Table 1: Retirees by age group and gender

	Percent of all observations		
Age Groups	Males	Females	Total
Age 0 to 24	0.1	0.1	0.1
Age 25 to 29	0.2	0.1	0.2
Age 30 to 34	0.2	0.2	0.2
Age 35 to 39	0.4	0.4	0.4
Age 40 to 44	1.3	0.8	1.1
Age 45 to 49	2.6	1.9	2.3
Age 50 to 54	5.2	5.0	5.1
Age 55 to 59	25.0	40.7	31.0
Age 60 to 65	47.9	37.8	44.1
Age > 66	17.2	12.9	15.6
Total observations	4191	2565	6756

Gender representation is reversed for ages between 55 and 59, where women represent the majority of the sample considered. In general, female workers retire earlier than males. Nearly 50% of female workers enter into retirement before the age of 60, compared to 35% of the male workers.

In our sample, the age of retirement spans from a minimum of 20 years to a maximum of 80 years, with the average occurring at the age of 60.4 and the median at 60. Thus, the distribution of retirees is clearly skewed towards older people (see Table 2).

Table 2: Distribution of retirees by age

Percentiles		Smallest		
1%	41	20		
5%	50	20		
10%	55	21	Obs	6756
25%	57	22		
50%	60		Mean	60.453
		Largest	Std. Dev.	6.349
75%	64	80		
90%	67	80	Variance	40.312
95%	70	80	Skewness	-0.662
99%	78	80	Kurtosis	7.431

3. Empirical methodology

In this paper we employ a hazard based duration model framework to study the transitions from employment into retirement. The main advantage of using duration analysis is that it allows modelling the length of time spent in a given state (i.e. employment) before moving into another state (i.e. retirement). Relative to other approaches such as those that focus on the unconditional probability of an event taking place (e.g. a *probit* or *logit* models), our focus here is on the conditional probability, or, the probability that the spell of one particular status (e.g., employment) will end in the next short interval of time, given that it has lasted until recently.

As the analysis is concerned with the timing of the observed change from employment to retirement (or 'failure' event, see Section 2), it makes sense to conceptualize the length of each individual *j*'s employment spell as a random variable, Tj.[6] Assuming Tj has a continuous probability distribution $f(t)$, where *t* is a realisation of T_j, the cumulative distribution function of *T* will be given by $F(t) = Pr(T_j \leq t) = \int_0^t f(s)ds$.

This says that the survival function for the *j*-th individual, or the probability that his employment spell T is of length at least equal to *t*, is:

$$S(t) = 1 - F(t) = Pr(T_j > t) = \int_t^\infty f(s)ds \tag{1}$$

Conversely, the hazard rate (or instantaneous failure rate) for individual *j* at time *t*, is defined instead as the marginal probability of immediate retirement, conditional on not having retired before time *t*:

$$h(t) = P\left(t < T_j < t + dt \mid T_j > t\right) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} \tag{2}$$

This class of models can be distinguished between non-parametric, semi-parametric and full parametric models on the basis of whether they predict the probability distribution of a certain event by means of a set of additional covariates. While parametric models are widely used across numerous fields of economics, the Cox proportional hazard model (Cox, 1972; 1975) has proven particularly flexible. Compared to fully parametric approaches, a key benefit of this approach is that it allows to "avoid having to make assumptions about the nature of the duration times in the first place" (Box-Steffensmeier and Jones, 2004). In other

[6] See Box-Steffensmeier and Sokhey (2010) and Jenkins (2008) for a methodological overview.

words, the Cox model makes no assumption about the shape of the hazard function or about how covariates may affect this shape.[7] Thus, the Cox semi-parametric approach is regarded as a benchmark in this paper, whereas non-parametric (Section 4.1) and fully parametric (Appendix) approaches are employed for preliminary investigation of the data and robustness checks respectively.[8]

In the Cox model, the hazard for the *j*-th individual in the data is assumed to be:

$$h(t \mid x_j) = h_0(t)\exp(\beta' x_j) \tag{3}$$

where β' is the vector of regression coefficients; x a vector of covariates which influence the hazard rate; and $h_0(t)$ is the baseline hazard function.[9] By default, the model assumes a baseline hazard that is common to all the individuals in the study population. In this way, covariates act multiplicatively on the baseline hazard, adding additional risks on an individual basis, as determined by the individuals' prognostic information. This gives the model a simple and easily understood interpretation. The main idea behind it is the separation of the time effect in the baseline hazard function, on one side, and the effect of the covariates in an exponential term, on the other. In essence this assumption says that the hazard of failure at time t is related to individuals or groups of individuals by a proportionality constant which does not depend on t.

3.1 Frailty models

When observations are conditionally different in terms of their hazards due to unobserved heterogeneity, standard models, as the one just described, may lead to spurious duration dependence.[10] Hence, fitting a normal duration model, e.g. equation (3), would simply not recognise that some observations are more 'frail' (or, failure prone) than others.

A first possible solution would be to include fixed effects. However, it has been shown that fixed effects are not a viable alternative in this context, as there is an incidental parameter problem that leads to inconsistent and deflated standard errors (see Box-Steffensmeier and Zorn, 2001; Zorn, 2000). An alternative method is

---

[7] For further references see Cleves et al. (2010).

[8] Fully parametric models will be efficient only as long as the distributional assumptions are appropriately chosen in the class of parametric lifetime distributions (e.g., exponential, weibull, gompertz, log-normal, log-logistic or gamma). Clearly, if the hazard function shape is incorrectly specified, parameters can be seriously biased.

[9] In particular, when inference is dependent on the form of $exp(\beta' x)$ but still independent of $h_0(t)$, one speaks of a semi-parametric model (see Cox; 1972, 1975).

[10] The notion of unobserved heterogeneity amounts here to observations being *conditionally* different in terms of their hazards in ways that are unaccounted for in the systematic part of the model.

to use random effects or ‘frailty’ models instead. The basic idea behind frailty models is to introduce an additional random parameter into the hazard rate accounting exactly for unobserved heterogeneity. Frailties may be individual-specific or group-specific. Models constructed in terms of group-level frailties are typically referred to as ‘shared’ frailty models because observations within a sub-group are assumed to share unmeasured risk factors prompting them to fail earlier.

Lancaster (1979) proposed a parametric mixed proportional hazard model, accounting for unobserved ‘frailties’, which is a generalization of Cox’s (1972) approach. This specifies the hazard rate for the *j*-th individual as (see also Box-Steffensmeier and Jones, 2004; Zorn, 2000):[11]

$$h(t \mid x_j) = h_0(t)\exp(\beta' x_j)\exp(v_j) \tag{4}$$

where $v_j = W_j \psi$ describes the individual- or group-specific unobserved heterogeneity. For identification purposes, the mean of v is typically normalized to unity and its variance is assumed to equal an (unknown) parameter θ.[12] Compared to the standard Cox (1972) regression approach, integrating v out leaves with the only problem of estimating the additional parameter, θ, in the survivor function:[13]

$$S(t \mid v) = \int_0^t - h(s,v)ds = -v\int_0^t h_0(s)ds \tag{5}$$

4. Results

4.1 Non-parametric analysis

In order to summarise the data and visualise the distribution shape of employment duration for the sample or for separate groups, non-parametric estimation of the survivor and hazard functions relying on product-limit estimators are introduced.[14] Table 3 reports the survivor and cumulative hazard function for employment

[11] In essence the concept goes back to the work of Greenwood and Yule (1920) on *accident proneness*. The term frailty itself was introduced by Vaupel *et al.* (1979) in univariate survival models.
[12] As Box-Steffensmeier and Jones (2004) note, we always make assumptions about whether we use frailty models or not. When we do not take account of frailty, we are essentially assuming that v=1 with probability 1.
[13] To derive the expected value of the survivor function, a probability distribution for *v* needs to be specified. Albeit the gamma is the most common in the literature, any continuous distribution with positive support, a unit mean, and a finite variance θ – inverse Gaussian, log-normal etc. would be appropriate. Essentially, as long as we assume that *v* has some distribution, we can estimate the frailty model by estimating the frailty variance term θ.
[14] We use the Kaplan-Meier (1985) and the estimators dating back to Nelson (1972) and Aalen (1978) (referred to as Nelson-Aalen estimator) for the estimation of the survivor and cumulative hazard function respectively. For further details see also Kiefer (1988).

duration. The survivor function shows the proportion of people who remain in employment (i.e. do not 'fail' by entering retirement) as time proceeds, while the cumulative hazard shows the expected number of 'failures' at each observed time. On average, after 40 years of work, the survivor function starts decaying very rapidly, with the proportion of people still employed decreasing over time. This is in line with the idea that the definable pensionable age requires around 40 years of contribution, consistent with the evidence in OECD (2011) and European Commission (2011). Still, different conditions may apply depending on the number of years of contributions achieved at a certain date or the age of first entry into the pension system.

As shown in Table 3, after 45 years of work the probability of remaining in employment is around 0.64, indicating that roughly 36% of the sampled individuals where retired. Furthermore, the Nelson-Aalen cumulative hazard suggests that the hazard of exiting into retirement increases monotonically.[15] Survivor functions from employment to retirement across different categories, as well as by country, are plotted in Figure A1 and A2 in the Appendix.

Table 3: Kaplan-Meier survival and Nelson-Aalen cumulative hazard functions

Years of work	Beginning total	Failures	Survivor function	Standard error	Cumulative hazard	Standard error
20	86945	19	0.9989	0.0001	0.0011	0.0001
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
37	20953	362	0.9408	0.0014	0.0608	0.0015
38	18164	422	0.9189	0.0017	0.0840	0.0019
39	15431	394	0.8954	0.0021	0.1095	0.0023
40	13184	623	0.8531	0.0026	0.1568	0.0030
41	10509	435	0.8178	0.0030	0.1982	0.0036
42	8705	508	0.7701	0.0035	0.2565	0.0044
43	7076	373	0.7295	0.0039	0.3093	0.0052
44	5820	325	0.6888	0.0043	0.3651	0.0060
45	4751	333	0.6405	0.0047	0.4352	0.0072
46	3723	238	0.5995	0.0051	0.4991	0.0083
47	3010	173	0.5651	0.0054	0.5566	0.0094
48	2459	145	0.5318	0.0058	0.6156	0.0106
49	1995	110	0.5024	0.0061	0.6707	0.0118
50	1672	228	0.4339	0.0067	0.8071	0.0149
51	1232	103	0.3976	0.0071	0.8907	0.0170
52	999	79	0.3662	0.0073	0.9697	0.0192
53	819	63	0.3380	0.0076	1.0467	0.0215

[15] It is worth noting that for the survivor function and the cumulative hazard function, both the Kaplan-Meyer and the Nelson-Aalen estimators are consistent estimates of each function respectively, and their statistics are asymptotically equivalent (see Klein and Moeschberger, 2003).

54	684	38	0.3193	0.0078	1.1022	0.0233
55	583	54	0.2897	0.0080	1.1948	0.0265
56	473	32	0.2701	0.0082	1.2625	0.0291
57	386	32	0.2477	0.0084	1.3454	0.0325
58	324	22	0.2309	0.0086	1.4133	0.0356
59	261	15	0.2176	0.0087	1.4708	0.0386
60	218	23	0.1946	0.0090	1.5763	0.0444
61	167	17	0.1748	0.0093	1.6781	0.0508
62	123	9	0.1620	0.0095	1.7512	0.0564
63	98	10	0.1455	0.0099	1.8533	0.0650
64	79	11	0.1252	0.0102	1.9925	0.0773
65	62	35	0.0545	0.0091	2.5570	0.1228

Note: The standard error for the Kaplan-Meyer estimate is the one given by Greenwood (1926).

4.2 Semi-parametric analysis

In this section, estimates of the semi-parametric Cox proportional hazard models are presented. As discussed in Section 3, parametric analysis offers an advantage over the non-parametric methods, as it allows predicting the probability distribution of retirement by means of a set of additional covariates. In what follows the joint effect of various individual and labour market characteristics affecting the probability of exiting into retirement is analysed.

In Table 4, the estimated results from the Cox's proportional hazard model are reported.[16] The reported coefficients are hazard ratios.[17] As explained in Section 3, a 'shared' frailty model is employed where the sub-groups are selected according to the number of countries in our sample (26). Thus, the model assumes that observations share group-specific, unmeasured, risk factors that prompt exits into retirement. As the frailty terms explicitly account for the extra variance associated with such risk factors, we can evaluate the hypothesis that $\theta = 0$ to determine whether the choice of treating unobserved heterogeneity in the model is motivated. Supporting our concerns, the nested model under $\theta = 0$ is always preferred to the reference non-frailty model according to the relevant LR test at the bottom of Table 4.

Focusing on the regression results, the estimated hazard ratios indicate that there is no significant difference in the patterns of retirement between residents in the euro area (*EA*) and EU non-euro area countries. Moreover, the results suggest that the onset of the global financial crisis (*2009*) significantly increased flows

[16] A sensitivity analysis is performed in the Appendix showing that the results from the Cox proportion model are robust also when using full parametric models.

[17] A coefficient of, e.g., 0.5 for a dummy variable is interpreted as lowering the exit rate from employment to retirement by a half. For a continuous variable, a coefficient of 0.5 implies that a one unit change in the variable is associated with a hazard rate of 1/2 as large and an *n* unit change in the variable is associated with a hazard rate $(1/2)^n$ as large.

into retirement. However, it is only when controlling for household disposable income and personal benefits that we achieve this result. When omitting the income variables from the regression, the result rather suggest that the hazard to retire decreased in 2009.[18] The importance of the income variables for the result of the 2009 crisis on the hazard to retire may stem from the fact that wealth for people eligible to retirement generally became at risk in 2009, with income to cover basic expenses in retirement running short because of the financial crisis.[19]

Table 4: Cox regressions

Variable	Hazard ratio	(Std. Err.)	Variable	Hazard ratio	(Std. Err.)
EA	1.030	(0.320)	*Income variables*		
Dummy 2009	1.301***	(0.049)	Disposable income	0.982	(0.025)
Minimum retirement age	0.896***	(0.012)	Old age benefits	2.028***	(0.085)
			Unemployment benefits	1.864***	(0.117)
Individual characteristics			Disability benefits	2.806***	(0.230)
Female	1.479***	(0.075)	Sickness benefits	1.134	(0.088)
Married	1.024	(0.082)			
Skilled	1.676***	(0.073)	*Interaction*		
Part-time	0.127***	(0.061)	Part-time x disp. income	1.171***	(0.056)
Occupational group			*Health variables*		
2.Occ. group	0.656***	(0.033)	Health	1.682***	(0.201)
3.Occ. group	0.958	(0.048)	Health(-1)	1.460**	(0.224)
4.Occ. group	0.924	(0.055)			
5.Occ. group	0.853**	(0.053)	*Partner characteristics*		
			2.Partner unemployed	1.384***	(0.134)
			3.Partner retired	1.065*	(0.041)
			4.Partner inactive	1.155***	(0.055)
			Partner's health	0.706***	(0.083)
Statistics					
θ	0.577	(0.150)			
LR test (frailty terms)	Prob>=chi-bar-sq. = 0.000				
Wald $\chi 2$	1305.892				
Prob > $\chi 2$	0.000				
Log-likelihood	-32839.825				
Number of groups	26				
Observations	53,490				

Note: Standard errors in parentheses. *** denotes p<0.01, ** p<0.05, * p<0.1. See Table A3 for data and definitions

From Table 4 we also find that setting a *minimum retirement age* reduces the hazard to retire. This can be interpreted in the light of providing workers with a yardstick or a minimum number of years with payment to social security before they become eligible to retire.

[18] The finding of a decreasing hazard to retire in 2009 is consistent with the non-parametric Kaplan-Meier estimates reported in Figure A1 (Appendix). However, it should be borne in mind that the Kaplan-Meier estimates are not conditional on any covariates. Thus, the information extracted from a plain visual inspection of the plots in Figure A1 is very limited, compared to the semi-parametric analysis.

[19] In the present context it is however difficult to distinguish labour market quits from lay-offs.

Beyond more institutional factors, the participation of elderly workers is also affected by a wide set of socio-economic and environmental variables such as gender (*female*) and occupation groups (*occ. group*). While we find increased movements toward retirement among female workers, we do not find any statistical difference as to whether a person is *married* or not.

Besides, the personal characteristics typically associated with higher education (*skilled*) are not found to generally lead workers to work longer than their less-educated counterparts. This is somewhat consistent with the findings in Autor and Dorn (2009) reporting an inverted U-shape relationship between skills and changes in the mean age, suggesting that occupations in the bottom and top deciles of the skill distribution tend to work on average less than people with middle-skill jobs.

Working (or having worked) part-time plays an important role in reducing the hazard to retire. This is consistent with the recent evidence (see, *inter alia*, Machado and Portela, 2012) that retirement is no longer likely to be a discrete choice: with some workers exiting from full-time employment and making use of flexible working schemes before withdrawing completely from the labour market.

Furthermore, some occupation groups are found to have important explanatory power. Compared to the category of professionals, technicians and associate professionals (*occ. group =1*), those belonging to the category including service, skilled agricultural and fishery workers (*occ. group =2*), and those with elementary occupations (*occ. group =5*), show a significantly lower probability to retire. Albeit with such sectoral categories it is not possible to distinguish between private or public sector employees (see Table A3, in the Appendix), these results probably reconcile with the idea that formal workers are expected to retire earlier than casual workers and self-employed, typically belonging to some of the categories listed above (i.e., elementary occupations, agriculture and fishing).

Looking at the income variables, household *disposable income* ultimately does not exert an influence on retirement decisions in our sample. Nevertheless, the interaction between working part-time and disposable income (*part-time x disp. income*) significantly reduces employment duration. This result suggests that there exist a level effect of income when employment is not full-time, or, the exposure to the current (or past) level of income is higher when not working full-time (see also Blake, 2007; Montalto, 2001).

Income supports are very likely to influence the labour supply of the aged as well, given that unemployment schemes may induce older workers to seek part-time jobs or to withdraw earlier from the labour market.[20] The variation in age of eligibility for social security benefits (*old age, disability and/or sickness benefits*) can particularly affect the sustainability of the retirement status. It should be borne in mind that the effect of pension schemes and benefits are not exogenous to income, as pension scheme produce inter-temporal substitution effects (i.e. with a postponement of the retirement age today in favour of an expected higher pension return tomorrow). In this setting, receiving positive old age benefits or unemployment benefits significantly increases retirement decisions, in line with the idea that social insurance schemes such as disability benefits significantly increase flows out of employment (see estimated hazard ratios in Table 4).[21] Consistently, sickness benefits, representing cash benefits that replace (in whole or in part) loss of earnings during temporary inability to work, are not found to significantly affect the hazard to retire.

In line with the literature, our findings also point to the fact that *health* is an important determinant of retirement, as healthier people are found to continue to work and retire later (see *inter alia*, Bound, 1991; Jones *et al.*, 2008; 2010; Deschryvere, 2005; Disney *et al.*, 2006).[22] Overall, however – as highlighted by a growing literature (e.g., Jones *et al.* 2008; 2010) – measures of health are subject to an endogeneity problem. There are several reasons on why to expect an endogeneity bias when using self-reporting measures of health. First, self-reported health is based on subjective assessments which may not be comparable across individuals (Lindeboom, 2006; Lindeboom and van Doorslaer, 2004). Second, there is an obvious simultaneity problem between self-reported health and the labour market status, given that health problems may represent a legitimate reason for a person in the working age to be outside the labour market (Kerkhofs and Lindeboom, 1995; Kreider, 1999). Finally, for some individuals there may be incentives to report health

[20] On the other hand, as argued by Boskin and Hurd (1978), if higher social security taxes are needed to finance the increasing burden of an ageing population, this could create disincentives for people to reduce their labour force participation and withdraw earlier from the labour market.

[21] Note that, in the EU-SILC, unemployment benefits also include (see also Table A3):

(i) Partial unemployment benefits compensating for the loss of wages or salary due to formal short-time working arrangements, and/or intermittent work schedules, irrespective of their cause, and where the employer/employee relationship continues.

(ii) Early retirement for labour market reasons, including periodic payments to older workers who retire before reaching standard retirement age due to unemployment or to job reductions caused by economic measures such as the restructuring of an industrial sector or of a business enterprise. These payments normally cease when the beneficiary becomes entitled to an old age pension.

Thus, receiving unemployment benefits may unveil information about part-time working schemes and early retirement patterns in some cases.

[22] For a survey of the literature see Deschrivere (2005).

problems as a mean to obtain disability benefits (i.e. the so-called 'disability' route into retirement, see Blundell *et al.*, 2002).

Many studies in the literature typically use an instrumental variable approach, by adopting more 'objective' measures of health to instrument self-reported health measures. Along these lines, an 'individual health stock' is normally constructed, where self-reported health is regressed on a set of specific health problems (see also, Griliches, 1974; Fuller, 1987). As such questions concerning specific health problems are not available in the EU-SILC, we take into account the possibility that anticipated retirement may justify the reporting of bad health by, first, including a dummy whenever individuals receives disability benefits. This allow us to control for possible 'disability routes' into retirement (Blundell *et al.*, 2002). Further, to assess the robustness of our previous findings, alternative health measures are employed, along with the usual set of covariates. More specifically, in Table 5 measures arguably less prone to reporting bias than self-reported health are employed, such as a measure of health limitations (*limit*) and *chronic* diseases (see Jones *et al.*, 2008; 2010).

Using alternative health measures has generally a size effect on the coefficients of interest while it does not affect their sign and / or significance. Thus, independently of the proxy employed, health status is an important determinant of retirement decisions.

Table 5: Cox regressions using alternative health measures

	Hazard ratio	(Std. Err.)	Hazard ratio	(Std. Err.)	Hazard ratio	(Std. Err.)
Health	1.682***	(0.201)				
Health(-1)	1.460**	(0.224)				
Limit			1.210***	(0.076)		
Limit(-1)			1.213***	(0.083)		
Chronic					1.050	(0.041)
Chronic(-1)					0.919**	(0.036)
Wald χ^2	1305.892		1304.008		1282.473	
Prob > χ^2	0.000		0.000		0.000	
Log-likelihood	-32839.825		-32842.068		-32851.073	
Number of groups	26		26		26	
Observations	53,490		53,490		53,490	

Note: All regressions include a full set of covariates as in Table 4. The whole results are available upon request from the authors. Standard errors in parentheses. *** denotes p<0.01, ** p<0.05, * p<0.1. See Table A3 for data and definitions.

To assess whether a change in labour market status from employment to retirement is more influenced by a (negative) shock to an individual health or a level effect via slow health deterioration, a 'health shock', or a lagged health variable, is included in the regression, following the discussion in Jones *et al.* (2010). It seems plausible that the health lag is more informative about the decision to retire than current health as it normally takes time to entirely adjust to health limitations and to allow an individual to gauge his reduced ability to work over time. The use of health lag has the great advantage of reducing any endogeneity bias by observing the timing before the decision to effectively retire (see Jones *et al.*, 2010). In Table 4, the effect of the health shocks is significant. This is broadly consistent with the evidence obtained when using alternative health shock measures (see Table 5).[23]

Occupation statuses and health effects are important also as regard to individuals' partners. For instance, predictions regarding a joint labour market decision of old couples can derive from a family labour supply model like the one proposed by Killingsworth and Heckman (1986) were couples maximise a single utility function subject to a household budget constraint with pooled income. The analysis in this paper confirms the prediction that having a *partner retired* significantly increases the hazard to retire, compared to having a partner employed. This is in line with the idea that the primary reason for partners to retire together is shared preferences / substitution effect for leisure against working longer, with each partner valuing more retirement when the partner is retired as well (see Killingsworth and Heckman, 1986; Hurd, 1990; Michaud, 2003). Moreover, individuals with partners reporting bad health are generally associated with a lower probability to retire compared to individuals with partners reporting better health status (see also Wu, 2003).

4.2.1 Early-retirement decisions

Until now the analysis has focused on individuals retiring. However, understanding the motivations to retire *earlier* (before the legal retirement age), compared to standard retirement patterns, represent an important factor of analysis. This can be important, especially in the light of assisting the formulation of policies that might encourage early retirees to stay at work.

[23] Health is also important as concern the interaction with occupation groups (*occ. group*). Such an interaction term suggests that those who work in craft and related trades workers (including heavy works such as extraction and building) have higher incentives to retire due to (reported) health problems. For sake of brevity these results are not reported in Table 4, but are available upon request from the authors.

Table 6: Cox regressions, early retirement sample

Variable	Hazard ratio	(Std. Err.)	Variable	Hazard ratio	(Std. Err.)
EA	0.854	(0.273)	*Income variables*		
Dummy 2009	1.285***	(0.061)	Disposable income	1.082**	(0.036)
Minimum retirement age	0.848***	(0.015)	Old age benefits	3.330***	(0.156)
			Unemployment benefits	1.637***	(0.112)
Individual characteristics			Disability benefits	2.291***	(0.202)
Female	1.373***	(0.084)	Sickness benefits	1.318***	(0.121)
Married	1.135	(0.112)			
Skilled	1.914***	(0.102)	*Interaction*		
Part-time	1.201	(0.751)	Part-time x disp. income	0.962	(0.059)
Occupational group			*Health variables*		
2.Occ. group	0.746***	(0.045)	Health	2.383***	(0.338)
3.Occ. group	0.832***	(0.049)	Health(-1)	1.765***	(0.325)
4.Occ. group	0.796***	(0.054)			
5.Occ. group	0.784***	(0.058)	*Partner characteristics*		
			2.Partner unemployed	1.379***	(0.146)
			3.Partner retired	1.110**	(0.052)
			4.Partner inactive	1.101*	(0.060)
			Partner's health	0.682**	(0.109)
Statistics					
θ	0.604	(0.157)			
LR test (frailty terms)	Prob>=chi-bar-sq. = 0.000				
Wald χ2	1475.982				
Prob > χ2	0.000				
Log-likelihood	-22076.371				
Number of groups	26				
Observations	51,304				

Note: Standard errors in parentheses. *** denotes p<0.01, ** p<0.05, * p<0.1. See Table A3 for data and definitions

Cox estimates of early retirement decisions are reported in Table 6. Although the results are similar to those presented for the full sample, significant differences do exist. In particular:

- Working (or having worked) with a part-time contract does not play a significant role in reducing the hazard to retire early. This finding, combined with the result in Table 4, may suggests that a gradual reduction in hours worked over the last segment of the working life can contribute to increased employment of older workers, beyond the legal retirement age.

- Higher disposable household income and state / health benefits – including those temporary in nature, such as sickness benefits – significantly increase the hazard to retire early, compared to standard retirement decisions. This suggests that the choice of pre-retiring should be considered in the light of the expected retirement needs, or the evaluation of whether the accumulated income / wealth prior to retiring is considered adequate to sustain the future retirement status. In this vein, early retirements are more sensible to income effects (including short term benefits) compared to retiring after the legal retirement age. Along the same lines, and opposite to the results in Table 4, the interaction term between part time and income does not exert any significant effect on early retirement decisions.

. Finally, focusing on the individual health status, we show that – analogously to the findings in Table 4 – people with health problems are generally found to discontinue employment and retire earlier. However, the health coefficient for people retiring below the legal retirement age is twice as big the one reported in Table 4. This points to the fact that, among all the institutional measures scrutinized (long term and short term benefits, minimum retirement age, etc.), none could be sufficient alone if individuals withdraw earlier from the labour market due to a weakening of their health. Particularly for early retirees, policies aimed at advance retirement by improving the health of the workforce and at keeping those who experience health problems active may be essential.

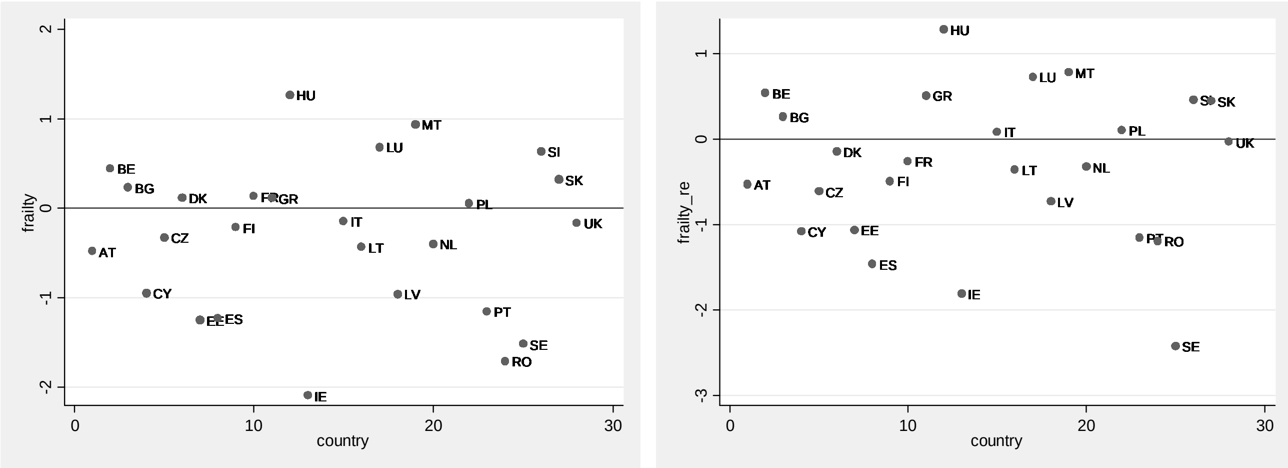
4.3 Frailty terms

To control for the fact that some countries may be more prone to retirement than others for unobserved reasons not captured by our covariates, a 'shared' frailty model has been used. The terms for the 26 EU member states from our 'shared' frailty model are shown in Figure 2. Particularly, the panel on the left-hand side of the figure show the estimated frailty terms from the regression in Table 4. The right-hand side panel of Figure 2 shows results the results from the regression in Table 6. Cases above the 0 line are the most failure-prone ones.

The results, in Figure 2, provide a mixed picture with some large euro area countries lying below the zero line (i.e. Italy, Spain, the Netherlands) while others, e.g. France and Belgium, lying slightly above zero. These results confirm our previous findings, suggesting that the hazard to retirement is mixed and can not be reconciled with membership to the euro area. For the early retirement regression, the picture changes only slightly with some countries moving around the zero line (e.g. France, Italy and Denmark).

Although, there are no significant differences across regions there are clear differences across countries. On average, however, more prone to retirement countries are also those who are more prone to retire earlier.

Figure 2: Frailty terms for EU member states, retirement (left-hand side) and early retirement (right-hand side)



5. Conclusions

Schemes to curb public expenditures by increasing the minimum retirement age represents important arguments of discussion in the bargaining set up (see also Hicks, 2011). However, understanding in greater details the motivations for retirements could assist the formulation of policies that might encourage the return of retirees to employment or decrease the incentives of withdrawing earlier from the labour market.

Workers are often assumed to dace the choice of leaving the labour market based on their own preferences (Fengler, 1975; Hayward, Grady and McLaughlin, 1988) and / or based on the trade-off between market work versus home production or leisure (for an overview see Bazzoli, 1985; Blöndal and Scarpetta, 1999; Duval, 2003; Gruber and Wise, 2002; Meghir and Whitehouse, 1997). In practice, however, different constraints can influence the labour force participation decision of the elderly.

In this paper a wide set of socio-economic and environmental variables is employed to study exits into retirement in the EU. Based on longitudinal data from the Eurostat Survey on Income and Living Conditions (EU-SILC), over the period 2004-2009, we analyse the probability of retiring at a given age, given that the person has not retired yet. A number of stylized facts are documented.

First, after controlling for (un)measured risk factor affecting the hazard to retire in each country, we find no significant differences, on average, in the patterns of retirement between residents in the euro area and the

EU non-euro area countries. Second, shifts into retirement have increased during the onset of the 2009 economic and financial crisis, when controlling for income effects. Income and benefits are found to be important also as regards early retirement decisions, when accumulated income / wealth is presumably lower, compared to retiring beyond the legal retirement age. In the same vein, flexible working arrangements are found to be particularly important for workers to keep working beyond the legal retirement age. Thus, making use of partial working arrangements could modify retirement patterns towards postponing the age of withdrawing from the labour market.

Finally, this analysis shows overall that, among all the institutional measures scrutinized (state/health benefits, minimum retirement age, etc.), none could be sufficient alone if individuals withdraw from the labour market before the legal retirement age due to a weakening of their health. Particularly, for early retirees, policies aimed at improving the health of the workforce and at keeping people who experience health problems active may be essential.

All in all, by jointly testing for a wide set of factors affecting retirement decisions in the EU, the results of this paper illustrate that adequate policies to retain old workers at work can only be appropriately formulated once the determinants of retirement decisions are fully understood and modelled. In our knowledge, this paper represents a first attempt in this direction for the whole EU.

References

Aalen, O. O. (1978). Nonparametric inference for a family of counting processes, *Annals of Statistics*, 6 701-726 in Cleves M., Gutierrez R. G., Gould W., Marchenko Y. V. (2010), An Introduction to Survival Analysis using Stata, Stata Press, Texas, US.

Akaike H. (1974). A new look at the statistical model identification, *IEEE Transaction on Automatic Control*, 19:716-723.

Autor D., Dorn D. (2009). This Job Is "Getting Old": Measuring Changes in Job Opportunities Using Occupational Age Structure, *American Economic Review*, American Economic Association, vol. 99(2), 45-51.

Bazzoli G.J. (1985). The Early Retirement Decision: New Empirical Evidence the Influence of Health. *The Journal of Human Resources*, Vol. 20, No. 2. (Spring), pp. 214-234.

Blake D. (2004). The impact of wealth on consumption and retirement behaviour in the UK, *Applied Financial Economics*, 14(8), 555-576.

Blöndal, S., Scarpetta, S. (1999). The Retirement Decision in OECD Countries, OECD Economics Department Working Papers, No. 202, OECD Publishing.

Bound J. (1991), Self-Reported Versus Objective Measures of Health in Retirement Models, *Journal of Human Resources*, University of Wisconsin Press, vol. 26(1), pages 106-138.

Boskin M. J., Hurd M. D., (1978). The effect of social security on early retirement, *Journal of Public Economics*, Elsevier, vol. 10(3), pages 361-377.

Box-Steffensmeier J., Jones B., (2004). Timing and Political Change: Event History Modeling in Political Science. Ann Arbor: University of Michigan Press.

Box-Steffensmeier, J., Zorn C., (2001). Duration Models and Proportional Hazards in Political Science, *American Journal of Political Science* 45:972–988.

Box-Steffensmeier, J., Sokhey A. E., (2010). Event History Methods in K.T. Leicht and J.C. Jenkins (eds.), H*andbook of Politics: State and Society in Global Perspective*, Handbooks of Sociology and Social Research.

Cleves M., Gutierrez R. G., Gould W., Marchenko Y. V. (2010). An Introduction to Survival Analysis using Stata, Stata Press, Texas, US.

Cox D. R (1972). Regression Models and Life Tables (with Discussion), *Journal of the Royal Statistical Society*, Series B 34:187—220 in Cleves M., Gutierrez R. G., Gould W., Marchenko Y. V. (2010), An Introduction to Survival Analysis using Stata, Stata Press, Texas, US.

Cox D. R. (1975). Partial likelihood, *Biometrika* 62, 269–76.

Deschryvere M. (2005). Health and Retirement Decisions: An Update of the Literature. ENEPRI Research Report. No 6.

Diamond, P. and Hausman, J. (1984). Individual Retirement and Saving Behavior, *Journal of Public Economics*, 23: 81-114.

Disney, R.. Emmerson, C., Wakefield, M. (2006). Ill health and retirement in Britain: A panel data-based analysis, *Journal of Health Economics*, vol. 25(4), pages 621-649, July.

Duggan J. E., (1984). The Labour-Force Participation of Older Workers, *Industrial and Labour Relations Review*, 37(3), 416-430.

Duval, R. (2003). The Retirement Effects of Old-Age Pension and Early Retirement Schemes in OECD Countries. EOCD Economics Department Working Paper No. 370. Paris: OECD.

European Commission (2011). The 2012 Ageing Report: Underlying Assumptions and Projection Methodologies, European Economy 4.

Fengler, A. P. (1975). Attitudinal orientation of wives towards their husbands' retirement. *International Journal of Aging and Human Development*, 6: 139- 152.

Fields G. S., Mitchell O. S. (1984). The effects of social security reforms on retirement ages and retirement incomes, *Journal of Public Economics*, Elsevier, vol. 25(1-2), pages 143-159.

Gehan E. A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples, *Biometrika*, 52: 203-223 in Cleves M., Gutierrez R. G., Gould W., Marchenko Y. V. (2010), An Introduction to Survival Analysis using Stata, Stata Press, Texas, US.

Grambsch P. M., Therneau T. M. (1994). Proportional hazard tests and diagnostics based on weighted residuals, *Biometrika*, 81: 515-526 in Cleves M., Gutierrez R. G., Gould W., Marchenko Y. V. (2010), An Introduction to Survival Analysis using Stata, Stata Press, Texas, US.

Greenwood M., Yule G.U. (1920). An Inquiry into the Nature of Frequency Distributions of Multiple Happenings, with Particular Reference to the Occurrence of Multiple Attacks of Disease or Repeated Accident, *Journal of the Royal Statistical Society A*, Vol. 83, pp. 255-279.

Greenwood M. (1926). The natural duration of cancer, Reports on Public Health and Medical Subjects, 33, 1–26 in Cleves M., Gutierrez R. G., Gould W., Marchenko Y. V. (2010), An Introduction to Survival Analysis using Stata, Stata Press, Texas, US.

Gruber, J., Wise, D.A. (2002). Social Security Programs and Retirement around the World: Micro Estimation, NBER Working Paper No. 9407.

Hagan, R., Jones, A. M., and Rice, N. (2009). Health and Retirement in Europe. *International Journal of Environmental Research and Public Health*, 6(10), pages 2676-2695.

Hanoch G., Honig M, (1983), Retirement, Wages, and Labor Supply of the Elderly, *Journal of Labor Economics*, 1, 131-51.

Hayward, M. D., Grady, W. R., McLaughlin, S. D (1988). The retirement process among older women in the United States. *Research on Aging*, 10: 358-382.

Hicks P. (2011). The surprisingly large policy implications of changing retirement durations, SEDAP Research Paper No. 284.

Hurd M. D., (1990) , Research on the Elderly: Economic Status, Retirement, and Consumption and Saving, *Journal of Economic Literature*, American Economic Association, vol. 28(2), pages 565-637.

Hurd M. D. (1993). The Effect of Labor Market Rigidities on the Labor Force Behavior of Older Workers, NBER Working Papers 4462.

Jenkins, S. P. 2008. “Survival Analysis.” Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester.

Jones A., Rice N., Roberts J., (2008), Early retirement and inequality in Britain and Germany: How important is health?, Working Papers 2008012, The University of Sheffield, Department of Economics.

Jones A., Rice N., Roberts J., (2010). Sick of work or too sick to work? Evidence on self-reported health shocks and early retirement from the BHPS, *Economic Modelling* 27 866-880.

Kaplan E.L., Meier P. (1958). Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, 53, 457-481 in Cleves M., Gutierrez R. G., Gould W., Marchenko Y. V. (2010), An Introduction to Survival Analysis using Stata, Stata Press, Texas, US.

Kerkhofs, M., Lindeboom, M. (1995). Subjective health measures and state dependent reporting errors. *Health Economics*, 4: 221–235.

Kiefer N. M. (1988). Economic Duration Data and Hazard Functions, *Journal of Economic Literature*, Vol. 26, No. 2, 646–679.

Killingsworth M., Heckman J., (1986). Female Labor Supply: A Survey, in O. Ashenfelter: Handbook of Labor Economics. Amsterdam: North-Holland, 103-204.

Klein J. P., Moeschberger M. L. (2003). Survival Analysis: Techniques for Censored and Truncated Data, New York: Springer-Verlag in Cleves M., Gutierrez R. G., Gould W., Marchenko Y. V. (2010), An Introduction to Survival Analysis using Stata, Stata Press, Texas, US.

Kreider B. (1999). Latent Work Disability and Reporting Bias, *Journal of Human Resources*, University of Wisconsin Press, vol. 34(4), 734-769.

Lancaster T. (1979), Econometric methods for the duration of unemployment, *Econometrica*, 47: 939-956.

Lindeboom M.. (2006). Health and work of older workers in A. Jones, editor, The Elgar Companion to Health Economics, pages 26–35. Edward Elgar Publishing.

Lindeboom M., van Doorslaer E. (2004). Cut-point shift and index shift in self-reported health, *Journal of Health Economics*, vol. 23(6), 1083-1099.

Machado C. S., Portela M. (2012). Hours of Work and Retirement Behavior, IZA Discussion Papers 6270.

Mantel N., Haenszel W. (1959). Statistical aspects of the analysis of data from the retrospective analysis of disease, *Journal of the National Cancer Institute*, 22: 719-748.

Meghir C., Whitehouse E. (1997). Labour-market transitions and retirement of men in the UK, *Journal of Econometrics*, vol. 79, pp. 327-354.

Michaud P. (2003). Joint Labor Supply Dynamics of Older Couples". IZA Discussion Paper 832.

Montalto C. P. (2001). Retirement Wealth and Its Adequacy: Assessing the Impact of Changes in the Age of Eligibility for Full Social Security Benefits, Center for Retirment Research at Boston College, CRR WP 20001-07.

Nelson, W. (1972). Theory and applications of hazard plotting for censored failure data, *Technometrics*, 14, 945-965 in Cleves M., Gutierrez R. G., Gould W., Marchenko Y. V. (2010), An Introduction to Survival Analysis using Stata, Stata Press, Texas, US.

OECD (2011). Pensions at a Glance 2011: Retirement-income Systems in OECD and G20 Countries, OECD Publishing, http://dx.doi.org/10.1787/pension_glance-2011-en.

Ruhm, C. (1990). Determinants of the timing of retirement. In P. B. Doeringer (Ed.) Bridges to retirement: Older workers in a changing labor market. (pp. 23-32). Ithaca, NY: ILR Press.

Schoenfeld D. (1981). The asymptotic properites of nonparametric tests for comparing survival distributions, *Biometrika*, 68: 316-319.

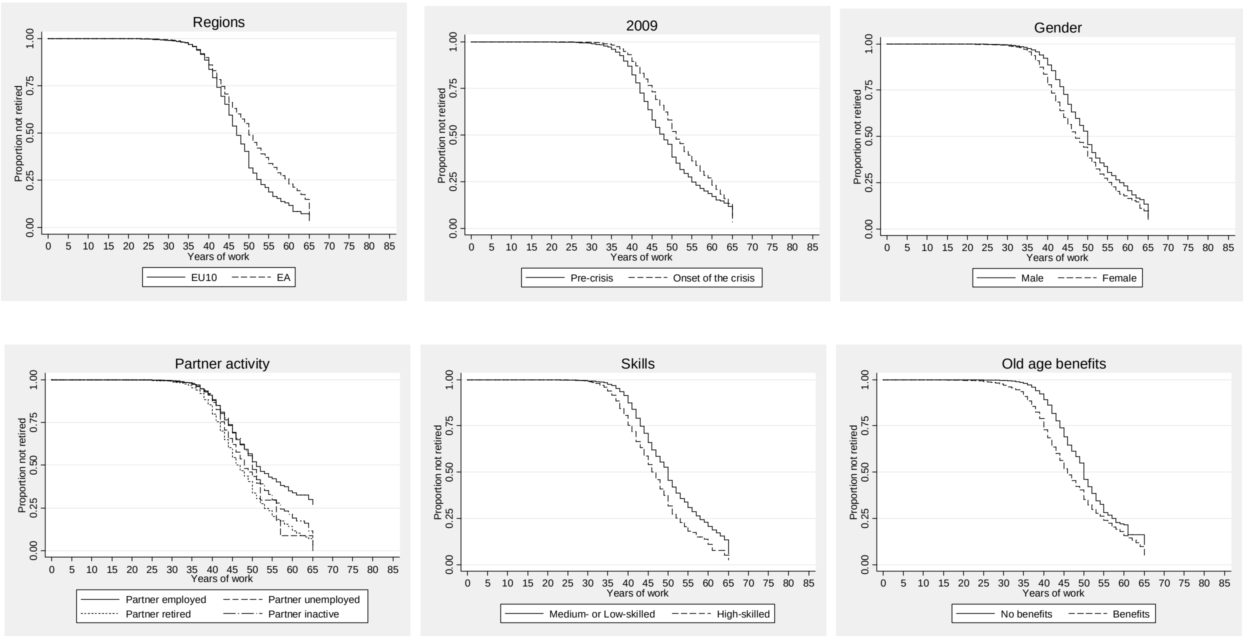
Vaupel, J.W., Manton, K.G., Stallard, E (1979). The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality, *Demography* 16, 439 – 454.

Wu S. (2003). The effect of health events on the economic status of married couples, *Journal of Human Resource* 38, 209–230.

Zorn, Christopher J. W. (2000). “Modeling Duration Dependence.” *Political Analysis* 8:367–380.

APPENDIX

Figure A1: Survival function for the transition from employment into retirement for separate groups



Note: Test results, i.e. log-rank (Mantel and Haenszel, 1959) and Wilcoxon (Breslow, 1970; Gehan, 1965), for the equality of the different survivor functions suggest that the equality of the survivor functions is rejected.

Figure A2: Survival function for the transition from employment into retirement by country

Sensitivity analysis of the model: full-parametric regressions

In order to investigate if the estimated Cox coefficients are robust, in what follows we present results from a full-parametric model.[24] As discussed in Section 3.1, when applying parametric models it is necessary to specify a certain functional form of the hazard rate that fits the data. The likelihood-ratio or Wald test can be used to discriminate between groups of nested models (Cleves *et al.*, 2010). In the present case, the results of the likelihood-ratio test indicate that the generalized gamma distribution fits well.[25] However, when models are not nested, likelihood-ratio or Wald test tests are not appropriate and an alternative statistic has to be used. The most common is the Akaike Information Criterion (AIC). Akaike (1974) proposed a method penalising each models' log-likelihood to reflect the number of parameters being estimated and then comparing them.[26]

In Table A1 an overview of the computed AIC scores is presented. There are slight differences in the value of the log-likelihood function between the models. Although the log-logistic distribution scores best, the results reveal that the Weibull model is the preferred specification in the proportional hazard form.[27] Note, also that the Weibull model has virtually the same AIC scores as the log-logistic one.[28] Furthermore, as shown in the non-parametric analysis above, the hazard of exiting into retirement exhibits a monotonically increase. Thus, based on these combined assessment, the more reasonable Weibull distribution is employed.[29] Estimates using the log-logistic distribution do not produce any relevant differences compared

---

[24] One of the assumptions underlying the Cox model is the proportional hazards assumption. Evaluating the robustness of the estimated Cox proportional hazard models, it is shown (from the results in Section 4.2) that the joint Wald test of all coefficients equal to 0 is rejected at a standard significance level in all cases. However, the test of the proportional hazards assumption using Schoenfeld's (1982) residuals is rejected. Since we are more interested in the parameter estimates than the shape of the hazard in this paper, the Cox proportional hazard model is, nevertheless, well-suited to this goal.

[25] We start from a generalized gamma model for evaluating and selecting an appropriate parametric model. We test the hypothesis that the ancillary parameters for the generalized gamma distribution (with standard deviation) *kappa* = 0 (model is log-normal); *kappa* = 1 (model is Weibull); and *kappa* = 1 and *sigma* = 1 (model is exponential). By testing the appropriate restrictions, it is found that we can reject the log-normal, the Weibull and the exponential distribution against the gamma for all samples.

[26] The AIC compares the likelihood scores while taking into account the degrees of freedom used in each model. AIC = -2* log-likelihood + 2 * (k + c), where k is the number of model covariates and c the number of model-specific distributional parameters.

[27] Since the Weibull can be specified in both the proportional hazard and accelerated failure time form we can compare it to other accelerated failure time distributional forms.

[28] This is the case also for the generalized gamma distribution.

[29] The Weibull model assumes a baseline hazard of the form $h_0(t) = pt^{p-1} exp(\beta_0)$, where p is some ancillary shape parameter estimated from the data, and the scale parameter is parameterised as $exp(\beta_0)$. The Weibull distribution can provide a variety of monotonically increasing or decreasing shapes of the hazard function, and their shape is determined by the estimated parameter p.

to the Weibull estimates, suggesting that the selection effect of using distributions other than the Weibull is limited.

Table A1: Model selection for full parametric regressions

Distribution (metric)	Log likelihood	k	c	AIC	Ranking
	Whole sample (retirement)				
Exponential (PH, AFT)	-10001	23	1	20049	6
Weibull (PH, AFT)	-2275	23	2	4599	3
Gompertz (PH)	-2544	23	2	5138	4
Log-normal (AFT)	-3445	23	2	6941	5
Log-logistic (AFT)	-2226	23	2	4503	1
Generalized gamma (AFT)	-2255	23	3	4563	2

Note: The models are estimated assuming gamma distributed frailty or heterogeneity.

In Table A2 the time ratios from the estimated – Weibull distributed – accelerated failure time model are presented.

Table A2: Full parametric regressions

Variable	Time ratio	(Std. Err.)	Variable	Hazard ratio	(Std. Err.)
EA	1.001	(0.037)	*Income variables*		
Dummy 2009	0.971***	(0.004)	Disposable income	1.001	(0.003)
Minimum retirement age	1.012***	(0.002)	Old age benefits	0.925***	(0.004)
			Unemployment benefits	0.926***	(0.007)
Individual characteristics.			Disability benefits	0.880***	(0.008)
Female	0.955***	(0.006)	Sickness benefits	0.988	(0.009)
Married	0.996	(0.009)			
Skilled	0.944***	(0.005)	*Interaction*		
Part-time	1.319***	(0.075)	Part-time x disp. income	0.979***	(0.006)
Occupational group			*Health variables*		
2.Occ. group	1.057***	(0.006)	Health	0.947***	(0.013)
3.Occ. group	1.002	(0.006)	Health(-1)	0.961**	(0.017)
4.Occ. group	1.006	(0.007)			
5.Occ. group	1.014*	(0.007)	*Partner characteristics*		
			2.Partner unemployed	0.964***	(0.011)
			3.Partner retired	0.990**	(0.004)
			4.Partner inactive	0.979***	(0.005)
			Partner's health	1.047***	(0.014)
			Constant	23.550***	(2.559)
Statistics					
ln(p)	8.521***	(0.084)			
ln(θ)	0.589**	(0.153)			
LR test (frailty terms)	Prob>=chi-bar-sq. = 0.000				
Wald χ2	1215.992				
Prob > χ2	0.000				
Log-likelihood	-1407.851				
Number of groups	26				
Observations	53,490				

Note: Standard errors in parentheses. *** denotes p<0.01, ** p<0.05, * p<0.1. See Table A3 for data and definitions

The results of the Weibull model are basically consistent with those of the Cox proportional hazard regression in Table 4. An important difference to bear in mind when interpreting the results is that in proportional hazard models (such as Cox's) the estimates are interpreted as the effect on the employment exit rate; while accelerated failure time models analyse the effect on the employment period.[30]

[30] Additionally, it should be noted from Table A2 that the Wald test on the ancillary shape parameter (p) indicates that we can reject the hypothesis that the hazard is a constant, suggesting a monotone increasing behaviour over time. The hypothesis that $ln(p)=0$ is rejected at the 1% significance level for all observations. The parameter p is the 'shape' parameter, as it defines the shape of the distribution. If $p= 1$, then the hazard is constant. For other values of p, the Weibull hazard is not constant; it is monotone decreasing when $p < 1$ and monotone increasing when $p > 1$.

Table A3: Variables used in the estimation

Variable	*Definition*
<u>Dependant variables</u>	
Employment duration	The amount of time that an individual spends in employment before entering into retirement.
Employment duration, early	The amount of time that an individual spends in employment before entering into early retirement (before the legal retirement age).
<u>Explanatory variables</u>	
Activity	Main activity status during the income reference period. If the main activity is not 'a job or business', the status is self-defined. The main activity status during the income reference period is 'at work' if the respondent worked (or was in paid apprenticeship or training) the majority of weeks during the income reference period. If a person spends the same number of weeks in different activities, priority should be given to economic activity ('main activity job or business') over non-economic activity and over inactivity.
Employed	Equals 1 if the individual is at work. A person is at work if he works at least 1 hour during the reference week.
Unemployed	Equals 2 if the individual is unemployed
Retired	Equals 3 if the individual is in retirement or early retirement
Inactive	Equals 4 if the individual classifies himself as any other inactive person.
Change activity	Most recent change in the individual's activity status. The variable records changes in the individual activity status over the last interview (or last 12 months for the first year of data collection).
Employed - retired	Equals 1 if the individual changed from employment to retirement.
Unemployed - retired	Equals 2 if the individual changed from unemployment to retirement.
Retired - employed	Equals 3 if the individual changed from retirement to employment.
Retired - unemployed	Equals 4 if the individual changed from retirement to unemployment.
Retired - inactive	Equals 5 if the individual changed from retirement to inactive other than retirement.
Inactive - retired	Equals 6 if the individual changed from inactivity other than retirement to retirement.
EA	Equals 1 if a country belongs to the euro area.
Dummy 2009	Equals 1 if year of the survey equals 2009.
Minimum ret. age	Countries' minimum retirement age according to OECD (2011).
Female	Equals1 if the interviewed is of female gender.
Married	Equals 1 if the interviewed is married.
Skilled	Equals1 if the interviewed has high education according to the highest ISCED level attained. This includes first stage of tertiary education (not leading directly to an advanced research qualification) and second stage of tertiary education (leading to an advanced research qualification.
Occ. Groups	The variable conforms to the ISCO-88 (COM) International Standard Classification of Occupations.

	1	Equals 1 if the individual belongs to legislators, senior officials and managers, professionals, technicians and associate professionals or clerks.
	2	Equals 2 if the individual belongs to service workers and shop and market sales workers, skilled agricultural and fishery workers.
	3	Equals 3 if the individual belongs to craft and related trades workers.
	4	Equals 4 if the individual belongs to plant and machine operators and assemblers.
	5	Equals 5 if the individual has a elementary occupation.
Part-time		Equals 1 if the individual works or worked part-time based on a self-defined economic status.
Disposable income		(Log) total disposable household income. This includes the sum for all household members of gross personal income components (gross employee cash or near cash income; gross non-cash employee income; company car; employers' social insurance contributions; gross cash benefits or losses from self-employment (including royalties); value of goods produced for own consumption; pensions received from individual private plans; unemployment benefits; old-age benefits; survivor' benefits, sickness benefits; disability benefits and education-related allowances plus gross income components at household level (imputed rent; income from rental of a property or land; family/children related allowances; social exclusion not elsewhere classified; housing allowances; regular inter-household cash transfers received; interests, dividends, profit from capital investments in unincorporated business; income received by people aged under 16) minus (employer's social insurance contributions interest paid on mortgage; regular taxes on wealth; regular inter-household cash transfer paid; tax on income and social insurance contributions).
Part-time X disp. income		Interaction term between part-time and disposable income.
Old age benefits		Equals 1 if the individual receives non-zero old age benefits. By definition, the old age function refers to the provision of social protection against the risk linked to old age, loss of income, inadequate income, lack of independence in carrying out daily tasks, reduced participation in social life, and so on. Old age benefits cover benefits that: provide a replacement income when the aged person retires from the labour market, or guarantee a certain income when a person has reached a prescribed age.
Disability benefits		Equals 1 if the individual receives non-zero disability benefits. Disability benefits refer to benefits that provide an income to persons below standard retirement age whose ability to work and earn is impaired beyond a minimum level laid down by legislation by a physical or mental disability. Disability is the full or partial inability to engage in economic activity or to lead a normal life due to a physical or mental impairment that is likely to be either permanent or to persist beyond a minimum prescribed period.
Sickness benefits		Equals 1 if the individual receives non-zero sickness benefits. Sickness benefits refer to cash benefits that replace in whole or in part loss of earnings during temporary inability to work due to sickness or injury. Being temporary in nature, those include only paid leave or cash benefits in case of self-reported sickness or injury or that of a dependent child.
Unempl. benefits		Equals 1 if the individual receives non-zero unemployment benefits. Unemployment benefits refer to benefits that replace in whole or in part income lost by a worker due to the loss of gainful employment; provide a subsistence (or better) income to persons entering or re-entering the labour market; compensate for the loss of earnings due to partial unemployment; replace in whole or in part income lost by an older worker who retires from gainful employment before the legal retirement age because of job reductions for economic reasons; contribute to the cost of training or re-training people looking for employment; or help unemployed persons meet the cost of travelling or relocating to obtain employment.
Health		Equals 1 if the individual assesses his health is 'very bad'. The measurement of self-perceived health is, by its very nature,

	subjective. The notion is restricted to an assessment coming from the individual and not from anyone outside that individual. The reference is to health in general rather than the present state of health, as the question is not intended to measure temporary health problems. It is expected to include the different dimensions of health, i.e. physical, social and emotional function and biomedical signs and symptoms. It omits any reference to an age as respondents are not specifically asked to compare their health with others of the same age or with their own previous or future health state.
Partner's activity	Main activity status of the partner (if any) during the income reference period. See *Activity* definition
Employed	Equals 1 if the partner is at work.
Unemployed	Equals 2 if the partner is unemployed
Retired	Equals 3 if the partner is in retirement or early retirement.
Inactive	Equals 4 if the partner is inactive.
Partner's health	Equals 1 if each individual's partner (if any) assesses his health to be 'very bad'. See *Health* definition.
<u>Alternative health measures</u>	
Limit	Equals 1 if the individual reports limitations in activities because of health problems. The purpose of the instrument is to measure the presence of long-standing limitations, as the consequences of these limitations (e.g. care, dependency) are more serious. The period of at least the last 6 months is relating to the duration of the activity limitation and not of the health problem. The answer to this question is yes (1 or 2) if the person is currently limited and has been limited in activities for at least the last 6 months.
Chronic	Equals 1 if the individual reports to suffer from any a chronic (long-standing) illness or condition.

Note: See also the EU-SILC's Guidelines.