Brunel University London

Department of
Economics and Finance

Economics and Finance Working Paper Series

Working Paper No. 16-17

Ted Aranki and Corrado Macchiarelli

The Impact of Income and Health on the Economic Decision to Retire (Early)

October 2016

http://www.brunel.ac.uk/economics

The Impact of Income and Health on the Economic Decision to Retire (Early)

Ted Aranki[a], Corrado Macchiarelli[b,c]

[a]*Sveriges Riksbank*
[b]*Brunel University London*
[c]*London School of Economics and Political Science*

Abstract

The economic decision to cease working is traditionally based on standard life-cycle models, where individuals trade-off between market work versus home production or leisure. In this paper, we study such transitions out of work for 26 EU countries over the period 2004-2011 in order to investigate the determinants of retirement: income and health, in particular. Controlling for these factors, we do not find any significant differences in the retirement pattern 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 2008-09 economic and financial crisis. Income, together with flexible working arrangements, as well as private pension plans, are found to be important as regards early retirement decisions, compared to retiring on (or beyond) the legal retirement age. Concerning health measures, we show that institutional factors (such as, state / health benefits, minimum retirement age) could not be sufficient alone if individuals withdraw earlier from the labor market due to a weakening of their health. Especially, these latter results can be of importance for structural and macroeconomic policy, for instance, in increasing the employment of both people and hours worked against the background of population aging.

Keywords: Retirement decision; early retirement decision; frailty terms; euro area, EU10

This version: Sept. 2015

Email address: Corrado.Macchiarelli@brunel.ac.uk (Corrado Macchiarelli)

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.[1] 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.

In standard life-cycle models, 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; Blndal 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. Financial considerations are often found to be one of the most important determinants of (early) retirement (see also Cahuc and Zylberberg, 2004; Montalto, 2011). Moreover, personal health or the need to provide

[1] Certainly, 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).

care to a household member can equally be an important reason to retire, especially before the legal retirement age (see *inter alia*, Bound, 1991; Bound *et al.* 1999; Jones *et al.*, 2008; 2010; Deschryvere, 2005; Disney *et al.*, 2006).

In this paper we employ a wide set of socio-economic and environmental variables to study exits into retirement in the EU, mainly based on the determinants above. Based on longitudinal data from the Eurostat Survey on Income and Living Conditions (EU-SILC) – covering the period 2004-2011 – we analyse the probability of retiring at a given age, given that the person has not retired yet (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 results for a large set of 26 EU countries, by providing a systematic, conditional approach on income and health variables to estimate labour market work-to-(early) retirement transitions. Secondly, we exploit cross country differences, 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 changes over time – especially after 2008-09, reflecting the extent to which the global economic and financial crisis hit in most countries – and across euro area versus EU non-euro area countries – reflecting region-specific dynamics and institutional set ups – we pay particular attention to those dynamics. Nonetheless, the results in this paper do not support any significant differences in the employment-to-retirement patterns between the average euro area and EU non-euro area countries. However, shifts into retirement seem to have increased during the onset of the crisis, particularly when controlling for income.

The main results also suggest that income variables, such as benefits (typically temporary in nature, e.g., sickness benefits), as well as the availability of private pension plans, are found to be important as regards early retirement decisions, 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.

In line with the literature, our findings also point to the fact that *health* is an important determinant of retirement in the EU, 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).[2] Moreover, individuals with partners in bad health are generally associated with a lower likelihood to retire compared to individuals with partners in better health status (see also Wu, 2003).

All in all, we conclude that institutional factors (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 may 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 theory behind the life-cycle model. Section 3 summarizes the EU-SILC data utilized. Section 4 and 5 present the econometric approach and the main results, respectively. Section 6 concludes.

2. The economic decision to retire

The economic decision by which a person ends the labor market participation fits well into standard life-cycle models in a neoclassical sense (see Cahuc and Zylberberg, 2004).

In a standard fashion, let us consider a person employed in year τ (i.e. equal to his / her age) and let us suppose that that person is willing to retire on a given future date $T_s \geqslant \tau$. In a classic model of labor supply, the evolution of the representative worker's wealth, starting from date τ, can be described as a dynamic decision. Here, it is normally assumed that the individual has a given endowment and we denote r_t as the real interest rate between periods $t-1$ and t. In each period, the available time (i.e. be it a day, a month, or a year) is normalized to unity to simplify the notation. Henceforth, the hours worked during time t are given by $(1 - L_t)$.

Defining A_t as the consumer's wealth (total disposable income, including assets) and B_t as the non-wage income for a given initial value of accumulated

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

income, A_0, allows us to describe the evolution of wealth of the consumer as:

$$A_t = (1 + r_t)A_{t-1} + B_t + w_t(1 - L_t) - C_t$$

where $A_t - (1 + r_t)A_{t-1}$ is the evolution of the consumer's wealth, B_t and $w_t(1 - L_t)$ are non-wage and wage income, respectively, and C_t is consumption. To simplify the problem, in these models it is normal to assume that the agent does not work after date T_s, thus the economic decision to retire becomes a discrete one. Hence, for $t \geqslant T_s$, leisure time will be $L_t = 1$.[3]

Following a standard approach (Cahuc and Zylberberg, 2004), we denote $B_t(T_s)$ as the income expected in period $t \geqslant T_s$, composed of pension payments over the period t and other sources of income. Equally, $B_t(t)$ will designate the non-wage income of the consumer while he / she is not working and denote C_{et} and C_{rt}, respectively, as the consumption of physical good during employment (e) and retirement (r).

For a given retirement age, T_s, the representative agent will solve the following problem (Cahuc and Zylberberg, 2004):

$$Max_{C_{et},C_{rt},L_t,T_s,t} \left[\sum_{t=\tau}^{T_s-1} U(C_{et}, L_t, t) + \sum_{t=T_s}^{T} U(C_{rt}, 1, t) \right]$$

subject to the constraints:

$$A_t = \begin{cases} (1 + r_t)A_{t-1} + B_t(0) + w_t(1 - L_t) - C_{et} & \text{if } \tau < t < T_s - 1 \\ (1 + r_t)A_{t-1} + B_t(T_s) - C_{rt} & \text{if } T_s \leqslant t \leqslant T \end{cases}$$

As we will show in the next section, in continuous time, T_s can be seen through the lenses of an event study.

2.1. Intrafamilial decisions

With respect to the problem set out above, the family, and the partner in particular, has considerable influence on the decision to retire. In the empirical approach that follows, we take into account the influence of the family structure. This is particularly relevant, as it give us ground for a framework going behind the *individual* decision to retire.

In order to extend the basic life-cycle model, we follow Cahuc and Zylberberg (2004) and consider a family composed of two individuals. Hence,

[3] Certainly, the process to retire can be gradual with people moving to part-time working schemes before retiring completely.

we assume that the preferences of family are represented by the following utility function:

$$U(C_{e1t}, C_{r1t}, C_{e2t}, C_{r2t}, L_{1t}, L_{2t})$$

where C_{ei} and C_{ri} represent the consumption during employment (e) and retirement (r) of physical goods and L_i (with $i = 1, 2$) denotes the leisure of individual i. This formalization, generally known as the "collective approach", assumes that decision making is something individuals do within the family (see also Killingsworth and Heckman, 1986; Hurd, 1990; Michaud, 2003).

As we shall discuss in our data section, this framework is particularly useful given the construction of our dataset.

A suitable model to model all intra-family decisions is the collective model presented by Chiappori (1988, 1992). In particular, the latter (Chiappori, 1992, *proposition 1*) shows that the efficient solution of the individual preferences of the people composing the household is also the solution to individual optimization problem in which each person would be endowed with a specific non-wage income, which will depend on the overall household disposable income / wealth (see also Cahuc and Zylberberg, 2004).
More precisely, by simply using Chappori's (1992) intuition above, the problem of agent 1 – or the agent taking the decision to retire, specifically – would take the form:

$$Max_{C_{e1t}, C_{r1t}, L_{1t}, T_s, t} U(C_{e1t}, C_{r1t}, L_{1t})$$

subject to the constraints:

$$\Phi_t = \begin{cases} (1 + r_t)A_{t-1} + B_{1t}(0) + w_{1t}(1 - L_{1t}) - C_{e1t} & \text{if } \tau < t < T_s - 1 \\ (1 + r_t)A_{t-1} + B_{1t}(T_s) - C_{r1t} & \text{if } T_s \leqslant t \leqslant T \end{cases}$$

where $\Phi_t = \Phi(A_t, w_{it}, B_{it})$ would have the typical interpretation of Chiappori's (1992) "sharing rule", depending, as before, on the household's wealth A_t, plus the parameters w_i (for $\tau < t < T_s - 1$) and B_i (for $t \leqslant T$) for i=2. In this way, w_2 and $B2$ will be exogenous to consumer's 1 decision. Hence, one can write down the welfare optimization problem of consumer 1 in the standard way (see Cahuc and Zylberberg, 2004). By designating the value of the welfare of consumer 1 at the optimum of the problem as $V_\tau(T_s)$ and denoting the legal age of retirement T_m, as the institutional age after which it will not be possible to work any more, agent 1 (aged τ) will choose the date T_s on which he / she will end his or her work pattern by solving the following problem:

$$Max_s V_\tau(T_s) \text{ subject to } T_m \geqslant T_s \geqslant \tau$$

This simple model does not consider explicitly the possibility of working beyond the legal retirement age (i.e. $T_m \leqslant T_s$). However, this is a strong theoretical assumption, which we will not enforce when estimating our econometric model.

The problem above has a difficult solution and it is normally solved numerically (see Cahuc and Zylberberg, 2004). In order to simplify the explanation, in the next sections we will rest on this baseline problem, trying to empirically measure what, beyond the personal characteristics, may affect the decision to retire. For this purpose, we will show that a hazard duration model can be quite helpful.

3. Data

In this paper we use the Eurostat Survey on Income and Living Conditions (EU-SILC) to model a simple model of retirement. The EU-SILC consists of a database available in yearly frequencies, broadly 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 2004-2011. The sample excludes Germany owing to data unavailability.[4] 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, including wage and non-wage income (benefits, cash transfers), interest, dividends, profit from capital investment, hence being particularly suited to test a simple model of retirement as detailed in Section 2. This represents an advantage compared to other datasets, given that income composition, and wealth at large, is indeed an important determinants of retirement decisions (see for instance, Hanoch and Honig, 1983; Mitchel and Fields, 1984; Dugan, 1984; Ruhm, 1990).[5]

Consistent with life-cycle theory, an individual's decision to retire, or the

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

[5]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.

employment-to-retirement transition, is the event of interest in our study. From the EU-SILC, we construct transitions from employment to retirement, or the probability 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.

To analyze employment-to-retirement transitions, we employ a hazard based duration model (see, Diamond and Hausman, 1984; Hagan, Jones and Rice, 2009; Jones, Rice and Roberts, 2010). This allows modeling 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) in a discrete fashion.[6]

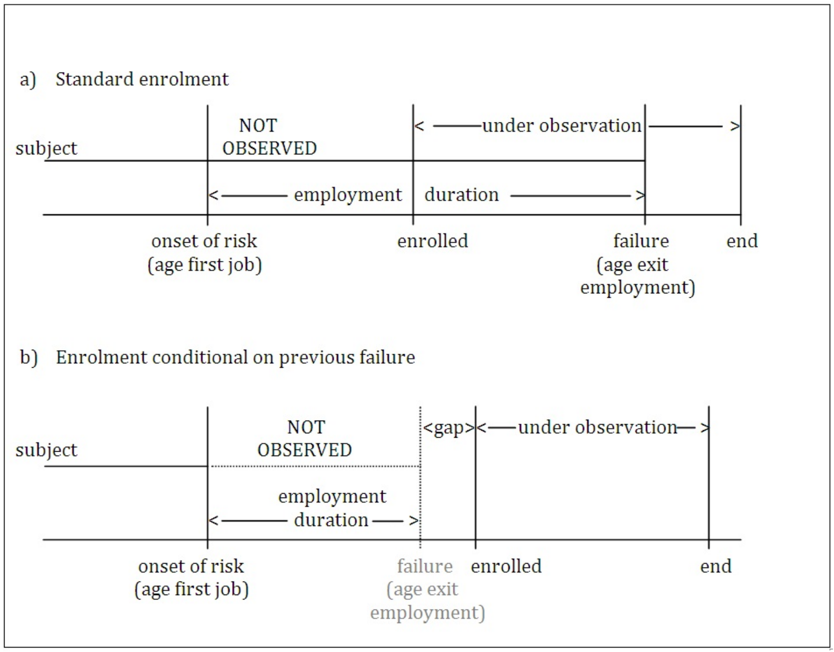
One statistical motivation for employing a duration analysis framework includes first of all the possibility of bringing to the data standard economic approaches, such as standard life-cycle models, where the decision to retire is indeed modeled discretely, and it is assumed to be influenced by a whole range of (wage and non-wage) income components, both individually and at the household level.

In addition, duration analysis is particularly useful given the presence of censored and left-truncated data in our dataset. In practice, not all observations' full history is observed until the 'failure' event (i.e. the decision to retire). 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 enroll them in the study (see Figure 1).

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 working 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, consistent with the discussion in Section 2.

[6] In Appendix A we detail the variables used in the estimation.

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



In the same vein, individuals can enter the observation period / being enrolled in the study upon having retired already. Here, an important difference of our study, 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 enrollment occurs, conditional on a previous retirement event (see Figure 1(b)), there may exist a positive difference between employment duration and the year of enrollment in the study, representing a *gap* of information about each individual's activity between the period he / she ceased to work and the period he / she became under observation. Only those with a *gap* $< |\text{G}|$, 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 $0 < gap < |G|$.[7]

As individuals' working histories are inferred based on retrospective questions, only the last spell is 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 more than 209.000 individuals. Out of this number, 6512 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 38% of the sample.

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	1.4	0.8	1.2
Age 45 to 49	2.6	2.0	2.4
Age 50 to 54	5.2	5.1	5.2
Age 55 to 59	24.8	41	31
Age 60 to 65	47.5	37.3	43.6
Age > 66	17.5	12.9	15.8
Total obs.	4,044	2,468	6,512

Gender representation is reversed for ages between 55 and 59, where women represent the majority of the sample considered. In general, female

[7]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.

workers retire earlier than males. Nearly 50% of female workers enter into retirement before the age of 60, compared to about 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	6512
25%	57	22	Sum of Wgt.	6512
50%	60		Mean	60.443
		Largest	Std. Dev.	6.442
75%	64	80		
90%	67	80	Variance	41.494
95%	71	80	Skewness	-0.656
99%	78	80	Kurtosis	7.322

4. Empirical methodology

Consistent with the discussion in the previous sections, the main advantage of using duration analysis in this framework is that it allows modeling 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), it makes sense to conceptualize the length of each individual j's employment spell as a random variable, T_{sj}.[8]

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

As in Section 2, we denote T_s as the actual retirement age, which is, in our framework, equal to the duration of the individual's employment spell.

Assuming T_{sj} has a continuous probability distribution $f(t)$, where t is a realization of T_{sj}, the cumulative distribution function of T_s will be given by:

$$F(t) = \Pr(T_{sj} \leq t) = \int_0^t f(s)ds$$

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

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

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:

$$h(t) = P(t < T_{sj} < t + dt | T_{sj} > t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} \tag{2}$$

The class of models above 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 'avoid[ing] having to make assumptions about the nature of the duration times in the first place' (Box-Steffensmeier and Jones, 2004). In other words, the Cox model makes no assumption about the shape of the hazard function or about how covariates may affect this shape.[9] Thus, the Cox semi-parametric approach is regarded as a benchmark in this paper, whereas non-parametric (Section 5.1) and fully parametric approaches are employed for preliminary investigation of the data and robustness checks, respectively. In particular, a fully

overview.

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

parametric regression corroborates our findings. The results of the latter are delegated to the Appendix C for sake of brevity.[10]

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

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

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.[11] 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.

4.1. 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.[12] Hence, fitting a normal duration model, for instance equation (3) would simply not recognize 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

---

[10] 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.

[11] 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).

[12] 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.

deflated standard errors (see Box-Steffensmeier and Zorn, 2001; Zorn, 2000). An alternative method is 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. This also allows us to cope with the idea that the decision to retire happens in an environment the characteristics of which are richer than the ones we can possibly observe (Section 2). 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):[13]

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

where v_j 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 to an unobserved parameter θ, which is to be estimated.[14] 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:[15]

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

[13]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.

[14]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.

[15]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.

5. Results

5.1. Non-parametric analysis

In order to summarize the data and visualize the shape of employment duration for the sample or for separate consumers' groups, we first introduce non-parametric estimation of the survivor and hazard functions relying on product-limit estimators.[16] Table 3 reports the survivor and cumulative hazard function for employment 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 44 years of work the probability of remaining in employment is around 0.63, indicating that roughly 37% of the sampled individuals where retired. Furthermore, the Nelson-Aalen cumulative hazard suggests that the hazard of exiting into retirement increases monotonically.[17] Survivor functions from employment to retirement across different categories, as well as by country, are plotted in Figure A1 and A2 in Appendix B.

5.2. Semi-parametric analysis

In this section, estimates of the semi-parametric Cox proportional hazard models are presented. As discussed in Section 4, a parametric analysis offers an advantage over the non-parametric methods, as it allows predicting the

[16]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).

[17]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).

probability distribution of retirement by means of a set of additional covariates. In what follows the joint effect of various individual and labor market characteristics affecting the probability of exiting into retirement is analyzed.

In Table 4, the estimated results from the Cox's proportional hazard model are reported.[18] The reported coefficients are hazard ratios.[19] As explained in Section 4, a 'shared' frailty model is employed where the subgroups are selected according to the number of countries in our sample (i.e. 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. Based on our discussion in Section 2, our income variable is indeed a pooled one and it is composed of wage and non-wage income (including public / private pension plans, age and state benefits) as well interest, dividends and profit from capital investment (see Appendix A for definitions).

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 and following years significantly increased flows 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 after 2009. 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 after 2009, with income to cover basic expenses in retirement running short because of the financial crisis.[20]

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

[19] 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.

[20] In the present context it is, however, difficult to distinguish labor market quits from

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.

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 a very important role in reducing the hazard to retire. This is in line with the empirical 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 labor 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 Appendix A), 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 standard retirement decisions in our sam-

lay-offs.

ple. Nevertheless, the marginal effect or the interaction between working part-time and disposable income (*part-time x disp. income*) is significant in reducing employment duration. This result suggests that there exist a marginal effect of income when employment is not full-time, or, the marginal exposure to the current (or past) level of income is higher when not working full-time (see also Blake, 2007; Montalto, 2001).

In addition, the availability of private pension plans (*private pension*), does play a significant effect, with the coefficient's magnitude making 16% more likely the decision to retire.

Along the same lines, other sources of income supports are very likely to influence the labor supply of the aged as well; in line with the evidence that unemployment schemes or temporary benefits, at large, may induce older workers to seek part-time jobs or to withdraw earlier from the labor market.[21] These overall results support the idea that workers generally tend to react to financial incentives that accompany early retirements (Cahuc and Zylberberg, 2004).

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 (for an overall discussion see also Macchiarelli and Ward-Warmedinger, 2015).[22] Consistently, sickness benefits, represent-

[21]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 labor force participation and withdraw earlier from the labour market.

[22]Note that, in the EU-SILC, unemployment benefits also include 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; and early retirement for labor 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 (see also Appendix

ing 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).[23] 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 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 labor market (Kerkhofs and Lindeboom, 1995; Kerkhofs *et al.*, 1999; Kreider, 1999). Finally, for some individuals there may be incentives to report health problems as a mean to obtain disability benefits (i.e. the so-called 'disability' route into retirement, see Blundell *et al.*, 2002; Burkhauser, Butler and Kim; 1995).

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 every time individuals receives disability benefits, as discussed. This allow us to control for possible 'disability routes' into retirement (Blundell *et al.*, 2002). Further, to assess the robustness of previous findings, alternative health measures are employed, along with the usual set of covariates. More specifically, in Table 5 alternative measures, some of which arguably less prone to reporting

---

A). 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.

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

bias than self-reported health are employed, such as a measure of health limitations (*limit*) and *chronic* diseases (see Jones *et al.*, 2008; 2010).

The former (*limit*) has generally a size effect on the coefficients of interest while it does not affect their sign and / or significance. This is less the case for the other proxy (*chronic*), which captures indeed chronic diseases at large, which may not affect directly the ability to work. Based on reports of bad health or a limited ability to work, health status seems indeed to be an important determinant of retirement decisions.[24]

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 Bound *et al.* (1999) and 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).[25]

Occupation statuses and health effects are important also as regard to individuals' partners. For instance, predictions regarding a joint labor market decision of old couples can derive from a family labor 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 does not confirm the prediction that having a *partner retired* significantly increases the hazard to retire, compared to having a partner employed. The latter idea is in line with the fact that the

[24]In addition, it should be noted that the the model where a measure of *chronic* disease is used is never preferred, based on the log-likelihood, to models where the alternative *health* or *limit* are employed.

[25]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 marginally 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.

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). This does not seem to be a strong factor in the EU or at least given the data sample (see Aranki and Macchiarelli, 2013). In particular, comparisons of these and our next results suggest that the partner activity is rather important for people retiring before the legal retirement age, when controlling for other factors (see next Section). 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).

To assess the overall robustness of our results, the Appendix A.3. replicates the same analysis using a fully parametric regression, with the full set of controls used in Table 4. The coefficients of interest are virtually unchanged when a fully parametric set up is used. The fact that controlling for unobservable (frailty) factors delivers the same set of results also in a fully parametric framework suggests that the estimates are likely to be sensitive to unobserved heterogeneity. In fact, using a rich set of covariates and specifying a functional form – as in the fully-parametric set-up – should reduce the significance of unobserved factors (under the assumption that the model is well specified; see Appendix A.3.). The significance of unobserved heterogeneity is confirmed in our model by the LR test at the bottom of Table A2., comparing the likelihood of the model with frailty terms with the model where θ is set equal to zero).

5.3. Early-retirement decisions

Until now the analysis has focused on individuals retiring. The analysis provided hitherto is indeed a useful benchmark for the analysis that follows. Understanding the motivations to retire *earlier* (before the legal retirement age), compared to standard retirement patterns, represents indeed 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.

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.

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, at least until (or beyond) the legal retirement age.

Higher disposable household income and state / health benefits – including those temporary in nature, such as sickness benefits – as well as the availability of private pensions plans 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 and private pension plans; which are now strongly significant) compared to retiring on or 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 (marginally) on early retirement decisions.

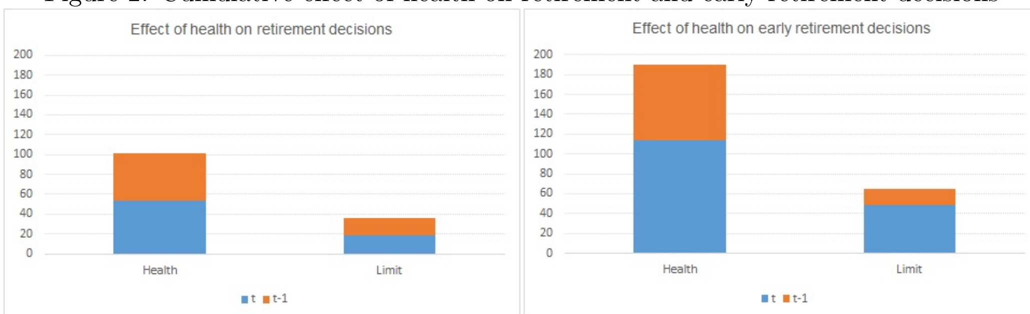
Finally, focusing on the individual health status, we show that – analogously to the findings in Table 4 – people with health problems are generally found to discontinue employment and retire earlier. The health coefficient for people retiring below the legal retirement age is twice as much the one reported in Table 4. In particular, for early retirees the hazard to retire based on bad health is 114% higher (or about 48% if using *limit* as a bad health proxy instead),[26] compared to about 53% for standard retirement decisions on or beyond the legal retirement age (or about 18% if using *limit* as a bad health proxy instead). Based on estimates of the hazard rates, the cumulative effect of health and its lag, also including alternative health measures, is even more severe when it comes to early retirement decisions (see Figure 3).

This seems to suggest 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 labor 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

[26]The results using alternative health measures for early retirement decisions are omitted for sake of brevity, but they are available upon request from the authors.

essential.

Figure 2: Cumulative effect of health on retirement and early retirement decisions



5.4. Frailty terms

As the result of employing a 'shared' frailty model, frailty terms are available from the estimation and presented in this section. As discussed, frailty terms control for the fact that some countries may be more prone to retirement than others for unobserved reasons not captured by our covariates (see Section 4.1). 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 3 shows results the results from the regression in Table 6. Cases above the 0 line are the most failure-prone ones.

Figure 3: Frailty terms for EU member states, retirement (left-hand side) and early retirement (right-hand side)

The results, in Figure 3, 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 enter retirement is mixed and cannot be reconciled with membership to the euro area or not. For the early retirement regression, the overall picture changes only slightly with some countries moving around the zero line (e.g. France, Italy and Denmark). The countries that move the most are Greece and Romania, denoting a much higher early-retirement tendency, compared to retiring on (or beyond) the legal retirement age. Sweden and France move in the opposite direction, denoting a lower early retirement tendency, compared to retirement decisions on (or beyond) the legal age.

Although there are no significant differences across regions (EA vs. EU10) there are clear differences across countries. With the due exceptions (Greece, Romania - on the one hand - and France and Sweden - on the other hand) discussed before, however, more prone to retirement countries are also those who are more prone to retire earlier on average.

6. 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 (early) 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; Blndal and Scarpetta, 1999; Duval, 2003; Gruber and Wise, 2002; Meghir and Whitehouse, 1997). In practice, however, several constraints can influence the labour force participation decision of the elderly.

In this paper we employ a wide set of socio-economic and environmental variables to study exits into retirement in the EU; income and health, in particular.

Based on longitudinal data from the Eurostat Survey on Income and Living Conditions (EU-SILC), over the period 2004-2011, we analyse the likelihood 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 in retirement have increased after the onset of the 2008-09 economic and financial crisis, when controlling for income effects. Income, private pension plans and temporary benefits are found to be important as regards early retirement decisions, when accumulated income / wealth is presumably lower, compared to retiring on or beyond the legal retirement age. In the same vein, flexible working arrangements are found to be particularly important for workers to keep working just until or beyond the legal retirement age.

Finally, concerning health measures, 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, bad health increases the hazard to retire by twice as much, compared to standard retirement decisions. Thus, policies aimed at improving the health of the workforce and at keeping people who experience health problems active may be essential.

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.

Aranki T., Macchiarelli C. (2013), Employment Duration and Shifts into Retirement in the EU, ECB Working Paper 1517.

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.

Blndal, 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.

Bound, J., Schoenbaum, M., Stinebrickner, T.R., Waidmann T., The dynamic effects of health on the labor force transitions of older workers, *Labour Economics*, 6(2), 179202.

Boskin M. J., Hurd M. D., (1978). The effect of social security on early retirement, *Journal of Public Economics*, 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.), *Handbook of Politics: State and Society in Global Perspective*, Handbooks of Sociology and Social Research.

Burkhauser, R.V., Butler, J.S., Kim, Y.W. (1995), The importance of employer accommodation on the job duration of workers with disabilities: A hazard model approach, *Labour Economics*, 2(2), 109130.

Cahuc P., Zylberberg A. (2004), Labor Economics, MIT Press, Chpt 1.

Chiappori, P.A. (1988), Rational household labor supply, *Econometrica*, 56(1), 6389.

Chiappori, P.A. (1992), Collective labor supply and welfare, *Journal of Political Economy*, 100(3), 437467.

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.

D'Errico M., Serafini R., Macchiarelli C. (2015), Differently Unequal: Zooming-In on the Distributional Dimensions of the Crisis in Euro Area Countries, *Economic Modelling*, Special issue: Current Challenges on Macroeconomic Analysis and International Finance Modelling, 48(1), 93115.

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. 29(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(4), 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.

Kerkhofs, M., Lindeboom, M., Theeuwesc, J., (1999), Retirement, financial incentives and health, *Labour Economics*, 6(2), 203227.

Kiefer, N.M. (1988). Economic Duration Data and Hazard Functions, *Journal of Economic Literature*, Vol 22(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.

Macchiarelli C., Ward-Warmedinger M. (2014), Transitions in labour market status in EU labour markets, with M. Ward-Warmedinger, *IZA Journal of European Labor Studies*, 3(17), 1-25.

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.

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.

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

Years of work	Beginning total	Failures	Survivor function	Standard error	Cum. hazard	Standard error
20	63804	21	0.9983	0.0001	0.0014	0.002
.	.	.	.	.	.	.
.	.	.	.	.	.	.
37	16451	367	0.9233	0.0018	0.0757	0.0832
38	14394	415	0.8966	0.0022	0.1036	0.113
39	12311	395	0.8679	0.0025	0.1347	0.146
40	10699	632	0.8166	0.0031	0.1922	0.2068
41	8571	443	0.7744	0.0035	0.2424	0.2599
42	7230	504	0.7204	0.004	0.3103	0.3316
43	5928	385	0.6736	0.0044	0.3734	0.3983
44	4866	320	0.6293	0.0048	0.4373	0.4661
45	4008	356	0.5734	0.0052	0.5234	0.5576
46	3164	249	0.5283	0.0055	0.5996	0.639
47	2565	181	0.491	0.0058	0.6677	0.7121
48	2114	143	0.4578	0.006	0.7327	0.7824
49	1761	106	0.4303	0.0062	0.7905	0.8452
50	1504	236	0.3627	0.0066	0.941	1.0088
51	1114	106	0.3282	0.0068	1.0317	1.1086
52	906	80	0.2992	0.0069	1.1155	1.2016
53	765	69	0.2723	0.007	1.2008	1.2969
54	635	45	0.253	0.0071	1.2675	1.3721
55	543	54	0.2278	0.0072	1.3608	1.4781
56	447	35	0.21	0.0072	1.4338	1.562
57	370	36	0.1895	0.0073	1.5239	1.667
58	309	24	0.1748	0.0073	1.5953	1.7514
59	249	15	0.1643	0.0073	1.65	1.8176
60	210	30	0.1408	0.0074	1.779	1.9753
61	159	17	0.1258	0.0075	1.874	2.0952
62	124	12	0.1136	0.0075	1.9586	2.2054
63	98	10	0.102	0.0076	2.0461	2.3234
64	77	9	0.0901	0.0077	2.1443	2.461
65	65	38	0.0374	0.0064	2.6478	3.1366

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

Table 4: Cox regressions

Variable	Hazard ratio	(Std. Err.)	Variable	Hazard ratio	(Std. Err.)
EA	0.949	(0.261)	*Income variables*		
Crisis	1.449***	(0.052)	Disposable income	1.038	(0.031)
Minimum retirement age	0.905***	(0.012)	Old age benefits	1.811***	(0.074)
			Unemployment benefits	1.852***	(0.114)
Individual characteristics			Disability benefits	2.734***	(0.216)
Female	1.513***	(0.075)	Sickness benefits	1.126	(0.087)
Married	1.047	(0.079)	Private Pension	1.164*	(0.097)
Skilled	1.617***	(0.070)	*Interaction*		
Part-time	0.115***	(0.054)	Part-time x disp. income	1.182***	(0.055)
Occupational group			*Health variables*		
2.Occ. group	0.666***	(0.032)	Health	1.536***	(0.175)
3.Occ. group	0.945	(0.047)	Health(-1)	1.472***	(0.218)
4.Occ. group	0.898*	(0.052)			
5.Occ. group	0.838***	(0.051)	*Partner characteristics*		
			2.Partner unemployed	1.344***	(0.129)
			3.Partner retired	1.045	(0.039)
			4.Partner inactive	1.117**	(0.052)
			Partner's health	0.7496***	(0.084)
Statistics					
θ	0.448	(0.119)			
LR test (frailty terms)	Prob>$\bar{\chi}2$ = (0.000)				
Wald $\chi 2$	1341.13				
Prob > $\chi 2$	(0.000)				
Log-likelihood	-34036.317				
Number of groups	26				
Observations	45,439				

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

Table 5: Cox regressions using alternative health measures

	Hazard ratio	(Std. Err.)	Hazard ratio	(Std. Err.)	Hazard ratio	(Std. Err.)
Health	1.536***	(0.175)				
Health(-1)	1.472***	(0.218)				
Limit			1.181***	(0.073)		
Limit(-1)			1.174**	(0.079)		
Chronic					1.025	(0.039)
Chronic(-1)					0.947	(0.036)
Wald χ^2	1341.13		1332.20		1315.48	
Prob > χ^2	(0.000)		(0.000)		(0.000)	
Log-likelihood	-34036.317		-34039.526		-34047.226	
Number of groups	26		26		26	
Observations	45,439		45,439		45,439	

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 A (Appendix A) for data and definitions.

Table 6: Cox regressions, early retirement sample

Variable	Hazard ratio	(Std. Err.)	Variable	Hazard ratio	(Std. Err.)
Crisis	1.512***	(0.068)	Disposable income	1.148***	(0.040)
Minimum retirement age	0.847***	(0.014)	Old age benefits	2.9327***	(0.134)
			Unemployment benefits	1.640***	(0.110)
Individual characteristics			Disability benefits	2.354***	(0.201)
Female	1.369***	(0.083)	Sickness benefits	1.347***	(0.124)
Married	1.056	(0.099)	Private pension	1.436***	(0.183)
Skilled	1.866***	(0.098)	*Interaction*		
Part-time	1.004	(0.640)	Part-time x disp. income	0.977	(0.061)
Occupational group			*Health variables*		
2.Occ. group	0.756***	(0.045)	Health	2.139***	(0.292)
3.Occ. group	0.811***	(0.047)	Health(-1)	1.762***	(0.326)
4.Occ. group	0.7656***	(0.051)			
5.Occ. group	0.749***	(0.055)	*Partner characteristics*		
			2.Partner unemployed	1.357***	(0.143)
			3.Partner retired	1.121**	(0.052)
			4.Partner inactive	1.112**	(0.060)
			Partner's health	0.779*	(0.118)
Statistics					
θ	0.506	(0.134)			
LR test (frailty terms)	Prob>$\bar{\chi}2$ = (0.000)				
Wald $\chi2$	1537.66				
Prob > $\chi2$	(0.000)				
Log-likelihood	-22779.249				
Number of groups	26				
Observations	43,335				

Note: 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 Appendix A.1. for data and definitions.

Appendix A. Variables used in the estimation

Dependent variables[27]

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).

Explanatory variables

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

[27] See also the EU-SILC's Guidelines.

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.
Crisis: Equals 1 if year of the survey equals 2009 or more.
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: Equals 1 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).

Private pension: Equals 1 if the household receive pensions from individual private plans (other than those covered under ESSPROS). The latter source is considered an extension of the standard income definition and it is cumulated for all members of the same household.

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 labor 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 traveling 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 bio-medical 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.

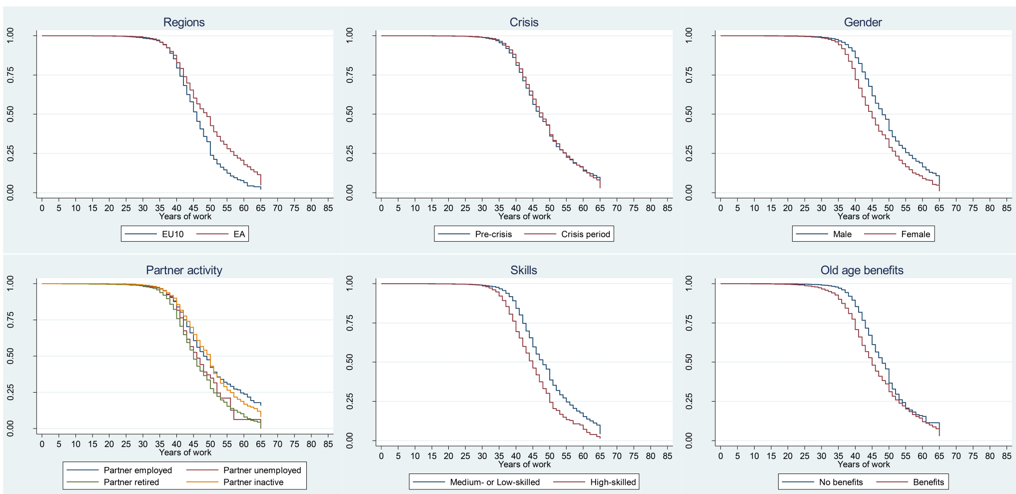
Partner's health: Equals 1 if each individual's partner (if any) assesses his health to be 'very bad'. See Health definition.

<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 chronic (long-standing) illness or condition.

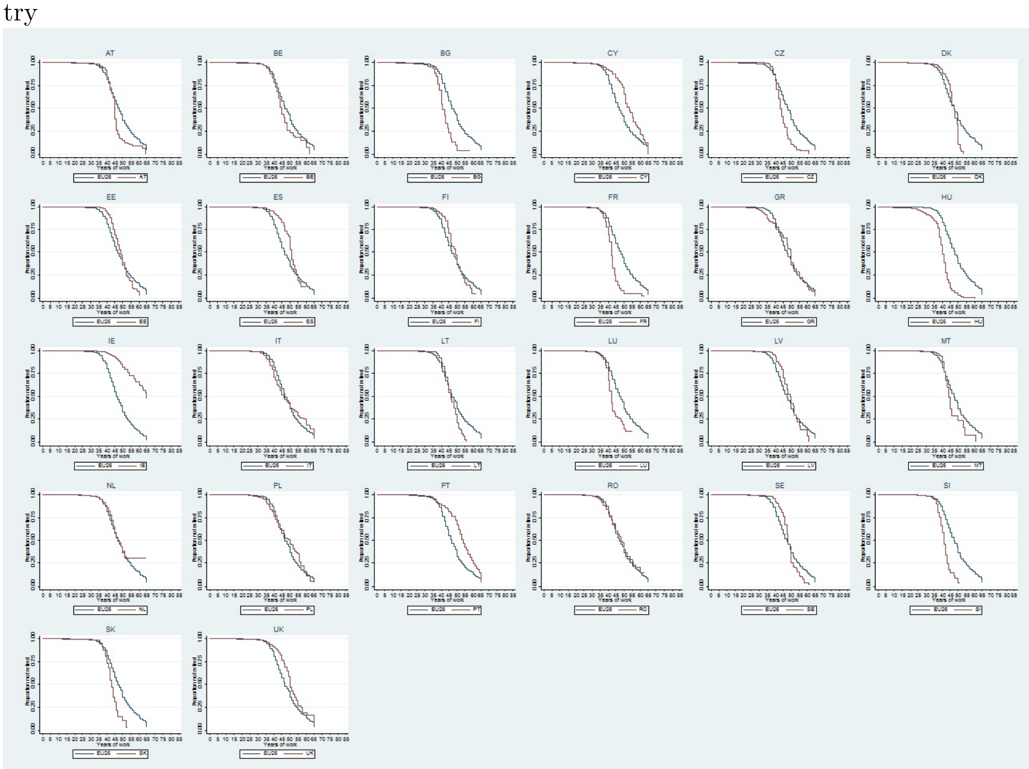
Appendix B. Additional Graphs

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



Appendix C. 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.[28] 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.[29] 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.[30]

In Table A.3 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.[31] Note, also that the Weibull model has virtually the same AIC scores as the log-logistic one.[32] Furthermore, as shown in the non-parametric

[28]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 5.2) 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.

[29]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.

[30]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.

[31]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.

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

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.[33] Estimates using the log-logistic distribution do not produce any relevant differences compared to the Weibull estimates, suggesting that the selection effect of using distributions other than the Weibull is limited.

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

The results of the Weibull model are qualitatively the same and 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.[34]

[33]The Weibull model assumes a baseline hazard of the form $h_0(t) = p^{tp-1} exp(\beta_0)$, 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.

[34]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 A1: Model selection for full parametric regressions

Distribution (metric)	Log likelihood	k	c	AIC	Ranking
Whole sample (retirement)					
Exponential (PH, AFT)	-9850.973	25	1	19751.95	6
Weibull (PH, AFT)	-1925.130	26	2	3902.259	3
Gompertz (PH)	-2210.403	26	2	4472.806	4
Log-normal (AFT)	-3090.066	26	2	6232.132	5
Log-logistic (AFT)	-1890.774	26	2	3833.547	1
Generalized gamma (AFT)	-1908.454	27	3	3870.908	2

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

Table A2: Full-parametric regression

Variable	Time ratio	(Std. Err.)	Variable	Hazard ratio	(Std. Err.)
EA	1.015	(0.033)	*Income variables*		
Crisis	0.960***	(0.004)	Disposable income	0.993**	(0.003)
Minimum retirement age	1.011***	(0.002)	Old age benefits	0.937***	(0.004)
			Unemployment benefits	0.925***	(0.007)
Individual characteristics.			Disability benefits	0.881***	(0.008)
Female	0.9519***	(0.006)	Sickness benefits	0.989	(0.009)
Married	0.992	(0.009)	Private pension	0.983*	(0.010)
Skilled	0.949***	(0.005)	*Interaction*		
Part-time	1.311***	(0.074)	Part-time x disp. income	0.980***	(0.005)
Occupational group			*Health variables*		
2.Occ. group	1.054***	(0.006)	Health	0.960***	(0.013)
3.Occ. group	1.003	(0.006)	Health(-1)	0.960***	(0.017)
4.Occ. group	1.010	(0.007)			
5.Occ. group	1.016**	(0.007)	*Partner characteristics*		
			2.Partner unemployed	0.966***	(0.011)
			3.Partner retired	0.992*	(0.004)
			4.Partner inactive	0.982***	(0.005)
			Partner's health	1.041***	(0.014)
			Constant	27.088***	(2.931)
Statistics					
ln(p)	2.130***	(0.009)			
ln(θ)	0.798***	(0.265)			
LR test (frailty terms)	Prob>=chi-bar-sq. = (0.000)				
Wald $\chi2$	1237.00				
Prob > $\chi2$	(0.000)				
Log-likelihood	-1181.2451				
Number of groups	26				
Observations	45,439				

Note: 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 Appendix A for data and definitions.