HOW FUNCTIONAL DATA CAN ENHANCE THE ESTIMATION OF HEALTH EXPECTANCY: THE CASE OF DISABLED SPANISH POPULATION

BY

IRENE ALBARRÁN, PABLO J. ALONSO-GONZÁLEZ, ANA ARRIBAS-GIL AND AUREA GRANÉ

ABSTRACT

The aging of population is perhaps the most important problem that developed countries must face in the near future. Dependency can be seen as a consequence of the process of gradual aging. In a health context, this contingency is defined as a lack of autonomy in performing basic activities of daily living that requires the care of another person or significant help. In Europe in general and in Spain in particular, this phenomena represents a problem with economic, political, social and demographic implications. The prevalence of dependency in the population, as well as its intensity and evolution over the course of a person's life are issues of greatest importance that should be addressed. The aim of this work is the estimation of life expectancy free of dependency (LEFD) based on functional trajectories to enhance the regular estimation of health expectancy. Using information from the Spanish survey EDAD 2008, we estimate the number of years spent free of dependency for disabled people according to gender, dependency degree (moderate, severe, major) and the earlier or later onset of dependency compared to a central trend. The main findings are as follows: first, we show evidence that to estimate LEFD ignoring the information provided by the functional trajectories may lead to non-representative LEFD estimates; second, in general, dependency-free life expectancy is higher for women than for men. However, its intensity is higher in women with later onset on dependency; Third, the loss of autonomy is higher (and more abrupt) in men than in women. Finally, the diversity of patterns observed at later onset of dependency tends to a dependency extreme-pattern in both genders.

KEYWORDS

ADL, cox regression, dependency, disability, functional data.

JEL codes: 62-07, 62-09, 62H20, 62H99, 62P05.

Astin Bulletin 0(0), 1-28. doi:10.1017/asb.2018.34

1. Introduction

Population aging is an ongoing global phenomenon and a powerful and transforming demographic force. Several reports have warned about the need of evidence-based policies sustained on rigorous research and on the importance to prioritize healthy aging and well-being (see for instance, WHO, 2011a,b; Lloyd-Sherlock *et al.*, 2012). In particular, one of the eight tackling societal challenges of the European program Horizon 2020 is concerned with these issues.

The decreasing mortality and increasing life expectancy in most Western European countries during the last decades are well documented (Eurostat, 2009). A key issue is to find out whether the increased life expectancy is healthy or either associated with an increase or decrease in disability (Fries, 1983).

Life expectancy is one of the most used indicators to measure quantity of life. However, if the aim is to measure quality of life, indicators related to healthy life expectancy should be used. These kinds of indicators introduce health status (morbility or disability) of the individual (Robine and Ritche, 1991; Robine *et al.*, 2003). For instance, Sanderson and Scherbov (2010) propose disability-free life expectancy as a consistent disability aging measure for many countries in order to provide better tools for policy makers. In this paper we are interested in dependency,[1] which is a more restrictive concept than disability. Therefore, our indicator will be life expectancy free of dependency (LEFD) (see Martel and Bélanger, 2000).

The aim of this work is to estimate the LEFD, that is, the expected number of years that a person can live free of this contingency based on mortality and morbility conditions. The evolution of dependency in the disabled Spanish population will be studied through a pseudo panel constructed from EDAD 2008, in the lack of longitudinal studies or the possibility to link different cross-sectional surveys.[2]

EDAD 2008 (Survey on Disability, Personal Autonomy and Dependency Situations 2008, undertaken by the Spanish Statistical Office—INE) is the most recent Spanish survey about disability and was the first Spanish survey that used the internationally accepted measures established by the 'International classification of functioning, disability and health'.[3] It was also the first time that the survey included information useful for studying the dependency phenomenon, such as the average hours per week of special care received by the dependent person.

Our main contribution is the estimation of LEFD based on gender, dependency degree (moderate, severe and major) and homogeneous groups of individuals with similar dependency pattern. The characterization of homogeneous groups of individuals is obtained through the proximity of the dependency trajectories (that are derived using the retrospective reported information of each individual from birth up to 2008, contained in EDAD 2008) to a central trend within each age–gender group. These central trends are computed via functional data techniques. To estimate LEFD in all the scenarios considered, we use the specification of Cox regression model in terms of the survival function, having in mind that the event of interest is not ‘survival’ itself but ‘being dependency-free at a given age’. Finally, we estimate the LEFD for disabled Spanish population within homogenous groups considering gender, dependency degree and ages from 30 to 100.

A very preliminary attempt to estimate life expectancy can be found in Albarrán *et al.* (2014). The present work is a novelty approach to solve the same problem and, as far as we know, this is the first time that dependency evolution is used to characterize the individuals in order to enhance the regular estimation of health expectancy. Other recent studies on dependency are Albarrán *et al.* (2015) and Albarrán-Lozano *et al.* (2017), regarding dependent children.

Other authors used Markovian multi-state models to study long-term care risk (see, for instance, Biessy, 2017; Fong *et al.*, 2017; Levantesi and Menzietti, 2018), where several states such as autonomy, different degrees of dependency and death must be established. However, this is not the case in database EDAD 2008 where all surveyed people was alive in 2008. To circumvent this problem, in this paper we present a novel approach via functional data techniques.

The main findings are as follows: first, the relative errors of the LEFD calculated using the partition by proximity-groups versus the global LEFD show evidence that the global LEFD may not be representative of the Spanish population. From economic and demographic points of view, this is a relevant finding, since the expected dependent population would demand care services (health care, pensions and other services) that should be covered and related expenditures should be financed. Second, in general, healthy life expectancy is higher for women than for men. However, the intensity of dependency is higher for those women with later onset of dependency. Third, the loss of autonomy is higher (and more abrupt) for men than for women. Fourth, for people with the earliest onset of dependency, having less than 50 points (out of 100) in the dependency rating scale is crucial for living a longer time free of major dependency. Finally, the loss of autonomy in people with the earliest onset of dependency tends to a singular extreme-pattern, characterized by few variable effects on LEFD estimation, whereas for people with the latest onset of dependency the loss of autonomy has a diversity of patterns that can be associated to a wide range of variable effects on LEFD. These diversities of patterns are higher in women than in men.

The paper proceeds as follows. Section 2 contains the definition of dependency and its graduation according to the Spanish legislation. Also some information about the Spanish survey EDAD 2008 is presented. Section 3 is devoted to explain the construction of the dependency trajectories from a pseudo panel from EDAD 2008, a description of the functional data techniques that we are going to use and the proximity measure that will help to characterize groups with homogeneous dependency trajectories. This is the most technical section of the paper. In Section 4 we propose the methodology to estimate LEFD and analyse the main results. Finally we conclude in Section 5.

2. Dependency situation in Spain: legislation and data-set EDAD 2008

2.1. Spanish legislation on dependency

When talking about dependency two fundamental aspects must be considered. First, the *definition* itself. In the Spanish case, article 2 of Act 39/2006, of 14th December, on the Promotion, Personal Autonomy and Care for Dependent persons states that dependency is a 'permanent state in which persons that for reasons derived from age, illness or disability and linked to the lack or loss of physical, mental, intellectual or sensorial autonomy require the care of another person/other people or significant help in order to perform basic activities of daily living or, in the case of people with mental disabilities or illness, other support for personal autonomy'.

Second, the *assessment of dependency*, which is usually solved using specific dependency rating scales that take into account the disabilities suffered by the person jointly with their intensity. Royal Decree 504/2007 rules the evaluation of dependency in Spain. The Spanish dependency rating scale goes from 0 to 100 points and it is categorized in four degrees: non-dependant (less than 25 points), I-moderate (greater or equal to 25 but under 50 points), II-severe (greater or equal to 50 but under 75 points), III-major (greater or equal to 75 points). See Table A1 in the Appendix for more details.

To acknowledge the entitlement to the benefits of the system, a person must reach at least the moderate degree, that is, at least 25 points are needed to be considered dependant in Spain. According to the dependency rating scale value or score reached by an individual, the Spanish legislation establishes a minimum level of protection, which is defined and financially guaranteed by the General State Administration.

2.2. EDAD 2008 survey

In order to provide reliable estimates at the national level, the EDAD 2008 survey was performed around the country using sampling. In particular, a two-stage sampling was performed, stratified and proportional to the size of the Spanish autonomous regions (with stratified sampling distribution proportional to population size in stratum, within each Spanish province). Therefore, each individual in EDAD 2008 is associated to a weight reflecting the population group that represents. See INE (2010) for more details on the sampling methodology.

EDAD 2008 gives information about people with disabilities that were living either at home or in institutions. In the first case, the survey was prepared interviewing 260,000 people who were living in 96,000 different houses, whereas for institutionalized people, 11,000 people in 800 centres were asked about their situation. Interviewed people were not only those suffering disabilities, but also their relatives and/or carers. This survey is based on the concept

TABLE 1

ESTIMATION OF CHILDREN AND ADULT POPULATION WITH DISABILITY LIVING AT HOME: 95% CONFIDENCE INTERVALS FOR NUMBER AND PREVALENCE RATE.

	Disabled people (in thousands) and prevalence rate (in %)		
Age (in years)	Total	Male	Female
Under 6	53–67.8 (1.8–2.3%)	30.9–41.9 (2.1–2.8%)	24.4–27.6 (1.1–1.7%)
Between 6 and 44	576.8–648.4 (2.3–2.7%)	316–374.2 (2.5–3.0%)	240.5–285.5 (2.0–2.4%)
Between 45 and 64	897.7–1005.9 (8.0–9.0%)	379.4–437.6 (6.9–8.0%)	505.2–580.4 (8.9–10.3%)
Between 65 and 79	1138–1264.6 (20.8–23.1%)	422.4–487.2 (17.1–19.8%)	703.5–789.5 (23.4–26.3%)
80 or more	971.8–1079.8 (45.8–50.9%)	277.6–326.2 (36.6–43.0%)	683.1–764.7 (50.0–56.0%)
Total	3740.4–3955.2 (8.2–8.6%)	1488.9–1605.5 (6.6–7.1%)	2226.6–2374.6 (9.3–10.2%)

Source: INE elaboration. Results derived from the weighted survey data.

of self-perceived disability, in accordance with the recommendations of the World Health Organization. So, the target people is identified through a set of questions about the possible difficulties they can find in doing some specific activities. Despite its drawbacks, the main advantage of this strategy is that it is focused on the daily activities of the individuals and the problems they may have while doing them, with no consideration of medical matters.

In 2008 the Spanish population ascends to 46.66 million people (23.10 million men and 23.56 million women). According to EDAD 2008, there are more than 4.1 million Spanish people suffering at least one kind of disability. Although the global prevalence rate is situated between 8.2% and 8.6% with a 95% of confidence, in the case of people living at home, this rate is lower than that for people living in institutions (8.4% and 17.7%, respectively). Disability is related to two main factors: gender and age; until 45 years old, the male prevalence is statistically significant greater than the female one. After that age, the relative incidence is greater for women. In general terms, more than 57% people with this problem are at least 65 years old, being most of them women. Table 1 contains an estimation (derived from the weighted survey data) of children and adult population with disability living at home.

The sample selected for the present study is formed by 7446 individuals and represents 2.35% of the Spanish population in 2008, that is, more than one million people (325,253 men and 773,079 women). We remind that each individual in the sample has a weight reflecting the population group that represents. These weights have been taken into account in all the computations of this paper. We give more details about the selected sample in Section 4.

The data set obtained from EDAD 2008 contains, among many other variables, the ages at which each person in the sample has suffered a change in his/her health condition and his/her current age. Although the survey includes the term 'dependence' in its title, the questionnaire does not consider any question on this topic. So, the dependency score is not reported in the survey but can be computed from the information provided in it and applying the Spanish legislation (Act 39/2006 and Royal Decree 504/2007). This computation is not straightforward. In particular, the Spanish dependency score is a sum of the weighted product of several factors that take into account the disabilities suffered jointly with their severity and the degree of supervision (average hours per week of special care received). The Spanish legislation classifies the disabilities into 11 main activities: eating and drinking, control of physical needs (excretion and urinate), washing, other personal tasks, dressing, maintaining health, mobility, moving inside home, moving outside home, housekeeping and taking decisions. In turn, each activity contains several tasks. For example, under mobility activity we find the following tasks: sitting down, lying down, standing up, changing posture from a siting position and changing posture from bed. Each activity and task contribute to the dependency score with different weights according to the age of the individual and the occurrence (or not) of mental disability or cognitive impairment. See Albarrán and Alonso (2009) for more details on the computation of the Spanish dependency score.

3. Searching for proximity-groups

One of the objectives of the paper is to search for different patterns within age–gender groups, that is, we are interested in identifying dependency trajectories that lie close/far/very far from a central trend of the group. The reason is that, as we will see in Section 4.2, LEFD can experiment huge variations as dependency trajectories depart from the central trend of the corresponding age–gender group. The central trend of each age–gender group will be obtained by using functional data techniques that we describe in Section 3.2. We start by obtaining the dependency trajectories.

3.1. From a pseudo panel to dependency trajectories

The aim of this section is to construct a dependency trajectory for each individual in the sample, that is, a curve describing the evolution of the personal dependency situation over time, and to use functional data techniques to analyse the database. Indeed, in functional data analysis, individual observations are real functions of time, observed at discrete time points. Each curve provides the evolution of a certain process for a given individual (see Ramsay and Silverman, 2005, for an overview). In our case, the process of interest is the evolution of dependency.

Notice that even if the available data come from a one-time survey, individuals were asked about their whole medical history, so we have information concerning their dependency situation/score from birth up to 2008. Then, for the ith individual we observe $(t_{i1}, y_{i1}), \ldots, (t_{in_i}, y_{in_i})$, respectively, the ages

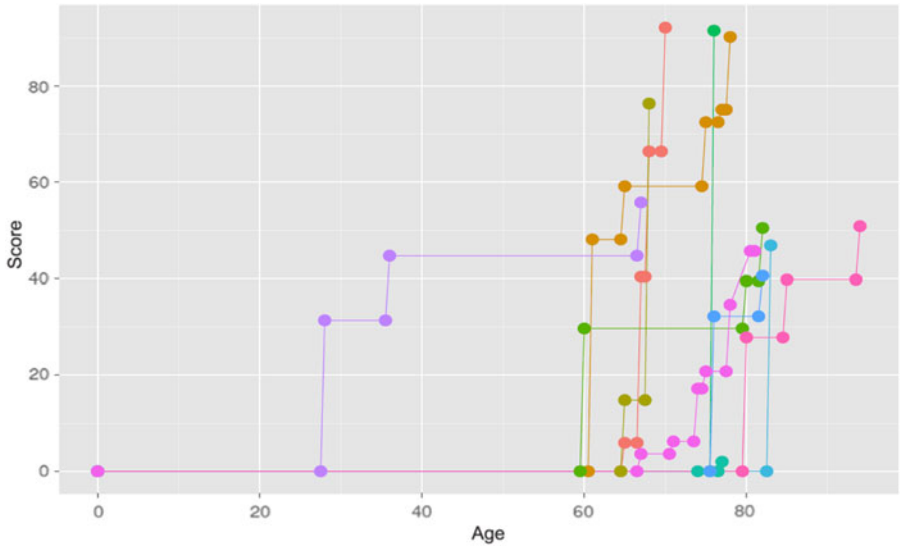

FIGURE 1: Examples of dependency curves from birth up to 2008. Dependency score is represented in the vertical axis and age in the horizontal one.

when changes occur and the dependency scores at these ages, and a_i, the current age (at 2008). From these data, in order to stress the step character of these curves, we add a first point $(0, 0)$ (only if $t_{i1} > 0$), intermediate points $(t_{ih} - \delta, y_{ih-1})$ between (t_{ih-1}, y_{ih-1}) and (t_{ih}, y_{ih}), where δ is a chosen short period of time, and a final point (a_i, y_{in_i}) (only if $t_{in_i} < a_i$). Indeed, even if the person's health/dependency condition can be seen as a smooth process, the dependency score function is piecewise constant since changes in score only take place once some particular disability status has been reached and recognized according to the Spanish legislation. These transformed sequences will make up our set of observations from now on. For the sake of simplicity, we will still refer to them as $(t_{ih}, y_{ih})_{h=1,\ldots,n_i}$, $i = 1, \ldots, n$. Thus, we have n discretely observed curves $y_1, \ldots, y_n$ defined in different time intervals $[0, a_i]$, $i = 1, \ldots, n$. We illustrate the step character of these curves in Figure 1. Notice that the curves are non-decreasing, meaning that recovery is not possible (in spite of adaptation strategies). In order to better interpret this figure, let us focus our attention on a particular trajectory, for example, the purple one. For this particular case, we can see three jumps in the score, taking place at ages 28, 36 and 67, which means that, this particular individual became dependent at the age of 28 with a dependency score of 31 points (Degree I); at the age of 36 another disability appeared increasing the score up to 45 points (Degree I); finally, at the age of 67 another disability took place and the score jumped to 56 points (Degree II). The trajectory also tells us that the current age of this individual at 2008 was 67 years old.

In order to apply any functional data analysis technique, we need functions that are defined over the same interval. One idea would be to consider

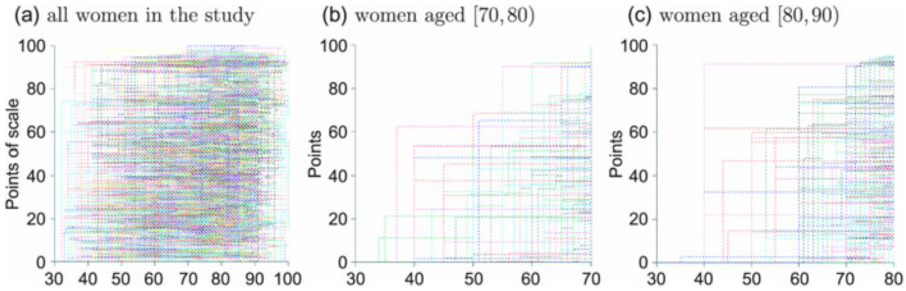

FIGURE 2: Dependency trajectories for all women under study and two selected groups of women. Dependency score is represented in the vertical axis and age in the horizontal one.

the different cohorts present in the sample and analyse the dependency trajectories within each cohort. However, this may lead to many different under-represented cohorts, since the age range of the individuals in the sample is large. Instead of that, we consider disjoint groups of people in age intervals of 10 years. Within each age interval $[A, A + 10)$ we truncate individual curves to get them defined in $[0, A]$. For all the analyses performed in this article, the first age interval is $[50, 60)$ and the last one is $[90, \infty)$. Moreover, we are particularly interested in those people with a dependency score of zero at the age of 30, and from now on, they will be grouped in 10 age–gender intervals (5 groups per gender) according to their current age at 2008. Notice that we consider such a group of 30-year-old non-dependent people in order to obtain LEFD estimates for different dependency scenarios, which are useful for health, economic, demographic and insurance contexts. In Figure 2 we depict the resulting dependency trajectories for all women under study and two selected groups of women with dependency score of zero at the age of 30. The aim of this figure is to illustrate different patterns of dependency. For example, from panels (b) and (c) we can see that the density of curves in the last 10 years is greater for women aged 80 than for women aged 70, meaning that the dependency situation tends to worsen with age. That is the reason why we have considered disjoint groups of people according to their current age at 2008, otherwise, due to the great amount of curves (see panel (a)) it would be not possible to track this phenomenon.

3.2. Estimating the central trend

Providing a measure of centrality when dealing with functional data is not an straightforward task. Indeed, not only the levels of the curves matter, but also their shapes, whose information is more difficult to incorporate to any numerical summary. The problem aggravates if we consider curves for which the main features are not aligned. It is well known that in this context the sample pointwise or cross-sectional mean is a poor estimator of the mean behaviour (Gasser *et al.*, 1984; Kneip and Gasser, 1992; Gasser and Kneip, 1995).

In this context, it is extremely important to use measures of centrality that can take into account the misalignment between the curves of the sample.

Indeed, in the particular case of the dependency evolution curves that we study in this work, it is very natural to consider that the evolution of dependency may present a common pattern which is accelerated or delayed in some individuals with respect to others. A general framework for modelling such trajectories is the so-called *time warping* model, since it includes any kind of parametric model in which the individual parameters allow for variations in scale and phase with respect to some given functional form, such as growth models, and also semi-parametric models in which this functional form is unknown and estimated from the data, such as shape-invariant models (see Wang and Gasser, 1997, for details). Also, notice that we can assume that observations are free of measurement error since they correspond to the evaluation, on an official numerical scale, of the particular conditions suffered by each individual at each moment.

In the time warping model, two approaches to estimate the central trend or mean behaviour of the data are possible: (i) to align or register the curves and to compute any desired sample statistic on the registered sample; and (ii) to define appropriate estimators directly on the observed sample, taking into account the nature of the data. For the analysis of the dependency data set we will consider an estimator of the second kind that we describe in the following.

3.2.1. *Deepest curve.*
The literature on estimators directly defined on the unregistered sample is relatively small. Liu and Müller (2004), Dupuy *et al.* (2011) or Arribas-Gil and Romo (2012) are works which are particularly concerned by the definition of suitable population centrality measures, and their corresponding sample statistics, in the time warping model. However, there might be curves with a typical shape but taking atypical values (abnormally high or low at some locations) and, in this case, a registration procedure would neutralize the effect of those curves with an atypical shape (due to the fact that they may be delayed or accelerated with respect to the rest). Therefore, for the analysis of the dependency data set we will consider the approach of Arribas-Gil and Romo (2012) since it provides a robust estimator of the central trend for a set of curves.

A way to provide a centrality measure that is robust against the two types of atypical curves is to use functional depth. Indeed, the deepest curve of a sample in terms of modified band depth (López-Pintado and Romo, 2009) has been proven to be an accurate and robust estimator of the central pattern of a sample of curves in the time warping model (Arribas-Gil and Romo, 2012). It can be understood as a generalization of the median to functional data because, intuitively, it is the curve that is most surrounded by other curves. Therefore, it provides an accurate measure of centrality since: (i) it is a curve geometrically located in the centre of the sample and (ii) it presents a typical shape because it is one of the observed curves.

As it was mentioned before, we are interested in estimating the central trend of each age–gender group. Therefore, for each one of these 10 groups we compute the deepest curve in terms of modified band depth using the `roahd` package in `R` by Tarabelloni *et al.* (2016). As an example, in Figure 3 we depict

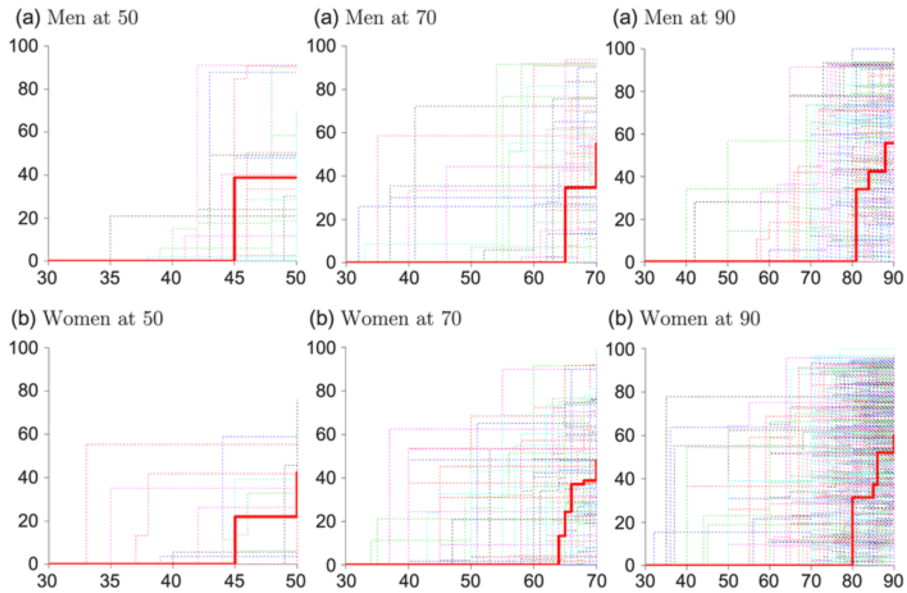

FIGURE 3: Dependency trajectories for men and women with their corresponding deepest curves (in bold red). Dependency score is represented in the vertical axis and age in the horizontal one.

the dependency trajectories with the corresponding deepest curve (in bold red) for several age–gender groups, where we observe that for people aged 50 the first score value reached by the deepest curve is lower for women than for men, meaning that the loss of autonomy is stronger in men than in women at earlier ages. The contrary happens for later ages, reproducing somehow the behaviour observed in Table 1.

3.3. Distance to the deepest curve

Once we have estimated the central trend within each age–gender group, we propose to search for different patterns of trajectories within each group by computing a proximity measure to the deepest curve.

Therefore, within each group we compute the L^2-distance of each trajectory to the corresponding deepest one multiplied by 1 (or −1) if the trajectory is most of the time above (or below) the deepest curve. Indeed, if we note $y_{ij}(\cdot)$ the dependency curve of individual i in group j, and $m_j(\cdot)$ the deepest curve of group j, we obtain

$$d_j(i) = \sqrt{\sum_t \left(y_{ij}(t) - m_j(t)\right)^2} \cdot \operatorname{sign}\left(\sum_t \left(y_{ij}(t) - m_j(t)\right)\right), \qquad (3.1)$$
$$i = 1, \ldots, n_j, \; j = 1, \ldots, 10,$$

where n_j is the number of individuals in group j. This yields a numerical summary for each one of the trajectories that can be used to establish

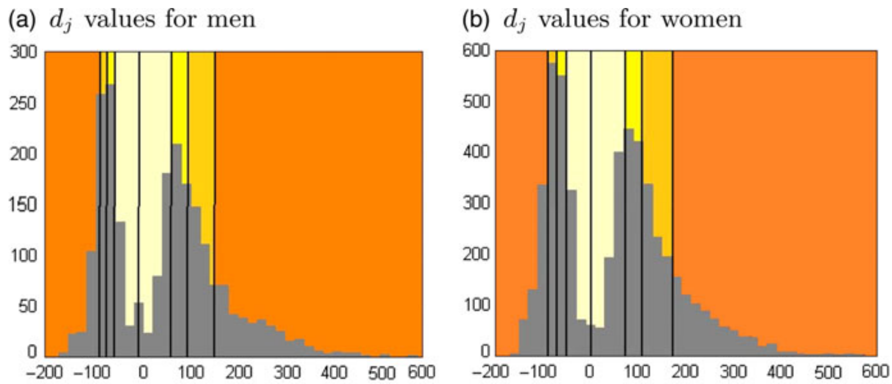

FIGURE 4: Histogram for d_j for men and women. Vertical lines stand for the quartiles of sets $\{d_j < 0\}$ and $\{d_j \geq 0\}$, respectively. Different colours indicate the degree of departure from $d_j = 0$.

different patterns. As we will see later, these patterns will exhibit quite different dependency-free life expectancies.

In Figure 4 we depict the histograms for the proximity measure d_j defined in Equation (3.1) computed over all the trajectories by gender. Notice that the sign of d_j indicates whether the trajectory is below or above the deepest curve within its age–gender group. In particular, negative values of d_j correspond to trajectories below the deepest curve and, therefore, to individuals with lower dependency scores than those of the central trend of their age–gender groups (later onset of dependency compared to the central trend). In fact, the best situations are expected for the left-tail values of d_j. On the other hand, positive values of d_j correspond to trajectories above the deepest curve and, hence, to individuals with higher dependency scores than those of the central trend of their age–gender groups (earlier onset of dependency compared to the central trend). In this case, the worst situations are expected for large values of d_j. With vertical lines and different colours we highlight eight regions corresponding to the division established by the quartiles of sets $\{d_j < 0\}$ and $\{d_j \geq 0\}$, respectively.

Since each set of values $\{d_j < 0\}$ and $\{d_j \geq 0\}$ has a different meaning, later/earlier onset of dependency, we propose to compute the LEFD for the groups of individuals established by the quartiles in each set, yielding to eight groups for each gender, that we call proximity-groups (see Figure 4 where the regions defined by the quartiles are in different colours). Remind that, we are also interested in estimating LEFD for the three dependency degrees (moderate, severe, major) within each proximity-group. Depending on the researcher, other partitions or sets of particular interest, such as extreme observations, can be considered.

3.4. Summary of the procedure

In the following we briefly describe the main steps to find the proximity-groups for a given set of dependency trajectories. However, the procedure can be applied to track the occurrence and evolution of a certain phenomenon in a group of individuals, whenever other methodologies cannot be applied.

Remind that, in our case, for each individual we have registered the age when a disability occurs and the dependency score. We also know their current ages at 2008.

Step 1 Group the individuals in age–gender groups, since dependency can be seen as a consequence of gradual aging and it is commonly assumed that the evolution of this phenomenon is different in men than in women. In our case, we have considered 10 disjoint groups (5 per gender) according to their current ages at 2008. In particular, $[50, 60), [60, 70) \ldots, [90, \infty)$.

Step 2 Calculate the deepest curve of each age–gender group. The `R` package `roahd` by Tarabelloni *et al.* (2016) contains several possibilities. In our case, we have computed the deepest curve in terms of the modified band depth (López-Pintado and Romo, 2009).

Step 3 For each individual compute the proximity to the corresponding deepest curve using measure d_j from formula (3.1). Remind that we are only interested in two sets of values per gender: $\{d_j < 0\}$, meaning later onset of dependency, and $\{d_j \geq 0\}$, earlier onset of dependency.

Step 4 To establish the proximity-groups per gender, use certain quantiles within sets $\{d_j < 0\}$ and $\{d_j \geq 0\}$. In our case, we have used the quartiles, leading to eight proximity-groups per gender.

4. Estimating LEFD

In this section we propose to use Cox proportional hazards regression model to estimate the LEFD (Cox, 1972; Cox and Oakes, 1984). The reason is that we are interested not only in estimating healthy life expectancy, but also in identifying the variables with persistent, high and low impacts on LEFD. Cox regression model allows to explore the determinants of life expectancy (or survival probability) and to estimate hazard ratios of the covariates included in the model, such as gender, disabilities, etc. (Czado and Rudolph, 2002). This model is one of the most popular regression techniques in survival analysis and can be written as follows:

$$h(t, \mathbf{x}) = h_0(t) \exp\left(\boldsymbol{\beta}'\mathbf{x}\right), \tag{4.1}$$

where $\mathbf{x}$ is a vector of explanatory variables, $\boldsymbol{\beta}$ is the vector of coefficients of the explanatory variables, function $h(t)$ is the expected hazard at time t and $h_0(t)$ is the baseline hazard function that represents the hazard when all the explanatory variables are equal to zero.

In our case, since we are interested in estimating the LEFD, the first step is to consider the specification of Cox regression model in terms of the survival function $S(t)$. Using the well-known relationship

$$S(t) = \exp\left(-\int_0^t h(t)\, dt\right),$$

and modelling the hazard function by Equation (4.1), we have that

$$S(t, \mathbf{x}) = \exp\left(-\int_0^t h_0(t)\exp\left(\boldsymbol{\beta}'\mathbf{x}\right) dt\right) = \left[\exp\left(-\int_0^t h_0(t)\, dt\right)\right]^{\exp\left(\boldsymbol{\beta}'\mathbf{x}\right)}$$

that can be written as follows:

$$S(t, \mathbf{x}) = S_0(t)^{\exp\left(\boldsymbol{\beta}'\mathbf{x}\right)},$$

where

$$S_0(t) = \exp\left(-\int_0^t h_0(t)\, dt\right)$$

is the baseline survival function at time t, and, as before, $\mathbf{x}$ is the vector of explanatory variables and $\boldsymbol{\beta}$ is the vector of coefficients of the explanatory variables.

In our case, the event of interest is not 'survival' itself, but 'being dependency-free at a given age'. That is, we interpret function $S(t, \mathbf{x})$ as the time spent free of dependency at a given age. At this point we must remind that EDAD 2008 only contains records of alive people at 2008, hence the effect of death is ignored. That is, the estimated being dependency-free probability at a given age is in fact the probability of being dependency-free at a given age given that a person is alive at that age. Thus, the next step consists in correcting these estimates by survival probabilities given by the disabled Spanish pensioners' mortality table Orden TAS/4054/2005 (2005). Finally, we estimate life expectancy at a given age x using the following formula:

$$\hat{e}_x = 0.5 + \sum_{t=1}^{\omega - x} {}_t\hat{p}_x,$$

where ${}_t\hat{p}_x = S(t, \mathbf{x})/S(0, \mathbf{x})$ and ω is the limit of life.

We consider several scenarios to estimate LEFD attending to gender, eight proximity-groups (IV–I for $\{d_j < 0\}$ and I–IV for $\{d_j \geq 0\}$) and three dependency degrees (I-moderate, II-severe, III-major) given by the Spanish legislation. Indeed, for a non-dependent person we compute three different LEFDs, which are the expected number of years that a person can live outof each one of the three dependency degrees. Notice that the dependency history of an individual may not reach all the states, that is, the first score reached by an individual can be greater than 50 or 75 points. This is the reason why, in the following tables and for the sake of simplicity, we call 'degree I' to the expected number of years that a person can live out of any dependency degree (score under 25 points); 'degree II' stands for the expected number of years that a person can live out of severe or major dependency (score under 50 points); 'degree III' stands for the expected number of years that a person can live out of major dependency (score under 75 points).

4.1. Searching for main effects on LEFD

The variables included in the Cox regression model are gender, proximity-group and the disabilities recorded in EDAD 2008 related to dependency, that is: to present difficulties in performing postural changes, bathing/hygiene, control of physical needs, conducting household life, maintaining interaction and interpersonal relationships, following medical treatments and mobility difficulties (inside and outside the house).

We include in the analysis people that in 2008 were between 50 and 100 years old, with a dependency score of zero at the age of 30. We remind that we are particularly interested in such a group of 30-year-old non-dependent people in order to obtain LEFD estimates for different dependency scenarios, which are useful for health, economic, demographic and insurance contexts. We also remind that the sample is formed by 7446 individuals (2230 men and 5216 women), that represent more than one million Spanish people, according to INE.

We estimate and validate Cox regression model with `survival` package in `R`.

Figure 5 contains a heat map of the main effects derived from Cox regression results estimated for each proximity-group and dependency degree (see Table A2 in the Appendix for detailed results). Light colours represent low, sometimes null, impacts on the LEFD, whereas dark colours stand for high effects on the LEFD.

In the following we interpret the main variable effects on LEFD using Figure 5. In both genders we observe that the loss of autonomy, that is, the occurrence of disabilities, is more relevant for people with later onset of dependency compared to a central trend. The variables with the most persistent impact on both genders' LEFD are control of physical needs and bathing/hygiene. Additionally, medical treatments present a lasting effect on men's LEFD and the same happens on maintaining interaction and interpersonal relationships regarding women's LEFD. On the other hand, mobility difficulties inside and outside the house have the lowest impact on both genders' LEFD. The second to low for men's LEFD is maintaining interaction and interpersonal relationships. Finally, from Table A2 in the Appendix, we observe that the highest impact on men's LEFD is registered by difficulties

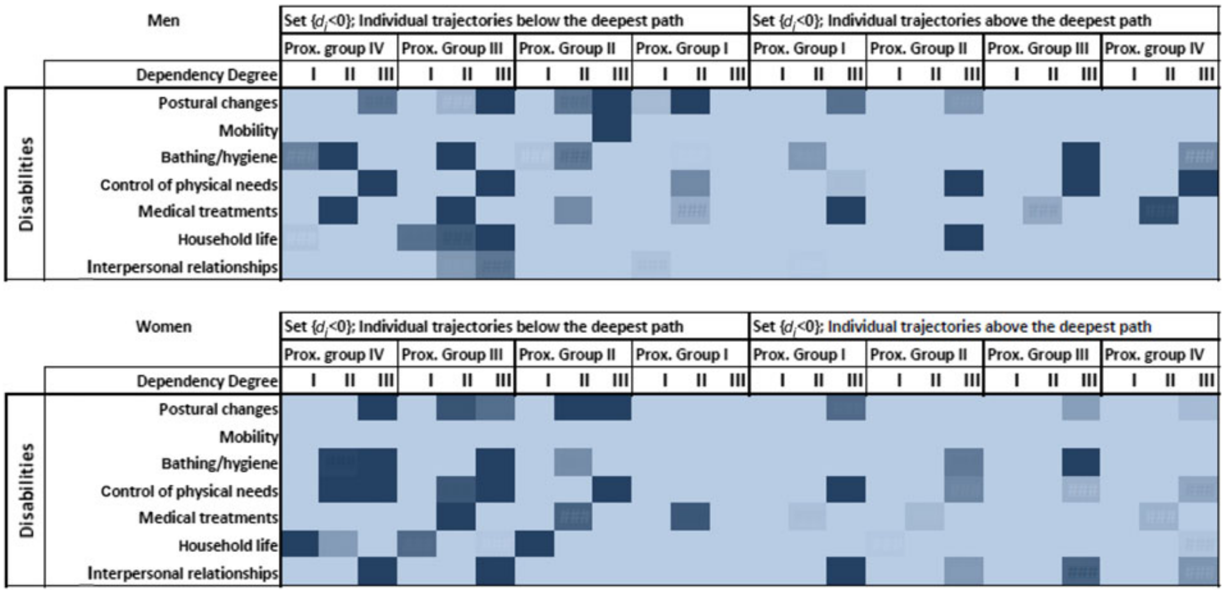

FIGURE 5: Heat map of main effects derived from Cox regression results (based on Table A2 in Appendix).

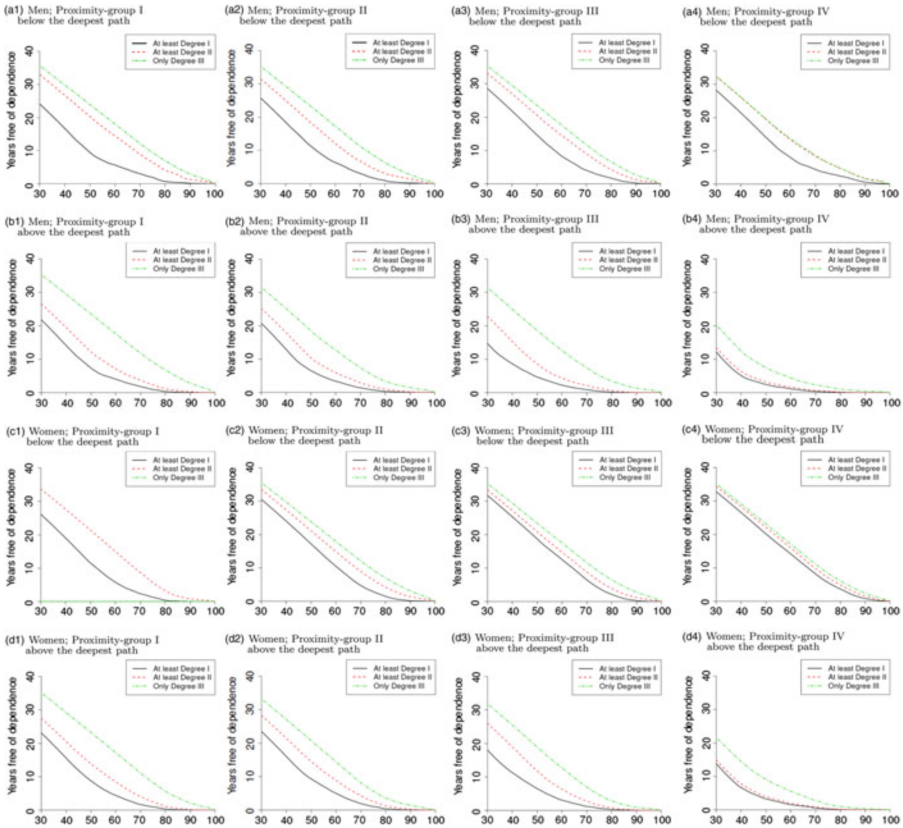

FIGURE 6: LEFD for men (panels a1–b4) and women (panels c1–d4). Age is represented in the horizontal axis and years free of dependency in the vertical one.

in performing postural changes and by control of physical needs regarding women's LEFD. Both variables present a similar behaviour, having high impacts on people with later onset of dependency compared to a central trend and fading afterwards.

4.2. Contribution of the proximity-groups on estimated LEFD

In Figure 6 we depict the estimated LEFD for the scenarios considered for the eight proximity-groups for men and women. In particular, each panel contains three curves corresponding to the evolution of the LEFD along age in three situations (At least Degree I, At least Degree II and Only Degree III).[4] In general, looking at LEFD curves we can observe that, first, they reach higher values for women than for men, meaning that the LEFD is higher for women than for men; second, the decreasing rate is higher (and more abrupt) for men than for women, which means that the loss of autonomy is higher (and more abrupt) in men than in women; finally, we focus on the extreme proximity-groups IV-below the deepest path (a4–c4 panels) and IV-above the deepest path (b4–d4 panels). The former correspond to individuals with the best dependency situation (latest onset of dependency), where we observe that LEFD curves for 'At least Degree I' reach lower values for men than for women, indicating that dependency phenomenon tends to appear earlier in men than in women. Additionally, the estimation of women's LEFD is very similar for the three dependency situations (notice the proximity of the three curves), meaning that, although the phenomenon tends to appear later in women than in men, its intensity is higher in women than in men. A similar behaviour can be observed in panels a3–c3 and a2–c2. The latter corresponds to individuals with the worst dependency situation (earliest onset of dependency), where we can observe an analogous shape of the LEFD for both genders in the three dependency situations, reaching slightly higher values for women, indicating a slightly worse life expectancy for them. Additionally, we observe that the time spent free of low or moderate dependency degree is quite similar in both genders. This may suggest that for people with the earliest onset of dependency, having less than 50 points in the dependency score is crucial for spending (a maximum of) 8–10 years free of major dependency.

As a summary, Table 2 contains the estimated LEFD for men and women at three particular ages jointly with the LEFD calculated without taking into account the partition by proximity-groups (rows LEFD for men and LEFD for women). We may remind that these LEFD estimations are computed from survey EDAD 2008 that contains only disabled people. Therefore, if we want to extend these estimates to the Spanish population, they must interpreted as a lower bound, that is, as the 'at least' expected numbers of years free of dependency. Nevertheless, the methodology that we propose in this paper is not restricted to this database.

In Table 2 we observe that the variance of LEFD increases with age and tends to decrease with dependency degree. In general, the variance is greater for women. The relative errors[5] show evidence that the global LEFD by gender, calculated without taking into account the partition by proximity-groups, is far from any of the LEFD values estimated by proximity-groups. This means that the global LFED may not be representative of the dependent Spanish population, not even for those individuals within the most central proximity-groups, that is, for those that are the nearest to the corresponding central trend.

To illustrate the contribution of the partition by proximity-groups on the LEFD estimation, we consider the following example. The global LEFD for a 30-year-old women with dependency degree I is 21.6 years, which means that the expected number of years that a 30-year-old woman can live out of any dependency degree is 21.6. In other words, it is expected that a 30-year-old woman can reach 51 years old out of any dependency degree. However,

TABLE 2

SUMMARY OF LEFD FOR THREE PARTICULAR AGES.

	Men	30 years old			50 years old			70 years old		
	Degree	I	II	III	I	II	III	I	II	III
$\{d_j < 0\}$	Proximity-group IV	28.234	32.320	32.416	14.117	19.501	19.623	4.417	8.424	8.648
	Proximity-group III	28.616	33.167	35.395	14.701	20.618	23.567	4.040	9.022	11.706
	Proximity-group II	25.664	31.426	35.186	11.352	18.389	23.290	3.065	6.745	11.234
	Proximity-group I	24.069	32.908	35.648	9.424	20.275	23.902	3.004	8.899	12.120
$\{d_j \geq 0\}$	Proximity-group I	21.848	26.600	35.375	7.378	12.228	23.541	1.838	3.711	11.812
	Proximity-group II	20.792	25.213	31.548	6.524	10.375	18.474	1.346	2.857	7.400
	Proximity-group III	14.762	22.884	31.427	4.695	8.647	18.659	0.878	2.127	7.425
	Proximity-group IV	12.191	13.509	20.194	2.426	3.183	7.497	0.359	0.600	2.337
LEFD for men		18.733	23.614	31.868	6.475	10.429	19.391	1.556	3.195	8.583
	Women	30 years old			50 years old			70 years old		
	Degree	I	II	III	I	II	III	I	II	III
$\{d_j < 0\}$	Proximity-group IV	32.805	34.370	35.099	20.195	22.210	23.175	8.187	9.780	11.110
	Proximity-group III	31.714	33.353	35.264	18.694	20.864	23.394	6.465	8.783	11.589
	Proximity-group II	30.578	33.549	35.614	17.215	21.141	23.860	4.753	8.973	12.116
	Proximity-group I	26.172	33.743	–	11.557	21.380	–	2.343	8.699	–
$\{d_j \geq 0\}$	Proximity-group I	23.145	27.523	35.166	8.766	13.904	23.264	1.664	4.017	11.102
	Proximity-group II	23.527	28.316	33.279	9.340	14.536	20.765	1.938	4.283	8.204
	Proximity-group III	18.077	26.109	31.882	6.543	11.858	18.958	1.190	2.745	6.992
	Proximity-group IV	13.852	14.923	21.651	3.288	3.909	9.114	0.579	0.759	2.757
LEFD for women		21.601	26.657	32.860	8.781	13.623	20.577	2.092	4.337	9.045

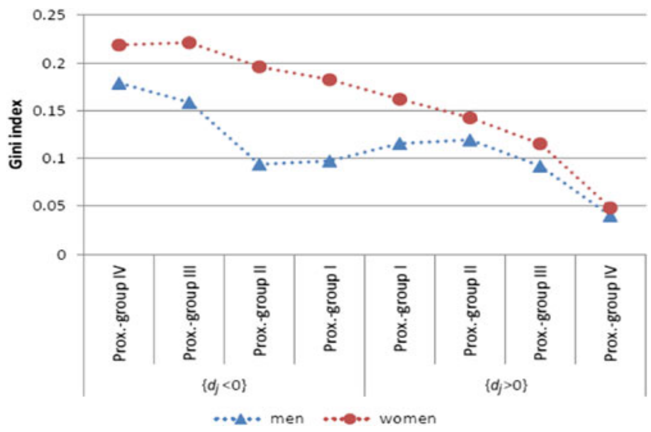

FIGURE 7: Gini index for men and women within the proximity-groups.

a more accurate estimation can be obtained by considering proximity-groups, ranging from 32.8 years (proximity-group IV within set $\{d_j < 0\}$) to 13.8 years (proximity-group IV within set $\{d_j \geq 0\}$). That is, in the best situation, a 30-year-old woman can live out of any dependency degree until 62 years old and, in the worst case, a 30-year-old woman becomes dependent at the age of 43 years old. Notice that this difference of around 20 years is relevant, at least, from demographic and economic points of view, in the sense that, the expected dependent population would demand care services (health care, pensions and other services) that should be covered and related expenditures should be financed.

Finally, we can use Gini index as a numerical measure to summarize dependency evolution per proximity-group. At this point it is important to remind that dependency curves are associated to individuals that represent population groups of different sizes.[6] Therefore, each individual is associated to a weight. The sum of all of these weights is an estimation of the whole dependent Spanish population.

Two variables are needed in order to calculate the Gini index: p variable for population and q variable for quantities or amounts. In our case, p variable has been computed summing the weights of all individuals with positive score and for each age ranging from 30 to 100 years old. On the other hand, q variable has been computed for each age ranging from 30 to 100 years old as the product of (positive) scores by weights and represents the total amount of score values at each age. We depict the results in Figure 7.

We can conclude the following. For proximity-groups within set $\{d_j \geq 0\}$, the concentration tends to reduce as d_j increases, which means that dependency curves tend to be more similar as their distance to the deepest curve increases, that is, the loss of autonomy in people with the earliest onset of dependency tends to a singular extreme-pattern. The opposite happens for proximity-groups within set $\{d_j < 0\}$, that is, Gini index tends to reduce as d_j

gets closer to zero, meaning that dependency curves tend to be less similar as their distance to the deepest curve increases. This may suggest that the loss of autonomy in people with the latest onset of dependency presents more than one pattern (similar results were found in Albarrán-Lozano *et al.*, 2017). Both findings can be related with the heat map given in Figure 5. In particular, the extreme-pattern observed for people with earliest onset of dependency corresponds to few variable effects on LEFD (estimated via Cox regression model). On the other hand, the diversity of patterns for people with later onset of dependency can be associated to a wider range of variable effects on LEFD. Finally, comparing both genders, women's Gini index experiments a regular decreasing, reaching always higher values than in men, meaning that women have a wider range of dependency patterns than men.

5. Conclusions and further research

Dependency, that is, lack of autonomy in performing basic ADL, can be seen as a consequence of the process of gradual aging. In Europe in general and in Spain in particular this phenomenon represents a problem with economic, political and social implications. The prevalence of dependency in the population as well as its intensity and evolution over the course of a person's life are issues of greatest importance that should be addressed.

The aim of this work is to estimate LEFD, that is, the expected number of years that a person can live free of this contingency based on mortality and morbility conditions. The evolution of dependency in the disabled Spanish population is studied through a pseudo panel constructed from EDAD 2008, in the lack of longitudinal studies or the possibility to link different cross-sectional surveys. In particular, individual dependency trajectories are obtained using the retrospective reported information of each individual from birth up to 2008, contained in EDAD 2008, and applying the Spanish legislation (Act 39/2006 and Royal Decree 504/2007).

The main contribution of this paper is the estimation of LEFD based on functional trajectories to enhance the regular estimation of health expectancy. Using the information of EDAD 2008, we estimate the number of years spent free of dependency for disabled people, according to gender, dependency degree (moderate, severe, major) and the later or earlier onset of dependency compared to a central trend.

In both genders we observe that the loss of autonomy, that is, the occurrence of disabilities, is more relevant for people with later onset of dependency compared to the central trend. Control of physical needs and bathing/hygiene present the most persistent impact on both genders' LEFD, whereas the lowest impact is registered by mobility difficulties inside and outside the house.

Concerning LEFD, we found that, first, in general, dependency-free life expectancy is higher for women than for men. However, the intensity of dependency is higher for those women with later onset of dependency. Second, the loss of autonomy is higher (and more abrupt) in men than in women. Third, for people with the earliest onset of dependency, having less than 50 points (out of 100) in the dependency score is crucial for living a longer time free of major dependency. Finally, the loss of autonomy in people with the earliest onset of dependency tends to a singular extreme-pattern, characterized by few variable effects on LEFD, whereas for people with later onset of dependency the loss of autonomy has a diversity of patterns that can be associated to a wider range of variable effects on LEFD. The diversity of theses patterns is higher in women than in men.

The proposed methodology can be applied in several directions. For example, dependency paths can be used in the health industry to help to discover the true cost of different dependency patterns that can be established by the proximity-groups and dependency degrees. Dependent population may demand care services that should be offered and financed (public health and/or private insurance companies).

Acknowledgements

Financial support is from research project MTM2014-56535-R by the Spanish Ministry of Economy and Competitiveness. The authors thank the Associate Editor and two Referees, whose comments helped to improve the quality of the paper.

Notes

1. Resolution R(98) of the Council of Europe defines dependency as 'such state in which people, whom for reason connected to the lack or loss of physical, mental or intellectual autonomy, require assistance and/or extensive help in order to carry out common everyday actions'. This definition has been translated into national legislations in a heterogeneous way (Kamette, 2011).
2. Three surveys about disability have been undertaken by the Spanish Statistical Office (INE) during the last 30 years. The first one was conducted in 1986 and was the Survey about Disabilities, Impairments and Handicaps. Then came the Survey about Disabilities, Impairments and Health Status, which was prepared using data from 1999. Finally, the last one was EDAD in 2008. Although all of them talk about disabilities, it is not possible to track this phenomenon in a homogeneous way along the years because the definition of that concept changed through the years depending on the classification that was used to prepare the survey.
3. In 2001, the World Health Organization (2011a) established a framework for measuring health and disability at both individual and population levels, which was known as the 'International classification of functioning (ICF), disability and health'. The ICF tries to establish a consensus in its understanding, by establishing a difference between the basic activities of living daily (ADL) and the instrumental ADL. The basic activities are defined as those activities which are essential for an independent life.
4. 'At least Degree I' stands for expected number of years that a person can live out of any dependency degree (score under 25 points); 'At least Degree II' stands for the expected number of years that a person can live out of severe or major dependency (score under 50 points); 'Only Degree III' stands for the expected number of years that a person can live out of major dependency (score under 75 points).

5. The relative error of a given LEFD is computed as its difference to the global LEFD by gender divided by the global LEFD by gender.

6. In EDAD 2008 a two-stage sampling was conducted by INE, leading to individuals that represent population groups of different sizes.

References

Albarrán, I. and Alonso, P. (2009) La población dependiente en España: estimación del número y coste global asociado a su cuidado. *Estudios de Economía*, **36**(2), 127–163.

Albarrán, I., Alonso, P., Arribas-Gil, A. and Grané, A. (2014) Can personal dependency paths help to estimate life expectancy free of dependency? In *Mathematical and Statistical Methods for Actuarial Sciences and Finance* (eds. M. Sibilio and C. Perna), pp. 1–5. Cham, Switzerland: Springer.

Albarrán, I., Alonso, P. and Grané, A. (2015) Profile identification via weighted related metric scaling: An application to dependent Spanish children. *Journal of the Royal Statistical Society A*, **178**, 1–26.

Albarrán-Lozano, I., Alonso-González, P. and Arribas-Gil, A. (2017) Dependence evolution in the Spanish disabled population: A functional data analysis approach. *Journal of the Royal Statistical Society A*, **180**(2), 657–677.

Arribas-Gil, A. and Romo, J. (2012) Robust depth-based estimation in the time warping model. *Biostatistics*, **13**, 398–414.

Biessy, G. (2017). Continuous-time semi-Markov inference of biometric laws associated with a long-term care insurance portfolio. *Astin Bulletin*, **47**, 527–561.

Cox, D. (1972) Regression models and life tables. *Journal of the Royal Statistical Society B*, **34**, 187–220.

Cox, D.R. and Oakes, D. (1984) *Analysis of Survival Data*. London: Chapman & Hall.

Czado, C. and Rudolph, F. (2002) Application of survival analysis methods to long-term care insurance. *Insurance: Mathematics and Economics*, **31**, 395–413.

Dupuy, J., Loubes, J. and Maza, E. (2011) Non parametric estimation of the structural expectation of a stochastic increasing function. *Statistics and Computing*, **21**(1), 121–136.

Eurostat (2009) Health statistics and atlas on mortality in the European Union. Eurostat. Luxembourg.

Fong, J., Sherris, M. and Yap, J. (2017) Forecasting disability: Application of a frailty model. *Scandinavian Actuarial Journal*, **2017**(2), 125–147.

Fries, J. (1983) The compression of morbidity. *The Milbank Memorial Fund Quarterly*, **61**, 397–419.

Gasser, T. and Kneip, A. (1995) Searching for structure in curve samples. *Journal of the American Statistical Association*, **90**, 1179–1188.

Gasser, T., Müller, H., Köhler, W., Molinari, L. and Prader, A. (1984) Nonparametric regression analysis of growth curves. *The Annals of Statistics*, **12**, 210–229.

INE (2010) Encuesta sobre Discapacidad, Autonomía personal y Situaciones de Dependencia (EDAD), Metodología. Ed. Subdirección General de Estadísticas Sociales Sectoriales (INE), Madrid, España.

Kamette, F. (2011) Dependency care in the EU: A comparative analysis. Foundation Robert Schumann. Social Issues European Issue, No. 196.

Kneip, A. and Gasser, T. (1992) Statistical tools to analyze data representing a sample of curves. *The Annals of Statistics*, **16**, 82–112.

Levantesi, S. and Menzietti, M. (2018) Natural hedging in long-term care insurance. *Astin Bulletin*, **48**(1), 233–274.

Liu, X. and Müller, H. (2004) Functional convex averaging and synchronization for time-warped random curves. *Journal of the American Statistical Association*, **99**(467), 687–699.

Lloyd-Sherlock, P., McKee, M., Ebrahim, S., Gorman, M., Greegross, S., Prince, M., Pruchno, R., Gutman, G., Kirwood, T., O'neill, D., Ferrucci, L., Kritchewski, S. and Vellas, B. (2012) Population ageing and health. *The Lancet*, **379**, 1295–1296.

López-Pintado, S. and Romo, J. (2009) On the concept of depth for functional data. *Journal of the American Statistical Association*, **114**, 486–503.
Martel, L. and Bélanger, A. (2000) Regression models and life tables. *Canadian Social Trends-Statistics Canada Autumn*, **2000**, 26–29.
Orden TAS/4054/2005 (2005) BOE num. 316, 28 de diciembre de 2005. Ministerio de Trabajo y Asuntos Sociales.
Ramsay, J. and Silverman, B. (2005) *Functional Data Analysis*, 2nd ed. New York: Springer Series in Statistics.
Robine, J. and Ritche, K. (1991) Healthy life expectancy: Evaluation of global indicator of change in population health. *BMJ*, **302**, 457–460.
Robine, J., Jagger, C., Mathers, C., Krimmins, E. and Suzman, R. (2003) *Determining Health Expectancies*. Chichester: Wiley.
Sanderson, W. and Scherbov, S. (2010) Remeasuring aging. *Science*, **329**, 1287–1288.
Tarabelloni, N., Arribas-Gil, A., Ieva, F., Paganoni, A. and Romo, J. (2016) roahd: Robust analysis of high dimensional data. Package on CRAN. Version 1.0.4. Published 2016-07-06.
Wang, K. and Gasser, T. (1997) Alignment of curves by dynamic time warping. *The Annals of Statistics*, **25**, 1251–1276.
WHO (2011a) Global health and aging. WHO (World Health Organization). US National Institute of Aging.
WHO (2011b) Global report on disability. WHO (World Health Organization). US National Institute of Aging.

Irene Albarrán
Statistics Department
Universidad Carlos III de Madrid
C/Madrid 126, 28903 Getafe
Spain
E-Mail: irene.albarran@uc3m.es

Pablo J. Alonso-González
Economics Department
Universidad de Alcalá
Plaza de la Victoria 2, 28802 Alcalá de Henares
Spain
E-Mail: pablo.alonsog@uah.es

Ana Arribas-Gil
UC3M-BS Institute of Financial Big Data
C/Madrid 135, 28903 Getafe
Spain
E-Mail: ana.arribas@uc3m.es

Aurea Grané (Corresponding author)
Statistics Department
Universidad Carlos III de Madrid
C/Madrid 126, 28903 Getafe
Spain
E-Mail: aurea.grane@uc3m.es

APPENDIX

TABLE A1

DEPENDENCY GRADUATION ACCORDING TO SPANISH LEGISLATION.

Dependency	Degree	Level	Score	Dependency	Degree	Level	Score
Non-dependent	–	–	[0, 25)	Severe	II	1	[50, 65)
					II	2	[65, 75)
Moderate	I	1	[25, 40)	Major	III	1	[75, 90)
	I	2	[40, 50)		III	2	[90, 100]

Moderate dependency	The person needs help in order to perform various basic ADL[a] at least once a day or the person needs intermittent or limited support for his/her personal autonomy.
Severe dependency	The person needs help in order to perform various basic ADL two or three times a day, but he/she does not want the permanent support of a carer or he/she needs extensive support for his/her personal autonomy.
Major dependency	The person needs help in order to perform various basic ADL several times a day or he/she needs the indispensable and continuous support of another person or he/she needs generalized support for his/her personal autonomy.

[a] ADL stands for Activities of Daily Living.

TABLE A2

COX REGRESSION RESULTS FOR MEN AND WOMEN: e^{β}'S AND β'S STANDARD DEVIATION (WITHIN PARENTHESIS).

Men	Set $\{d_j < 0\}$; Individual trajectories below the deepest path											
	Proximity-group IV			Proximity-group III			Proximity-group II			Proximity-group I		
Degree	I	II	III	I	II	III	I	II	III	I	II	III(*)
postural	1.283	1.566	3.107	1.496	2.321	17.951	1.120	3.094	5.323	2.240	4.678	–
changes	(0.02)	(0.026)	(0.047)	(0.014)	(0.02)	(0.068)	(0.016)	(0.031)	(0.102)	(0.015)	(0.029)	–
mobility	1.142	0.862	1.839	1.202	0.225	0.219	1.406	0.416	5.250	1.622	0.310	–
	(0.028)	(0.029)	(0.045)	(0.026)	(0.027)	(0.064)	(0.017)	(0.032)	(0.139)	(0.017)	(0.036)	–
bathing/	2.966	4.844	0.000	1.090	7.189	0.000	2.117	3.040	0.000	1.334	2.046	–
hygiene	(0.044)	(0.106)	(0)	(0.024)	(0.067)	(0)	(0.018)	(0.04)	(0.06)	(0.016)	(0.037)	–
physical	0.770	1.548	5.162	0.750	1.287	4.222	0.648	1.105	0.989	1.340	2.845	–
needs	(0.019)	(0.027)	(0.046)	(0.017)	(0.023)	(0.066)	(0.018)	(0.026)	(0.06)	(0.017)	(0.03)	–
medical	1.471	14.768	0.000	1.359	4.597	0.000	1.225	2.830	0.000	1.419	2.407	–
treatments	(0.025)	(0.086)	(0)	(0.018)	(0.043)	(0)	(0.016)	(0.034)	(0)	(0.016)	(0.037)	–
household	2.116	0.603	0.000	3.164	3.407	6.896	1.835	1.756	0.000	1.292	1.817	–
life	(0.04)	(0.045)	(0)	(0.032)	(0.064)	(0.106)	(0.019)	(0.033)	(0)	(0.017)	(0.037)	–
interpers.	1.189	1.521	0.000	1.555	2.667	3.091	0.797	1.178	0.000	2.134	1.317	–
rel.	(0.017)	(0.02)	(0)	(0.021)	(0.023)	(0.042)	(0.019)	(0.028)	(0)	(0.021)	(0.035)	–

TABLE A2

CONTINUED.

Men	Set $\{d_j \geq 0\}$; Individual trajectories above the deepest path											
	Proximity-group I			Proximity-group II			Proximity-group III			Proximity-group IV		
Degree	I	II	III	I	II	III	I	II	III	I	II	III
postural	1.183	1.547	3.163	1.304	1.580	2.689	1.161	1.320	1.779	0.592	0.816	1.787
changes	(0.01)	(0.012)	(0.048)	(0.011)	(0.014)	(0.025)	(0.01)	(0.012)	(0.022)	(0.012)	(0.013)	(0.019)
mobility	0.674	0.491	0.178	1.057	0.985	0.529	0.789	0.781	1.045	0.711	1.479	0.626
	(0.011)	(0.013)	(0.046)	(0.012)	(0.016)	(0.023)	(0.012)	(0.013)	(0.021)	(0.014)	(0.016)	(0.02)
bathing/	1.408	2.657	0.000	1.256	0.829	0.319	1.337	1.361	14.387	1.020	1.002	2.836
hygiene	(0.016)	(0.024)	(0)	(0.015)	(0.02)	(0.035)	(0.015)	(0.018)	(0.119)	(0.019)	(0.021)	(0.038)
physical	0.523	0.904	2.190	0.544	1.807	5.502	0.386	1.136	5.619	0.437	0.930	4.103
needs	(0.01)	(0.012)	(0.041)	(0.011)	(0.014)	(0.037)	(0.011)	(0.012)	(0.034)	(0.011)	(0.012)	(0.022)
medical	1.488	1.994	363.577	1.094	1.156	1.780	1.309	2.453	0.520	1.116	3.600	1.022
treatments	(0.013)	(0.017)	(0.395)	(0.013)	(0.017)	(0.038)	(0.014)	(0.017)	(0.028)	(0.016)	(0.021)	(0.022)
household	1.325	1.073	0.539	1.226	1.342	4.731	1.054	1.214	1.759	1.041	0.755	1.130
life	(0.013)	(0.016)	(0.047)	(0.013)	(0.017)	(0.053)	(0.013)	(0.016)	(0.028)	(0.017)	(0.019)	(0.025)
interpers.	1.264	2.045	1.901	0.935	1.205	1.872	0.680	0.793	1.818	0.753	0.939	1.995
rel.	(0.011)	(0.014)	(0.037)	(0.011)	(0.012)	(0.017)	(0.01)	(0.011)	(0.015)	(0.01)	(0.01)	(0.013)

Table A2

Continued.

Women	Set $\{d_j < 0\}$; Individual trajectories below the deepest path											
	Proximity-group IV			Proximity-group III			Proximity-group II			Proximity-group I		
Degree	I	II	III	I	II	III	I	II	III	I	II	III(*)
postural	1.069	1.536	3.633	1.709	2.841	2.599	1.027	3.249	11.905	1.032	1.704	–
changes	(0.012)	(0.019)	(0.028)	(0.01)	(0.016)	(0.029)	(0.01)	(0.018)	(0.112)	(0.008)	(0.017)	–
mobility	1.104	0.472	0.349	1.242	0.713	0.648	1.332	0.723	0.000	1.054	0.444	–
	(0.019)	(0.024)	(0.026)	(0.012)	(0.016)	(0.027)	(0.011)	(0.016)	(0)	(0.008)	(0.017)	–
bathing/	1.344	2.968	3.718	1.312	1.094	4.165	1.523	2.320	0.000	0.831	1.250	–
hygiene	(0.02)	(0.039)	(0.063)	(0.014)	(0.026)	(0.107)	(0.011)	(0.027)	(0)	(0.01)	(0.024)	–
physical	1.292	5.036	11.112	1.508	2.805	5.735	0.701	1.444	23.578	0.856	1.631	–
needs	(0.013)	(0.02)	(0.038)	(0.01)	(0.014)	(0.035)	(0.011)	(0.015)	(0.093)	(0.01)	(0.017)	–
medical	1.399	1.255	0.000	0.880	3.403	0.000	0.901	2.711	1.613	0.963	2.802	–
treatments	(0.016)	(0.024)	(0)	(0.011)	(0.023)	(0)	(0.01)	(0.02)	(0.055)	(0.008)	(0.021)	–
household	3.097	2.190	0.000	2.640	1.433	1.793	3.350	1.627	0.000	1.650	0.926	–
life	(0.025)	(0.036)	(0)	(0.018)	(0.028)	(0.079)	(0.018)	(0.029)	(0)	(0.012)	(0.023)	–
interpers.	1.133	1.438	3.128	0.819	1.142	3.460	0.795	1.038	1.122	0.769	1.082	–
rel.	(0.01)	(0.012)	(0.015)	(0.011)	(0.013)	(0.024)	(0.012)	(0.015)	(0.035)	(0.013)	(0.021)	–

(continued)

TABLE A2

CONTINUED.

Women	Set $\{d_j \geq 0\}$; Individual trajectories above the deepest path											
	Proximity-group I			Proximity-group II			Proximity-group III			Proximity-group IV		
Degree	I	II	III	I	II	III	I	II	III	I	II	III
postural changes	1.001	1.734	2.640	1.241	1.510	1.366	1.551	1.459	2.161	1.271	1.062	1.902
	(0.006)	(0.009)	(0.026)	(0.007)	(0.009)	(0.016)	(0.007)	(0.009)	(0.015)	(0.008)	(0.008)	(0.013)
mobility	0.787	0.644	0.716	0.918	0.933	1.059	0.639	0.902	0.618	0.987	0.930	1.373
	(0.007)	(0.008)	(0.022)	(0.007)	(0.009)	(0.015)	(0.008)	(0.009)	(0.013)	(0.009)	(0.009)	(0.013)
bathing/ hygiene	1.025	1.461	1.076	1.196	1.373	2.488	0.833	0.698	3.224	0.596	1.324	0.706
	(0.008)	(0.012)	(0.044)	(0.009)	(0.012)	(0.03)	(0.01)	(0.011)	(0.027)	(0.014)	(0.015)	(0.026)
physical needs	0.653	1.012	5.079	0.620	1.712	2.395	0.490	1.356	2.003	0.573	1.108	2.085
	(0.006)	(0.008)	(0.029)	(0.007)	(0.009)	(0.018)	(0.007)	(0.008)	(0.015)	(0.007)	(0.008)	(0.011)
medical treatments	0.973	1.857	0.468	0.910	1.916	1.046	0.859	1.612	0.584	0.969	1.988	0.884
	(0.007)	(0.009)	(0.027)	(0.008)	(0.012)	(0.017)	(0.008)	(0.01)	(0.015)	(0.01)	(0.01)	(0.013)
household life	1.395	0.722	0.935	1.791	0.708	1.566	1.489	1.304	1.538	1.231	0.928	1.823
	(0.01)	(0.011)	(0.039)	(0.01)	(0.012)	(0.031)	(0.01)	(0.011)	(0.022)	(0.012)	(0.011)	(0.021)
interpers. rel.	0.769	1.308	5.747	0.745	1.048	2.212	0.696	0.890	2.695	0.701	0.760	2.250
	(0.008)	(0.009)	(0.022)	(0.007)	(0.008)	(0.011)	(0.007)	(0.007)	(0.01)	(0.006)	(0.006)	(0.008)

(*) The sample size is not enough to obtain significative results.