NON-LINEAR MODELS OF DISABILITY AND AGE APPLIED TO CENSUS DATA

Irene Albarrán [1], Pablo J. Alonso [2] and J. Miguel Marín [3]

Abstract

It is usually considered that the proportion of handicapped people grows with age. Namely, the older the man/woman is, the more level of disability he/she suffers. However, empirical evidence shows that this assessment is not always true, or at least, it is not true in the Spanish population.

This study tries to assess the impact of age on disability in Spain. It is divided into three different parts. The first one is focused in describing the way disability is measured in this work. We used a former index defined by the authors that distinguishes between men and women. The second one is focused in a literature review about the methods used in this paper. This section emphasizes on local regression, feed forward neural networks and *BARS*. Finally, in the last section estimations are undertaken. Several methods are used and, therefore, there are fairly differences in the results, not only among the methodologies, but also between genders.

[1] Departamento de Estadística, Universidad Carlos III de Madrid, Av. de la Universidad Carlos III 22, 28270, Colmenarejo (Madrid), e-mail: ialbarra@est-econ.uc3m.es.

[2] Departamento de Estadística, Estructura Económica y O.E.I, Universidad de Alcalá, Plaza de la Victoria, 2, 28802, Alcalá de Henares (Madrid), e-mail: pablo.alonsog@uah.es

[3] Departamento de Estadística, Universidad Carlos III de Madrid, C/ Madrid 126, 28903, Getafe (Madrid), e-mail: jmmarin@est-econ.uc3m.es.

1 Introduction

Happiness is usually related to a situation in which people have covered their needs and enjoy of good health, namely, a time in which there is no illness and no physical or mental limitations to do anything. According to this, good health may be connected to not suffering disabilities.

Disability is an ambiguous term: we can consider medical and social points of views (Jiménez and Huete 2002). Medical view treats disabilities as a personal problem, due to illness or any other organic upsets, whereas the social one is worried about the integration in society of disabled people. In this context, disability can be defined as a gap between a person's abilities and environmental requirements. This is an important indicator of health status of people that has been shown to be strongly associated with mortality, specially in elderly people (Manton, 1988; Ruigomez et al., 1993; Wolinsky et al., 1993; Scott et al., 1997; Lamarcaa et al., 2003). It is an issue that affects individual, community, neighbourhood, and family (Institute of Medicine, 1991). However, people can acquire disabilities at any point in their lives (Jamoom et al., 2008). Disability increases the risk of need for home help (Avlund et al., 2001), hospitalization (Wolinsky et al., 1994, Avlund et al., 2001), nursing home admission (Laukkanen et al., 2000) and premature death (Jagger et al., 1993).

According to self-reports by people, disabilities are placed among the most important determinants of reduction in quality of life. This is the rationale for trying to develop or maintain a high quality of life after the acquisition of disability (Kemp, 1999). The critical importance of improving the well-being of people with disabilities is highlighted in many national health plans. Self-reported health status is reduced with age and among people with disabilities. Some health strategic plans in North America, Europe and Japan encourage the use of self-rated health as a measure to evaluate health status in the population. Factors that may impact self-reported health status include severity of disability or health condition, type of activity limitations, and age of the person with disabilities (Kemp, 1999; Trividi et al., 2006). The importance of disability measured by limitations in performing certain activities of daily living ($ADLs$) relies on the fact that it provides a measure of the burden caused by suboptimal health. $ADLs$ have been extensively used for assessing the need for health services (Fried and Bush, 1988) and as eligibility criteria for long-term care (Spector and Fleishman, 1998; Stone and Murtaugh, 1990).

Disability is a burden not only to individuals but also to society. It is also an important public health problem, linked to very substantial costs (Lamarcaa et al., 2003). This factor encourages the aim to get a right measurement of this phenomenon. It should also be one of the focuses of all public health policies and programs, trying to cover the whole continuum of services for preventing this situation and for restoring functional capacity through rehabilitative measures (Guralnik, 2005).

There are two characteristics that can affect on the impact of disability over people: age and gender of patients. There is a general idea that usually relates disability with age: the older the people, the greater the prevalence. In fact, studies of health transitions have suggested that older age is associated with a greater probability of functional decline and decreased odds of stability and improvement over time (Crimmins and Saito, 1993; Hogan et al., 1999). As our society ages, issues related to prevention of functional limitations and disability are emerging

as a public health priority. Some authors are interested in "age-related disability"(Morris and Doyle, 2008). The increase in the number of old people is presumably also increasing the number of old people with disabilities (University of Jyväskylä, 2004).

The effect of age has been studied specially in older people or in small groups of young people affected by a very specific disability. However, troubles related to disability can appear in any moment during the lifetime, not only in the third age.

Gender is another factor that can explain the impact of disability. Although the relationship between sex and disability has been studied for more than 20 years, there is no consensus about how this factor influences disability and through what mechanisms (Wray and Blaum, 2001). It is well known that there is a consistently higher prevalence of self-reported disability in women than in men, yet the independent effect of sex varies with the measures of disability examined and the covariates considered (Verbrugge, 1985, 1989; Fried et al., 1994; Johnson and Wolinsky, 1994; Merrill et al., 1997; Ferrer et al., 1999). The higher prevalence of disability reported by women, can be explained because women and men differ in the prevalence or severity of diseases causing disability, in their reporting of disability, and in their rates of mortality (Fried et al., 1994; Merrill et al., 1997; Verbrugge, 1985, 1989; Wingard, 1984). However, it is not clear why the independent effect of sex on disability differs across studies.

Most studies about gender and disability have only examined data on older adults. Comparisons of older versus younger age groups (Clark et al., 1997) may also provide insights into the underlying mechanisms and generalizability of the sex effect on different measures of disability and across age groups. While some researchers concentrate on explaining the relationship between sex and disability, others assume interaction effects by modeling separately men and women (Johnson and Wolinsky, 1994; Manton, 1988). Leveille et al., (2000) conclude that the disability incidence has greatest impact on the higher prevalence of disability in women compared with men.

This paper tries to describe the intensity of disability across the whole lifetime. It does not intend to analyse the effect of gender on the appearance of this contingency. Instead of, men and women have been separately studied. In each gender, it has been analysed the impact of age on the intensity that this contingency may be suffered. To do this, Spanish data from *EDDES*[1] were used. This survey was prepared in 2002, but the data were collected in 1999.

The remainder of the paper is set out as follows. Section 2 presents some empirical results and the disability index that relates the incidence of disability in the population with respect to their age. Section 3 is focused on the review of the methodologies that have been used in this paper: local regression, Bayesian spline regression and feed forward neural networks. In Section 4, the previous techniques have been used to estimate the relationship between age and the assessing-disability index. We show the main results that were obtained using the different methodologies. Finally, Section 5 concludes with some remarks.

---

[1]EDDES (Encuesta sobre Discapacidades, Deficiencias y Estados de Salud) can be translated as Disabilities, Impairments and Health Status Survey

2 Empirical Application: The Impact of Age on a Disability Index

As it has been pointed out in Section 1, caring disable people supposes a huge amount of funds. So, this fact encourages to rightly assess this contingency. Accurate measurement of disability is important not only for individuals but also at a national level. Knowledge about the prevalence and incidence of disability in the population is essential for anticipating demand for services and for planning. At the individual level, disability is assessed to determine eligibility for participation in care programs, and to assist in discharge and care planning (Fleishman et al., 2002).

One of the most extended methods to evaluate this contingency is asking people about their health and capabilities to do any activity. That is, the main source of information comes from self-reports of basic activities of daily living -*BADL*- (such as mobility and self-care, Katz et al., 1963) and instrumental activities of daily living -*IADL*- (such as preparing meals, shopping, telephoning, managing money, taking medication, Lawton et al., 1969). Nevertheless, there is no consensus about which activities can be considered as basic and which as instrumental. So the same activity can be classified into a group or in other depending on the author. For instance, in *RUG-III* index, clothing and washing by him/herself are both considered as *IADL* whereas in most classifications they are treated as *BADL*. Other examples can be found in Barthel Index (Mahoney and Barthel, 1965), where going up/down stairs are considered as *BADL* or in the *AGGIR* scale, which considers that speaking by phone is a *BADL*.

In order to avoid problems with classifications, the index used in this paper is elaborated with no distinction between basic or instrumental activities. So, the 36 possible disabilities are those reflected in the Classification of Impairments, Disabilities and Handicaps (*ICIDH*) from 1980. More recently performance-based measures of functioning have been applied to the assessment of functional limitation and consequent disability (University of Jyväskylä, 2004).

In this paper, self-report data are referred to population registered in *EDDES*, a survey made by the Spanish National Statistics Institute with people over 6 years old. Each person that has been recorded in this survey represents a number of people. So the sum of the all representation factors is the whole Spanish population over 6 in 1999. According to data in *EDDES*, almost 3.5 million people was suffering at least one disability in 1999. However, not all these individuals are in a similar situation. Most of them can live by themselves and only a group needs to be helped by a third person. This group is composed by almost 1.5 million Spaniards in this situation in 1999. A classification of them, dividing by age and gender is shown in Table 1:

Table 1: Handicapped people who needs third person aid.
Classification according age and gender

| | Men | | Women | |
Age	Number	%	Number	%
Between 6 and 16	22,260	4.6	18,670	1.9
Between 17 and 25	23,910	4.9	90,184	9.4
Between 26 and 50	90,803	18.6	90,184	9.4
Between 51 and 65	78,962	16.2	142,996	14.9
Between 66 and 75	104,631	21.5	233,060	24.2
Between 76 and 85	111,963	22.9	297,412	30,9
Over 86	54,913	11.3	164,967	17.2
Total	487,173	100.0	961,198	100,0

Whole population:		
Men	487,173	33.6
Women	961,198	66.4

Source: own elaboration based in *EDDES*

Information about this group is used to prepare the Disability, Severity and Care Index -*DSCI*- (Albarrán et. al 2007) for each age. This index tries to measure the personal situation in people with any disability who needs to be aided by third person. It is expressed as:

$$DSCI = \frac{\sum_{i=1}^{36} S_i f(h_i)}{k} 100$$

where:

S_i is the i-th level of disability severity. It can takes values 1, 2 or 3 for moderate, serious or extreme severity, respectively.

$f(h_i)$ is a measure of the hours per week hours of received care. We use the median of weekly hours of care related to the i-th disability level.

k is the normalization constant. It is the maximum value for the index numerator. This value is reached when all the disabilities suffered by someone are extremely severe. In that case, k will be equal to $3 \sum_{i=1}^{36} f(h_i)$

Due to the use of this constant, the index value is always between 0 and 1. Statistical information collected in *EDDES* about weekly hours has two shortcomings. The first one is that it is referred to time for caring all the disabilities suffered by someone, no matter how many of them the individual has. So, it is need to divide this amount of time amongst the

4

disabilities. The second problem is that *EDDES* does not reflects the accurate number of hours. Instead of, they are grouped into six tranches and the limits among them are 7, 14, 30, 40 and 60 hours. To solve this problem, time for care is measured by the class mark of each interval. This number is divided by the number of disabilities suffered by the handicapped person. Replicating this procedure by everyone, it can be obtained the empirical distribution for time dedicated to each disability and from it, all the main statistics can be estimated (such as mean, variance or median).

[Insert Figure 1 about here]

Results for Men   Results for Women
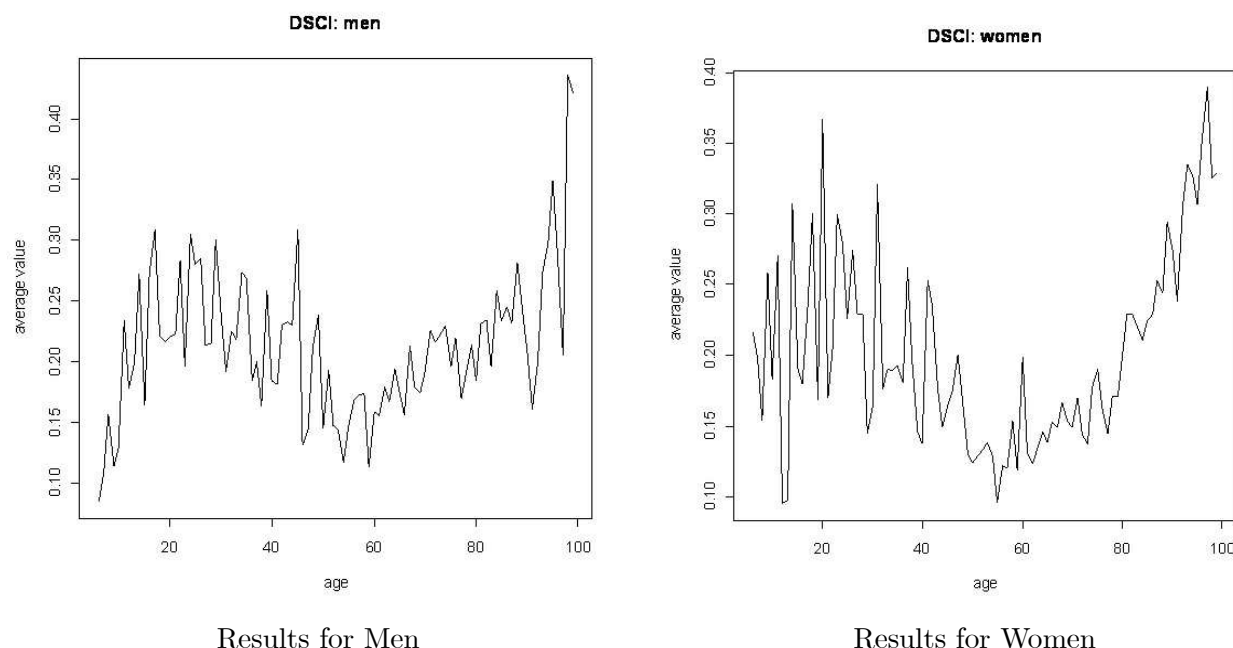
These graphs show that over lifetime, the index both increases and decreases in value. Based on this graph, we can define the following phases:

1. An ascending phase since 6 years old to a local maximum at around 20-25 years old

2. A descending phase since 20- 25 years old to a total minimum at around 50-55 years old

3. An ascending phase since 50-55 years old to the end of life at 99 years old

3   Methodological Review

Smoothing methods are used to find the functional expression of the relationship between a dependent, or response variable, and some predictor variables. Usually, linear functional functions are applied (i.e. linear regressions) to represent the relationship between dependent and independent variables. However, other techniques, as smoothing methods, provide a more flexible approach as they allow complex relationships in data.

In this paper, we focus in three representative techniques: local regression techniques (*LOESS*), Bayesian spline regression and Bayesian feed-forward neural networks. We will briefly describe them and highlight their properties.

Usually, the simplest estimation technique is to adjust a polynomial regression by raising an independent variable to consecutive powers. Polynomial adjustment offers simplicity and easiness of calculation, but it has some shortcomings. For example, if we keep adding powers, polynomial regressions can lead to a perfect multicollinearity (Marsh and Cormier, 2001). On the other hand, polynomial regression with lower powers is not flexible enough to adjust to different changes in slope.

Spline regressions are mathematical models with dummy variables that involve a set of restrictions about the continuity conditions of the function. In fact, the use of artificial variables without these continuity conditions may result in gaps (see Pindyck and Rubinfeld (1998), Smith (1979) and Suits et al., (1978)). In general, it is possible to consider spline regression as a method that uses a continuous predictor variable defined over some interval on the variable domain, and a dependent variable that is a continuous function of the predictor in all the intervals, with different slopes in each of them.

These models are commonly used when the regression line is split into segments separated by nodes. Each of these nodes shows a change in slope but there are no gaps between them. This continuity is the main difference between a spline method and a traditional artificial variable method. These models are commonly used in situations where the norms or some other phenomena cause gradual changes in the behaviour of the dependent variable. Therefore, as suggested by Pindyck and Rubinfeld (1998), the first step in the development of these models is the definition of dummy variables. Their value is zero until they reach the critical point or node, in which they change to one. The advantage of this regression over the polynomial one is its flexibility; by using low power, it prevents the probability of perfect multicollinearity of higher powers. However, the main challenge is the node detection. When it is possible to determine the position of nodes, then the estimation is constrained. When this is not possible, then nodes become additional parameters to estimate. Moreover, if we fail to find all the nodes, it is necessary to use the estimation method proposed and revised by Marsh (1983 and 1986).

Local regressions are an approximation to curves and surfaces by fitting the data through smoothing. The fit for x is the value of a parametric function fitted by observations around x (Kendall and Stuart, 1976; Stone 1977; Cleveland 1979; Hastie and Tibshirani 1990; Cleveland 1993). The underlying model in local regression is $E\left(y_i\right) = g\left(x_i\right) \ i = 1, \ldots, n$, where y_i are observations of the dependent variable and $\mathbf{x}_i$ is the set of independent variables. In general, the distribution of y_i is unknown, and it is necessary to set some hypotheses related to g and the distribution of y_i. For example, it is commonly assumed that y_i has a constant variance and a local approximation of g is performed by a polynomial function of some degree.

The use of local regressions is fairly advantageous. According to Hastie and Loader (1993), these methods show the following strengths:

(***i***) They properly solve bias problems in high curvature regions.

(***ii***) They are easy to understand and interpret.

(***iii***) Some methods permit quick calculation of one or more variables.

(***iii***) They are simple enough to adjust to different hypotheses related to distributions.

(***iv***) They do not require regularity and smoothness as other methods do, such as boundary kernel.

(***v***) If they use ordinary least squares as an estimation method, the estimates are lineal in the dependent variables.

For local adjustment, it is necessary to define the following elements: polynomial degree, weighing function, bandwidth and fitting criteria. The first three elements depend on the hypotheses related to g, while the last one depends on the hypotheses related to y_i. To choose these elements we do not need any previous knowledge about them, as it is sufficient to perform data analysis with graphic diagnosis tools such as coplots, residual-dependence graphs or residual quantiles (Cleveland 1993).

Several methods can be applied for the bandwidth selection. The most commonly used are fixed bandwidth and closest neighbour. Local regressions use closest neighbour bandwidth (Cleveland 1979; Stone 1977), since fixed bandwidth often provides estimators with significant changes in variance, due to big changes in data density.

Finally, the selection of fitting criteria depends on the distribution of y_i. For the normal distribution, which is the simplest case, the main problem is the lack of robustness and sensitivity of results to the residual distributions in fat tails. In this case, Cleveland and Devlin (1988), Tsybakov (1986), and Cleveland (1979) suggest the use of locally weighted and robust locally weighted regressions.

Another option (see Hjort (1995)), is to use Bayesian adaptive regression splines ($BARS$) based on a reversible jump technique. It can fit curves with irregular variation and it was firstly introduced in DiMatteo et al., (2001), where they showed that $BARS$ could reduce the mean squared error below other existing methods. This kind of regression uses data $y_1, \ldots, y_p$ obtained at $t = t_1, \ldots, t_p$, with each y_j is assumed to depend probabilistically on $f(t_j)$ with f being a spline having an unknown number of knots at unknown locations $x_{i_1}, \ldots, x_{i_k}$. The previous model includes a vector of nuisance parameters x_i to indicate generality. $BARS$ is a Bayesian $MCMC$-based algorithm that searches for an optimal knot set and it generates a posterior distribution on the space of splines, obtaining an improved spline estimate based on model averaging. DiMatteo et al., (2001) showed that $BARS$ has a good behavior with respect to other non-parametric methods, and it has been used in biological and medical applications (see, e.g. Kass et al., (2003) and Kaufman et al., (2005)).

Finally we consider, as a third non-linear technique, feed forward neural networks. They are a semi-parametric alternative to regression modeling, where there may be a complex relationship between the explanatory and dependent variables. Neural nets have very interesting properties related with a great adaptability to noisy observations and non-linear modeling (see Bishop (1995)).

Most work in feed forward neural networks is based on results that assure that these models are universal approximators as Cybenko (1989) points out. The basic model of a feed forward

neural network, where a logistic activation function is used, can be expressed as,

$$y_i = b_0 + b_1 x + \sum_{k=1}^{M} \beta_k \frac{1}{1 + \exp\{-\gamma_{k0} - \gamma_k x_i\}}.$$

In this paper, we will use a Bayesian approach to deal with the inference tasks. Bayesian inference for neural network models is a fairly well developed topic (for general reviews see Neal (1996) and Lee (2004)). The parameters of the model are estimated, by using a simulation-based procedure with a $MCMC$ general algorithm (for a general description see Muller and Ríos-Insua (1998) and Lee (2004)).

4   Results

Once the index used to assess the relationship between age and disability hs been introduced, this section is focused in evaluating the effect of age. First of all, we will study if the intensity of disabilities is the same or not in men and women. To do this task, $BARS$ procedure will be used.

The $BARS$ method is used under two circumstances. The first one is referred to the level in the index in men and women. As sample sizes are not small, we can use a Gaussian approximation to apply a Gaussian process test (see Behseta and Kass (2005)), obtaining an asymptotic chi-square distribution when the null hypothesis of equality of functions is true. In our data a $p\text{-}value$ equal to 3.060845e-05 is obtained. Therefore, equality between the two curves is not accepted.

In order to get a higher degree of smoothing, this technique has also been used in the first difference of the estimated $DSCI$ in both genders. As before, it is obtained an asymptotic chi-square distribution when the null hypothesis of equality of functions is true. In the first difference data a $p\text{-}value$ close to 0 is obtained. In this case, therefore, equality between these two curves is not accepted either.

Once the difference between genders has been checked, the next step is the estimation of the impact of age on the intensity of disability. Two different techniques have been used: $LOESS$ regression (Cleveland, 1979) and feed forward neural networks.

Estimation with $LOESS$ is done according to the following steps. The first one is to estimate the parameter that controls the degree of smoothing, f. Afterwards, index values are estimated with $LOESS$ using a linear function in age. Furthermore, $LOESS$ is used again on a second regression of the differences estimated values on age, in order to get a higher degree of smoothing. Obviously, it is needed again to assess the optimal value of f for this new series. The results are shown in table 2.
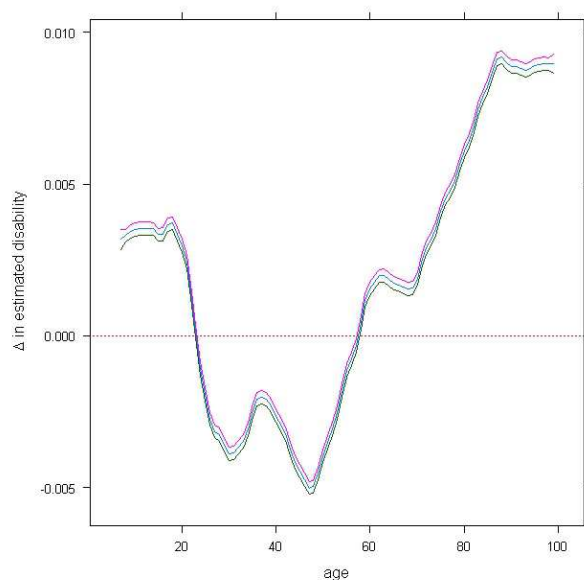
Table 2: Smoothing parameter and standard deviation using $LOESS$

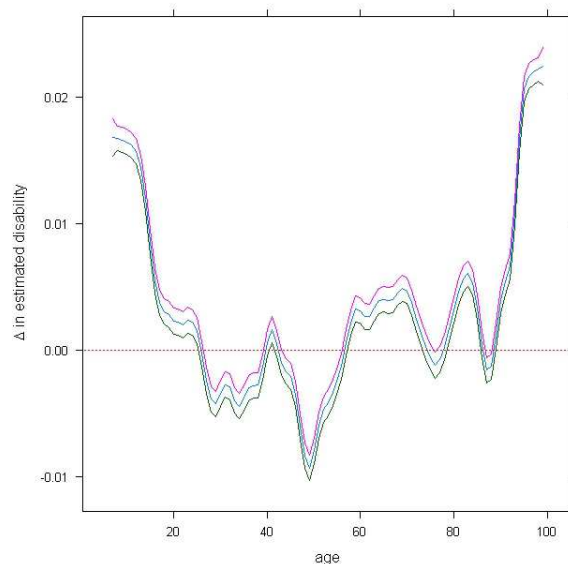|	$DSCI$		$\Delta DSCI$	
	f	$\hat{\sigma}(\%)$	f	$\hat{\sigma}(\%)$
men	.15425	4.0744	.05915	0.0856
women	.27125	4.2128	.05915	0.0185

and the estimated values for $\Delta DSCI$ are shown in Figure 2:

[Insert Figure 2 about here]

Results for Women   Results for Men

With regards to the use of neural networks, the relationship between age and $DSCI$ are estimated with a general non-linear model consisting in two parts: a linear part plus a one-layer feed forward neural network. Although feed forward neural nets can uniformly approximate any continuous function (see Cybenko (1989)), when functions are near linearity estimators in neural nets converge slowly, and it is better to introduce a linear term (see e.g. Muller and Rios-Insua (1998) and Lee (2004)).

In this case the feed-forward neural net model with a linear term can be expressed as

$$ y_i = b_0 + b_1 x + \sum_{k=1}^{M} \beta_k \frac{1}{1 + \exp\{-\gamma_{k0} - \gamma_k x_i\}}. $$

where M is the number of nodes. This term has a relevant effect over the smoothness of the model.

We use a Bayesian approach to estimate the parameters of the model (for a revision of the technical aspects see Lee (2004)). In order to program the model we have used `WinBugs` (see Lunn et al., (2000)) (the code is available from the authors, upon request). We consider, as the prior distributions of parameters normal distributions with zero means and high variances.

We use data from a previous official questionnaire from year 1986, to find out a sensible prior distribution of the parameters of the models, related with the official data from the last

9

year. The number of nodes, M, has been chosen according to the criteria of minimum deviance information (DIC) which is implemented in `WinBugs`. The preferred model has been that of $M = 2$ nodes. Results for that criteria and sum of square errors (SSE) are collected in Table 3:

Table 3: DIC and SSE in Neural Network models

Nodes	1	2	3	4	5	6
DIC	227.432	116.488	142.117	167.839	125.382	127.571
SSE	58.27759	53.03188	54.83788	53.80078	54.69006	53.48059

Source: own elaboration

The corresponding two curves derived from men and women are shown in Figure 3
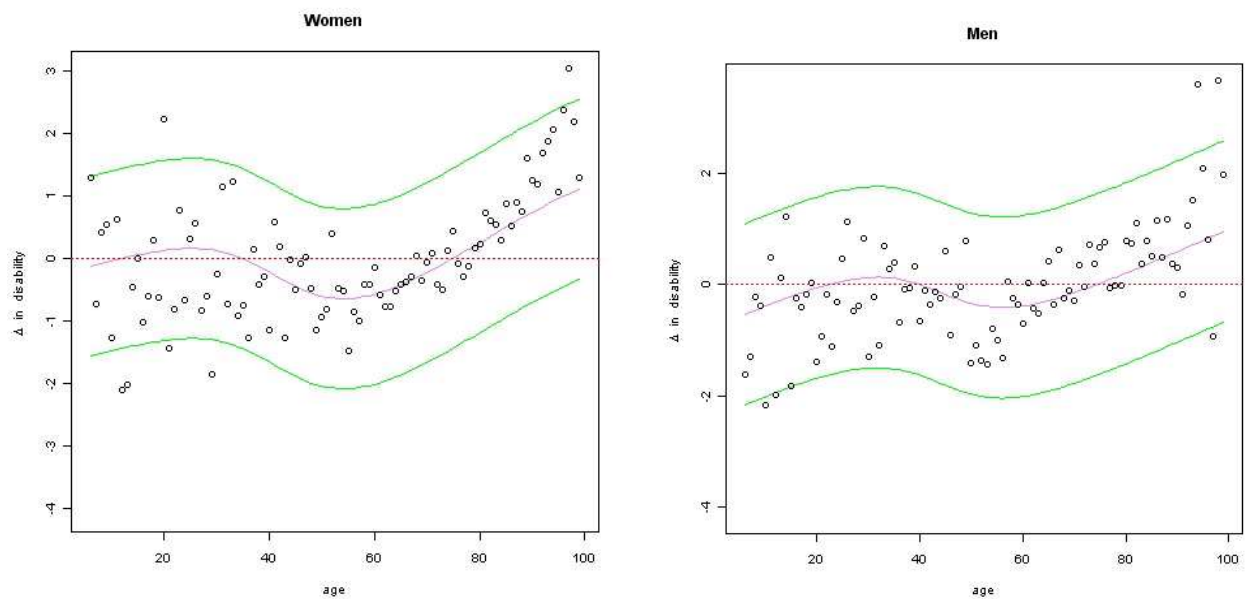
Results for Women   Results for Men

As in $LOESS$ estimation, a neural network model is estimated for increased in disability in each gender. The results for this case are shown in tables 4 and 5:

Table 4: estimated parameters for the neural net model applied for first difference Men data

	mean	std. dev.	2.5% HPD	median	97.5% HPD
b_0	-0.15	0.03	-0.19	-0.14	-0.10
b_1	0.07	0.02	0.04	0.07	0.10
β_1	-0.29	0.04	-0.38	-0.28	-0.23
β_2	1.49	0.08	1.35	1.48	1.65
γ_1	-1.62	0.07	-1.75	-1.62	-1.50
γ_2	-0.69	0.04	-0.75	-0.69	-0.62
γ_{01}	-0.60	0.16	-0.85	-0.63	-0.30
γ_{02}	-1.59	0.05	-1.68	-1.59	-1.49

Table 5: estimated parameters for the neural net model applied for first difference Women data

	mean	std. dev.	2.5% HPD	median	97.5% HPD
b_0	-0.37	0.02	-0.41	-0.38	-0.32
b_1	0.14	0.02	0.11	0.15	0.17
β_1	0.99	0.12	0.71	1.02	1.19
β_2	0.54	0.06	0.40	0.55	0.63
γ_1	-0.63	0.06	-0.75	-0.63	-0.48
γ_2	-0.35	0.07	-0.48	-0.36	-0.19
γ_{01}	-1.76	0.19	-2.06	-1.80	-1.36
γ_{02}	-0.32	0.03	-0.37	-0.32	-0.26

The corresponding curves for first differences derived for men and women are those reflected in Figure 4:

[Insert Figure 4 about here]

Results for Women

Results for Men
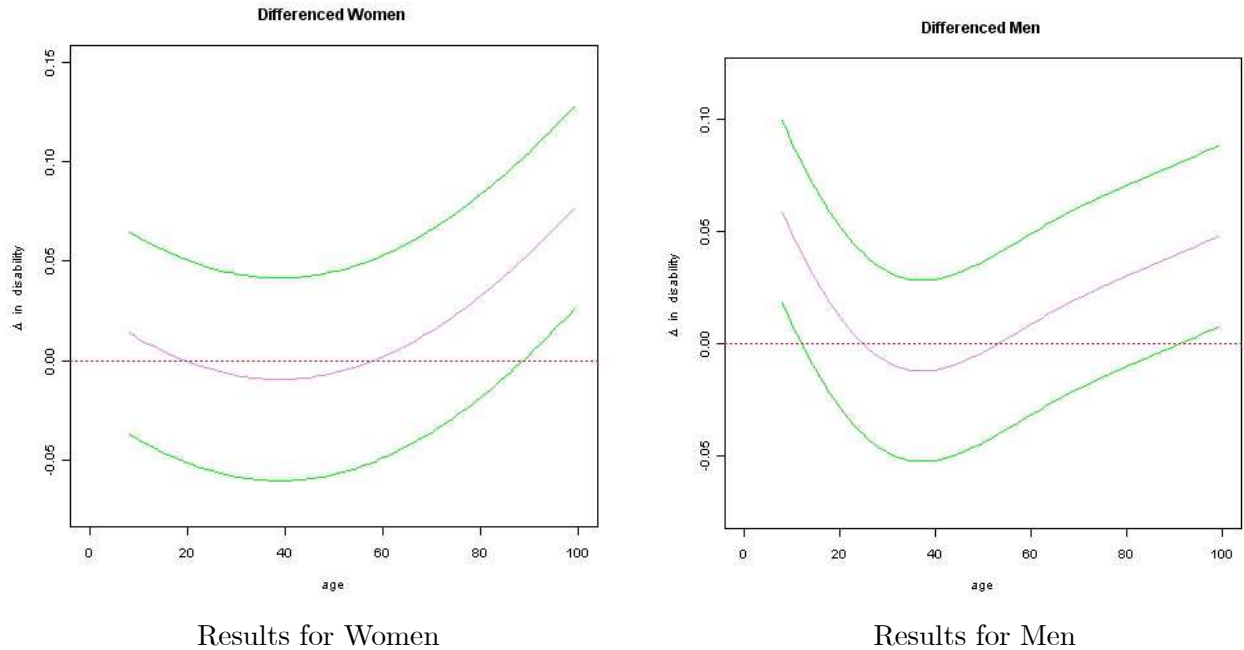
5   Concluding Remarks

The results obtained with estimations suggest that there exist a non-linear relationship between age and level of disability. However, the estimated parameters are quite sensitive to the used methodology, *LOESS* or Neural Network and the parameters estimated in every case are different for each gender. The intensity of changes in the dependent variable is greater with feed-forward neural networks. For instance, the expected increase in *DSCI* in men during the last years of lifetime is around 2% using *LOESS*, whereas it is near 5% with neural networks. In the female case, these increases are 1% and 7% using *LOESS* and neural networks, respectively.

A common place in literature is the assessment that disability increases with age. However, the results suggest that this situation can not be always expected. For men and women, the estimations reflect that a decrease in measured disability can be reached between ages over 20 and under 60 years old, regardless of the method used in estimations. The resulting profile is similar for both genders: it can be seen an increase from the beginning -6 years old- till certain age. After this moment, there is a period in which the index decreases and finally, a new increase is experienced in the index. This last span finishes at the end of the lifetime. This evolution es similar in both, men and women, when Neural Network are used. However, it is needed to do a little qualification about this behaviour when men are estimated with *LOESS* . In this case, the downward period is broken in two smaller intervals because there exists a local maximum in increase $\Delta IDSC$ near 40 in which the dependent variable is positive although close to zero. The final increase till the deathtime begins about 60 but, as in the former interval, this last increasing period is divided into several smaller intervals, because the increased index experiences two little decreases around 80 and 90. Obviously, after that age, it can be seen a continuous upward process in the estimated values of the increased *DSCI*, as in

the rest of estimations. The age at which the increase in *DSCI* reaches the expected minimum value is in a range between 40 and 50 years old, depending on the method used for estimation. This age is reached before in men than in women when Neural Network is used. In this case, there no exists significative differences due to gender.

Bibliography

Albarrán I., Alonso P. and Fajardo M.A. (2007). Valoración global de la discapacidad. Propuesta de un índice y su aplicación a la población española recogida en la EDDES. *Estudios de Economía Aplicada*, 25(2), 523–549.

Avlund K., Damsgaard M.T. and Schroll M. (2001). Tiredness as determinant of subsequent use of health and social services among non-disabled elderly people. *J. Aging Health*, 13(2), 267–286.

Behseta S. and Kass R.E. (2005). Testing Equality of Two Functions using BARS. *Statistics in Medicine*, 24, 3523–3534.

Bishop C.M. (1995). *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford.

Bishop M. (2005). Quality of Life and psychosocial adaptation to chronic illness and disability: Preliminary analysis of a conceptual and theoretical synthesis. *Rehabil. Couns. Bull.*, 48, 219–31.

Clark D.O., Stump T.E. and Wolinsky F.D. (1997). A race and gender specific replication of five dimensions of functional limitation and disability. *J. Aging Health*, 9, 28–42.

Cleveland W.S. (1979). Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.*, 74, 829–836.

Cleveland W.S. (1993). *Visualizing data*. Hobart Press, N. York.

Cleveland W.S. and Devlin S.J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *J. Am. Stat. Assoc.*, 83, 596–610.

Crimmins E.M. and Saito Y. (1993). Getting better and getting worse: Transitions in functional status among older Americans. *J. Aging Health*, 5(1), 3–36.

Cybenko G. (1989). Approximation by superposition of sigmoidal functions. *Mathematics of Control Systems an Signals*, 2, 303–314.

DiMatteo I., Genovese C.R. and Kass R.E. (2001). Bayesian curve-fitting with free-knot splines. *Biometrika*, 88, 1055–1071.

Ding J. (2009). Population with intellectual disability based on 2000–2007 national registers in Taiwan: Age and gender. *Research in Developmental Disabilities*, 30, 294–300.

Ferrer M., Lamarca R., Orfila F. and Alonso J. (1999) Comparison of performance based and self-rated functional capacity in Spanish elderly. *Am. J. Epidemiol.,* 149(3), 228–235.

Fleishman J.A., Spector W.D. and Altman B.M. (2002). Impact of Differential Item Functioning on Age and Gender Differences in Functional Disability. *J. of Gerontology,* 57(2), 275–284

Fried L.P. and Bush T.L. (1988). Morbidity as a focus of preventive health care in the elderly. *Epidemiol. Rev.,* 10, 48–64.

Fried L., Ettinger W., Lind B., Newman A. and Gardin J. (1994). Physical disability in older adults: A physiological approach. *J. Clin. Epidemiol.,* 47, 747–760.

Guralnik J.M. (2005). The evolution of research on disability in old age. *Aging Clin. Exp. Res.,* 17(3), 165–167.

Hastie T. and Tibshirani R.J. (1990). *Generalized Additive Models.* Chapman and Hall, London.

Hastie T. and Loader C. (1993). Local regression: automatic kernel carprentry (with discussion). *Statistical Science*, 8, 120–143.

Hjort, N.L. (1995). Bayesian approaches to semiparametric density estimation (with discussion). *Bayesian Statistics,* V, 223–253, Oxford University Press.

Hogan D.B., Ebly E.M. and Fung T.S. (1999). Disease, disability, and age in cognitively intact seniors: results from the Canadian Study of Health and Aging, *J. of Gerontology,* 54(2), 77–82.

Institute of Medicine (1991). *Disability in America: Toward a National Agenda for Prevention.* Washington, DC.

Jagger C., Spiers N.A. and Clarke M. (1993). Factors associated with decline in function, institutionalization and mortality of elderly people. *Age and Ageing*, 22(3), 190–197.

Jamoom E.W., Horner-Johnson W., Suzuki R., Andresen E.M., Campbell V.A. and the RRTC Expert Panel on Health Status Measurement (2008). Age at disability onset and self-reported health status. *BMC Public Health,* 8, 10.

Jiménez A. and Huete, A. (2003). La Discapacidad en España: Datos Estadísticos Madrid. *Real Patronato sobre discapacidad.*

Johnson R.J. and Wolinsky F.D. (1994). Gender, race, and health: The structure of health status among older adults. *The Gerontologist*, 34, 24–35.

Kass, R.E. Ventura V. and Cai, C. (2003). Statistical smoothing of neuronal data. *Network: Computation in Neural Systems*, 14, 5–15.

Katz S., Ford A.B., Moskowitz R.W., Jackson B.A. and Jaffe M.W. (1963). Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. *J. of the American Medical Association*, 185, 914–919.

Kaufman C.G., Ventura V. and Kass, R.E. (2005). Spline-based nonparametric regression for periodic functions and its application to directional tuning of neurons. *Statistics in Medicine*, 24, 2255–2265.

Kemp B. (1999). Quality of life while aging with a disability. *Assist. Technol.,* 11, 158–163.

Kendall M.G. and Stuart A. (1976). *The Advanced Theory of Statistics*, Vol. 3. Hafner, N. York.

Lamarcaa R., Ferrera M., Andersenb P.K., Liestolc K., Keidingb N. and Alonso J. (2003). A changing relationship between disability and survival in the elderly population: differences by age. *J. Clin. Epidemiol.,* 56, 1192–1201.

Laukkanen P., Leskinen E., Kauppinen M., Sakari-Rantala R. and Heikkinen E. (2000). Health and functional capacity as predictors of community dwelling among elderly people. *J. Clin. Epidemiol.,* 53(3), 257–265.

Lawton M.P. and Brody E.M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist*, 9, 179–186.

Lee H.K.H. (2004). *Bayesian nonparametrics via neural networks*. Asa-Siam, Philadelphia.

Leveille S.G., Penninx B.W., Melzer D., Izmirlian G. and Guralnik J.M. (2000). Sex differences in the prevalence of mobility disability in old age: the dynamics of incidence, recovery, and mortality, *J. of Gerontology,* 55(1), 41–50.

Lunn D.J., Thomas A., Best N. and Spiegelhalter D. (2000). WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10, 325–337.

Mahoney F.I. and Barthel D. (1965). Functional evaluation: the Barthel Index. *Maryland State Medical Journal*, 14, 56–61.

Manton K.G., Stallard E. and Corder L.S. (1998). The dynamics of dimensions of age-related disability 1982 to 1994 in the U.S. elderly population. *J. of Gerontology*, 53(1), 59–70.

Manton, K.G. (1988). A longitudinal study of functional change and mortality in the United States. *J. of Gerontology*, 43, 153–175.

Marsh L.C. (1983). On estimating spline regressions. *Proceedings of SAS Users Group International*, 8, 723–728.

Marsh L.C. (1986). Estimating the number and location of knots in spline regressions. *J. of Applied Business Research*, 3, 60–70.

Marsh L.C. and Cormier D.R. (2001). *Spline regressions models.* Sage University Paper.

Merrill S.S., Seeman T.E., Kasl S.V. and Berkman L.F. (1997). Gender differences in the comparison of self-reported disability and performance measures. *J. of Gerontology*, 52, 19–26.

Morris K. and Doyle P. (2008) Age-related disability: etiology and evidence of resilience. *The Gerontologist*, 48, 655–656.

Muller P. and Ríos-Insua D. (1998). Issues in Bayesian Analysis of Neural Network Models. *Neural Computation,* 10(3), 749–770.

Neal R.M.(1996). *Bayesian Learning for Neural Networks. Lecture Notes in Statistics No. 118.* Springer-Verlag, New York.

Pindyck, R.S. and Rubinfeld D.L. (1998). *Econometric models and econometric forecasts (4th ed).* Irwin/McGraw-Hill, N. York.

Ruigomez A., Alonso J. and Anto J.M. (1993). Functional capacity and five-year mortality in a sample of urban community elderly. *Eur. J. Public Health,* 3, 165–171.

Scott W.K., Macera C.A., Cornman C.B. and Sharpe P.A. (1997). Functional health status as a predictor of mortality in men and women over 65. *J. Clin. Epidemiol.,* 50(3). 291–296.

Smith P.L. (1979). Splines as a useful and convenient statistical tool. *The American Statistician*, 33(2), 57–62.

Spector W.D. and Fleishman J.A. (1998). Combining activities of daily living with instrumental activities of daily living to measure functional disability. *J. of Gerontology*, 53(1), 46–57.

Stone R.I. and Murtaugh C.M. (1990). The elderly population with chronic functional disability: implications for home care eligibility. *The Gerontologist*, 30(4), 491–496.

Stone C.J. (1977). Consistent nonparametric regression (with discussion). *The Annals of Statistics*, 5, 595–620.

Suits D.B., Mason A. and Chan L. (1978). Spline functions fitted by standard regression methods. *Review of Economics and Statistics*, 60, 132–139.

Trividi M.H., Rush A.J., Wisniewski S.R., Warden D., McKinney W., Downing M., Berman S.R., Farabaugh A., Luther J.F., Nierenberg A.A., Callan J.A., Sackeim H.A. (2006). Factors associated with health-related quality of life among outpatients with major depressive disorder: a STAR*D report. *J. Clin. Psychiatry*, 67, 185–195.

Tsybakov A.B. (1986). Robust reconstruction of functions by the local approximation method. *Problems of Information Transmission*, 22, 69–84.

University of Jyväskylä (2004). *Disability in old age, Final Report: Conclusions and Recommendations*, Finland.

Verbrugge L. (1985). Gender and health: An update on hypotheses and evidence. *J. of Health and Social Behavior*, 26, 156–182.

Verbrugge L. (1989). The twain meet: Empirical explanations of sex differences in health and mortality. *J. of Health and Social Behavior*, 30, 282–304.

Wallstrom G., Liebner J., and Kass R.E. (2008). An implementation of Bayesian Adaptive Regression Splines (BARS) in C with S and R wrappers. *J. of Statistical Software*, 26.

WHO (1997). *International Classification of Impairments, Disabilities and Handicaps (ICIDH)*. IMSERSO, Madrid.

Wingard D. (1984). The sex differential in morbidity, mortality and lifestyle. *Annual Review of Public Health*, 5, 433–458.

Wolinsky F.D., Callahan C.M., Fitzgerald J.F. and Johnson R.J. (1993). Changes in functional status and the risks of subsequent nursing home placement and death. *J. of Gerontology*, 48, 94–101.

Wolinsky F.D., Culler S.D., Callahan C.M. and Johnson R.J. (1994). Hospital resource consumption among older adults: a prospective analysis of episodes, length of stay, and charges over a seven-year period. *J. of Gerontology*, 49(5), 240-252.

Wray L.A. and Blaum C.S. (2001). Explaining the Role of Sex on Disability: A Population-Based Study. *The Gerontologist*, 41(4), 499–510.