UNIVERSIDAD CARLOS III DE MADRID
working papers

Working Paper 05-25
Economics Series 15
April 2005

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624 98 75

DO TRAINING PROGRAMMES GET THE UNEMPLOYEDBACK TO WORK?: A LOOK AT THE SPANISH EXPERIENCE*

F. Alfonso Arellano[1]

Abstract

This study analyzes the effect of some training courses for economic disadvantaged unemployed workers elaborated by the Spanish Department of Employment (INEM) on exit rate from unemployment. Two groups of Spanish unemployed workers are compared between April 2000 and February 2001, one of them did training courses in the first quarter of 2000. Non-parametric techniques, parametric and semi-parametric continuous time duration methods are used to analyze this relationship. The results suggest a higher positive effect of some training courses for women than for men, especially with medium level courses. The lower the age and the period of active labour demand are, the higher exit rate to a job is. However, education and disabilities do not affect significantly the exit rate to employment.

JEL Codes: J24, J64

Keywords: training programmes, unemployment, non-parametric methods,

***Acknowledgements**. I would like to thank to César Alonso and Juan José Dolado for the supervision in this paper, participants in seminar at CEMFI, the VI IZA Summer School in Labour Economics and the XXVIII Simposio de Análisis Económico. Also thanks to Almudena Durán and Antonio Hernando (INEM) for giving to me the data bases used in this work. Financial support from predoctoral fellowship AP2000-0853 of the Spanish Ministry of Education is gratefully acknowledged. The usual disclaimer applies.

[1] **Address for correspondence:** F. Alfonso Arellano. Department of Economics. Universidad Carlos III de Madrid, C/ Madrid, 126, 28093 Getafe, Madrid, Spain. E-mail: farellan@eco.uc3m.es.

1 Introduction

Governments spend great amounts of resources, basically from taxes, to develop social programmes and other public activities. Different groups improve and worsen with such programmes. The study of these profits and losses plays an important role on the public decision taking.

For analytical and policy purposes, the OECD splits this spending into so-called **active** and **passive** labour market measures. The active measures comprise a wide range of policies aimed at improving the access of the unemployed workers to the labour market and jobs, job-related skills and the functioning of the labour market and the passive measures relate to spending on income transfers.

Public expending on active labour market programmes absorbs significant resources in most of OECD countries. It supposes more than a third of the total resources dedicated to unemployment benefits, and it exceeds such benefits in some countries. Table 1 shows a wide variation across EU countries in the share of the main categories of labour market programmes. The OECD data base covers five main categories of these programmes: Public employment services and administration, youth measures, subsidized employment, measures for the disabled and labour market training[1]. This last category constitutes one of the most attractive (and expensive) public interventions.

	Austria	Belgium	Denmark	Finland	France	Germany	Greece	Netherlands	Portugal	***Spain***	Sweden	U.K.
1. Public employment services and administration	0,14	0,17	0,12	0,12	0,18	0,23	0,06	0,26	0,11	**0,09**	0,23	0,13
2. Labour market training	0,20	0,24	0,85	0,29	0,25	0,34	0,21	0,31	0,15	**0,14**	0,30	0,05
3. Youth measures	0,03	-	0,10	0,16	0,42	0,09	0,10	0,04	0,22	**0,06**	0,02	0,15
4. Subsidised employment	0,11	0,77	0,17	0,29	0,37	0,25	0,08	0,38	0,09	**0,40**	0,24	0,01
-Hiring subsidies	0,06	0,27	0,02	0,15	0,18	0,03	0,05	0,05	0,01	**0,25**	0,19	0,01
5. Measures for the disabled	0,06	0,12	0,33	0,09	0,09	0,29	0,01	0,58	0,04	**0,03**	0,31	0,02
Active measures (from 1 to 5)	0,53	1,30	1,56	0,95	1,31	1,20	0,46	1,58	0,61	**0,73**	1,09	0,36
Passive measures (*)	1,07	2,18	3,00	2,02	1,65	1,92	0,47	1,86	0,90	**1,33**	1,19	0,56
Labour market policies	1,60	3,48	4,56	2,96	2,96	3,13	0,93	3,44	1,52	**2,06**	2,28	0,92
Labour market policies for one point of unemployment rate	0,44	0,53	1,06	0,33	0,34	0,40	0,12	1,43	0,37	***0,16***	0,45	0,18
Active policies for one point of unemployment rate	0,15	0,22	0,36	0,10	0,15	0,15	0,06	0,66	0,15	***0,06***	0,21	0,07

(*) It includes unemployment benefits and early retirement pensions for labour market reasons.
Source: OECD, Employment perspectives, June 2002

Table 1: Expenditure on labour market programmes in EU countries, 2001

[1] See Martin (2000) for further details on the public spending of labour market programmes in OECD.

Spain jointly with U.K. and Greece dedicate a small quantity of resources to (active) labour market policies, as can be inferred from the proportion of the public expenditure on active labour market policies for one percent of unemployment rate. Apart from subsidized employment, labour market training is the most important labour market programme in Spain. Taking into account these figures, the goal of the paper concerns the analysis and evaluation of a labour market training programme that the Spanish Department of Employment (INEM), or the autonomous region with the corresponding competence, carries out annually in Spain: the National Plan for Training and Professional Insertion.

The paper is organized as follows: Section 2 incorporates information about the training courses in Spain. A descriptive analysis of data base is included in Section 3. In Section 4, non-parametric techniques are applied to obtain preliminary information about data. In order to study the relationship of the variables in data, parametric and semi-parametric techniques are used in Section 5. Conclusions and extensions are given in Section 6.

2 Training courses in Spain

The National Plan for Training and Professional Insertion was included as one of the labour market actions defined by the Spanish government in 1980, but its original structure comes from the rearrangement of training programmes to lay stress on professional reinsertion of unemployed workers since 1993. This plan does not belong neither to the Educational System, which depends on the Spanish Ministry of Education, nor the training dedicated to employed workers, controlled by FORCEM[2]. Although any unemployed worker can profit from these courses, the plan includes a preferable set of groups:

- Unemployed workers who receive unemployment benefit.
- Long-term unemployed workers above 25 years (over 1 year of registered unemployment).
- Young unemployed workers (below 25 years) whose last job had a minimum duration of 6 months.
- Women who want to work.
- Disabled workers.
- Migrant workers.

The management and planning of the programmes, and preselection of candidates correspond to INEM or regional governments with this competence. The selection of individuals depends on training centres which carry out the courses. Any institution can be qualified as a training centre if some conditions are satisfied. Once a worker passes evaluation, she obtains an official professional certificate. The training courses are divided into four levels:

Type of course	Target
Course 1: Broad basis	appointed preferably to youth to provide knowledge and skills to facilitate insertion in the labour market, but these courses do not provide a specific qualification for an occupation.
Course 2: Occupation	assigned to unskilled workers, it provide knowledge and skills to hold an occupation.
Course 3: Specialization	assigned to skilled workers who need to train for a new occupation
Course 4 : Adaptation and Occupation	Improve and bring up to date knowledge such that skilled workers can be promoted to a superior job level.

Table 2: Description of training courses

[2]FORCEM is the Fundation for Continuous Training, constituted by employer's organizations and trade unions in May 1993. It takes charge of driving and spreading continuous training among firms and workers, promoting assistance, and controlling this activity.

3 Data base

I use administrative data provided by INEM. They are distributed in three data sets: (i) a file constituted by workers with active labour demand, who are controlled in three dates (31 March 2000, 30 September 2000, 31 March 2001); (ii) other file contains detail information about the workers who did the courses in the first quarter of 2000; and finally (iii) a set of contracts along the period between 31 March 2000 and 31 March 2001.

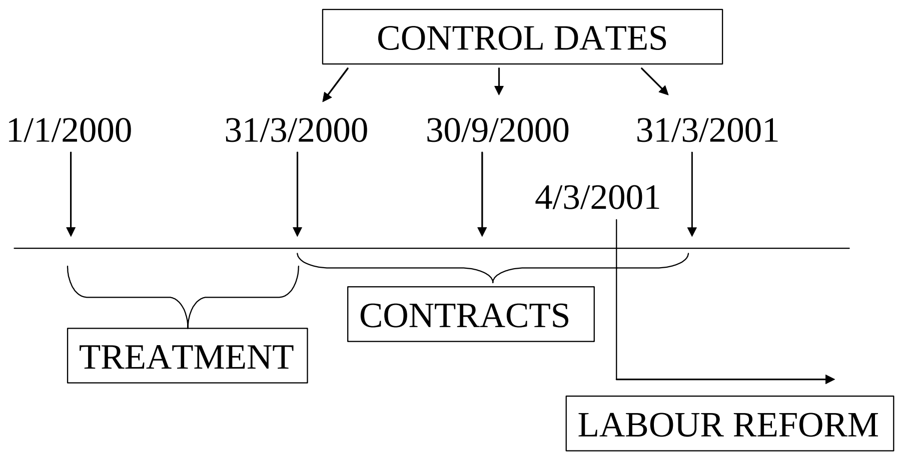

Figure 1: Evolution of data

Labour and personal information is available between 31 March 2000 and 31 March 2001 for each worker (Figure 1). The workers of the sample are supposed to begin the unemployment period at the start of this period[3]. Those workers with incomplete information in at least two consecutive control dates and the final date of the unemployment period are eliminated. I concentrate on unemployed workers at 31 March 2000, whose age does not exceed 60 years and with active labour demand period below 2,000 days. This subsample is also reduced because of some limitations from censored data[4]. Data with censorship at 30 September 2000 is eliminated. Comparing the results with and without this set of individuals, they do not vary substantially.

The final sample is constituted by 18,617 individuals, part of them (6,190 individuals) did a training course in the first quarter of 2000. The worker who did a training course is included in the so-called *treatment* group, otherwise

[3] As Ham and Lalonde (1996) consider for duration models.

[4] Considering the theory exposed by Miller (1981) and Kalbfleisch and Prentice (1980), censored data at 28 February 2001 present Type I censorship, because the experiment finished at this day. Miller (1981) considers this assumption is valid in the case of random losses to follow-up, as it is supposed in this study. Asymptotic results would be considered with Type I censorship if confidence intervals and tests were used.

the worker belongs to the *control* group. Several homogeneity conditions are accepted to avoid distortions of the estimates and identify the treatment effect appropriately. Control workers are assume not to do any training course for the treatment period and no worker of the sample did any training course before the first quarter of 2000[5]. Lastly, given the selection mechanism of the treatment group explained in the previous section, the effect of treatment coincides with the intention to be treated.

The analysis period is reduced because of the application of an important labour market reform at 4 March 2001. This labour market reform introduced urgent measures to increase and improve the employment quality, given the high use of temporary contracts[6]. The measure consisted of extending a new permanent contract with smaller dismissal costs introduced in 1997 to other groups of workers. Old permanent contracts are characterized by a severance payment of 20 days' wages per year of job tenure (up to 12 months) in the case of fair dismissals, and 45 days' wages per year of job tenure (up to 42 months) in the case of unfair dismissals. New permanent contracts present the same figures as old ones for fair dismissals, but they allow a reduction of 33 days' wages per year of job tenure (up to 24 months) in the case of unfair dismissals. In order to avoid collateral effects of this measure in the sample, the period of study is limited to 28 February 2001. The group of workers affected by this reduction is minuscule (around 1% of the sample). Given the limitations of the time interval, the conclusions derived from the training courses are constrained to the short run.

Appendix A describe the variables used in the paper. Table A1 presents a descriptive analysis considering the final sample. There exists homogeneity between the treatment and control groups, except for the worker's residence and the economic activities workers choose as their first preference. Treated workers prefer jobs related to White-collar workers and are more concentrated in Madrid than control workers, who prefer jobs associated to Catering, Protection and Sales. The difference of residence between these two groups is small for the most populated provinces.

Previous differences between the treatment and control groups are maintained when men and women are compared (Table A2). Although basic characteristics appear in both groups, women select jobs related to Services Sector and men prefer Industry. Men show a greater weight of disability, low education and living in the most populated provinces than women. The opposite relationship occurs to workers without benefits and the active labour demand period. Given the important differences between men and women, each group will be analysed and estimated separately.

[5] A less restrictive assumption is the homogeneity of the knowledge derived from the courses. Other option is that the alternative training is captured by other observed variables. The effect of the courses would be net in these cases.

[6] For a resume about the Spanish labour market reforms, see Dolado et al. (2001), Kugler, Jimeno and Hernanz (2002) and Arellano (2005).

4 Non-parametric estimates of survivor functions

4.1 Introduction

Let T be a non-negative continuous random variable representing the waiting time until the leaving unemployment. T has probability density function $f(t)$ and cumulative distribution function $F(t)$, whose complement is the **survivor function** $S(t) = 1 - F(t)$. An alternative description of the distribution of T is given by the **hazard function**, or instantaneous rate of occurrence of the event, defined as

$$h(t) = \lim_{dt \to 0} \frac{P[t \leq T < t + dt \mid T \geq t]}{dt}$$

where the numerator is the conditional probability that the individual gets a job in the interval $(t, t + dt)$ given that this event has not occurred before, and the denominator is the width of the interval. The former expression can be written as

$$h(t) = \frac{f(t)}{S(t)}$$

Given this result, $h(t)$ may be estimated using an estimate of $F(t)$ or $S(t)$. Let n_j be the population in the unemployment situation at time t_j and d_j the number of changes at t_j. The non-parametric maximum-likelihood estimate of the survivor function is (Kaplan and Meier (1958))

$$\hat{S}(t) = \prod_{j|t_j \leq t} \left(\frac{n_j - d_j}{n_j} \right)$$

Kaplan-Meier estimates are used taking into account that the moment of occurring an event corresponds to the day the worker gets a job.

4.2 Kaplan-Meier estimates

Estimates of the survivor functions are represented in the following figures of this subsection. Probability appears in the vertical axis and time in the horizontal axis. Figure 2 presents a Kaplan-Meier estimate of the survivor function from total sample and the number of spells affected by right censorship at the end of the period. The introduction of confidence intervals by Greenwood's formula does not generate important differences in the figure, because the confidence intervals are near the estimate.

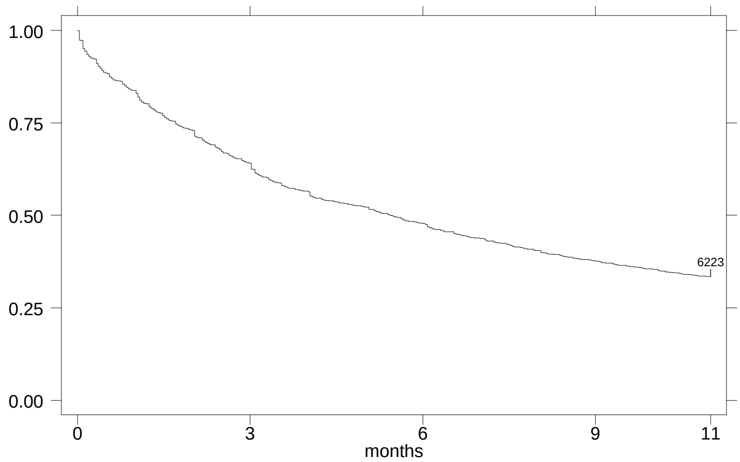

Figure 2: Kaplan-Meier estimate

The estimates may contain bias if a set of circumstances are produced, as the existence of a great amount of censoring points and the lack of independence of the sample because of implicit factors. Given that the problem of non-independent censored points was solved, it is possible to control implicit factors partially using different variables.

Considering the training courses, the control group has a higher probability of being unemployed before any moment than the treatment group, as can be seen in Figure 3:

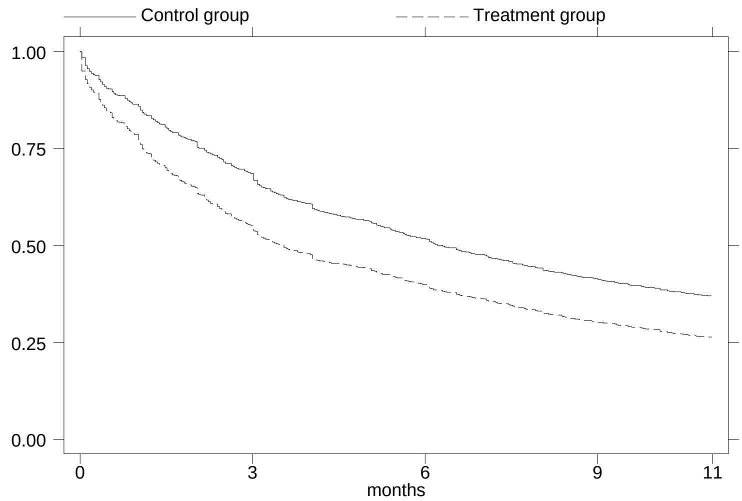

Figure 3: Treatment

With respect to gender, Figure 4 shows one of the facts of the Spanish labour market, women have more problems to get a job than men.

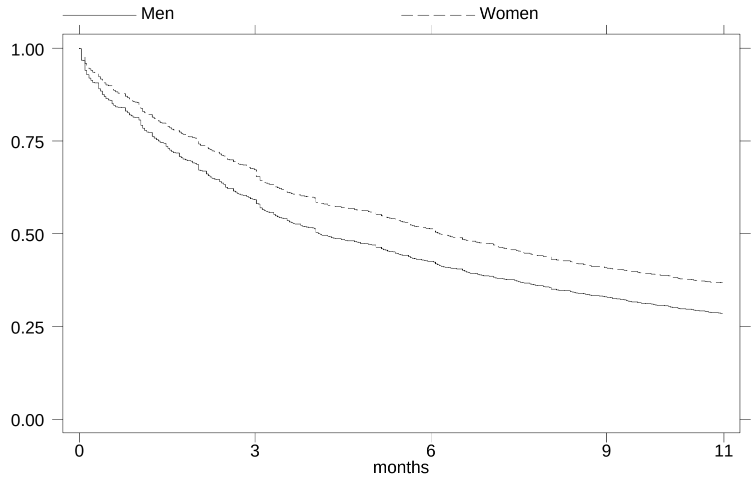

Figure 4: Gender

Distinguishing among several education levels in Figure 5, the worst groups are the less educated workers and the group with the best behaviour is associated to the most qualified Technical College (*Level* 7).

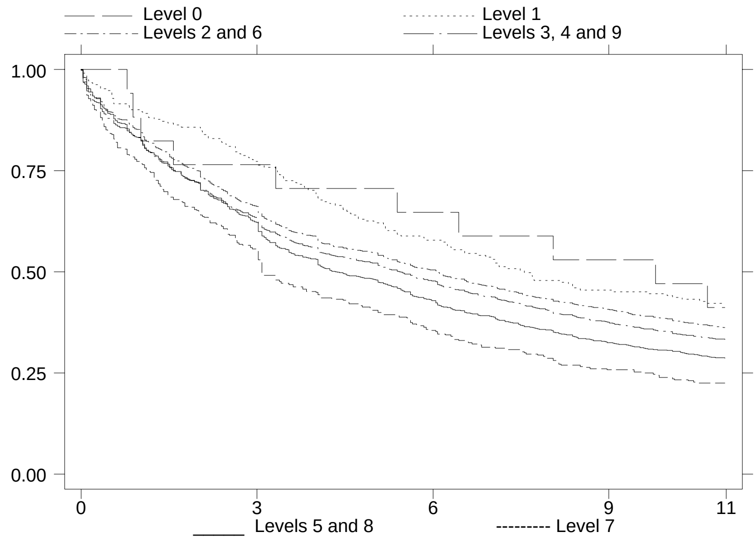

Figure 5: Education Levels

The differences for the rest of education levels are less significative. The difference between poorest levels and the rest of levels except *Level* 7 is smaller for men than for women. Men with the most qualified Technical College present better results than women with the same education level. Therefore, there is no clear linear relationship between education and the probability to abandon unemployment.

Worker's age is divided into six intervals in order to show its effect on the probability of changing the situation of unemployment appropriately (Figure 6). There exists a inverse relationship between age and the probability of leaving unemployment, except for the case of the youngest workers (a possible explanation is the lack of experience of this group).

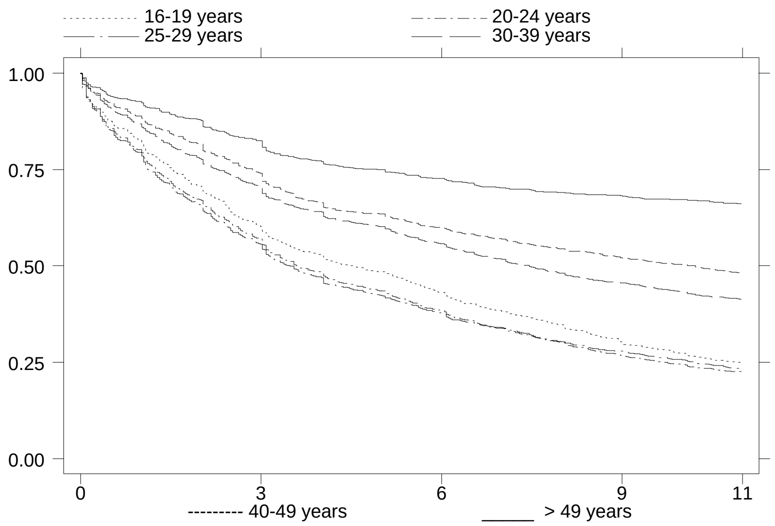

Figure 6: Age Groups

Another interesting variable is the existence of disability. Figure 7 presents the estimates distinguishing between disabled and non disabled workers. The result is in favour of the latter group.

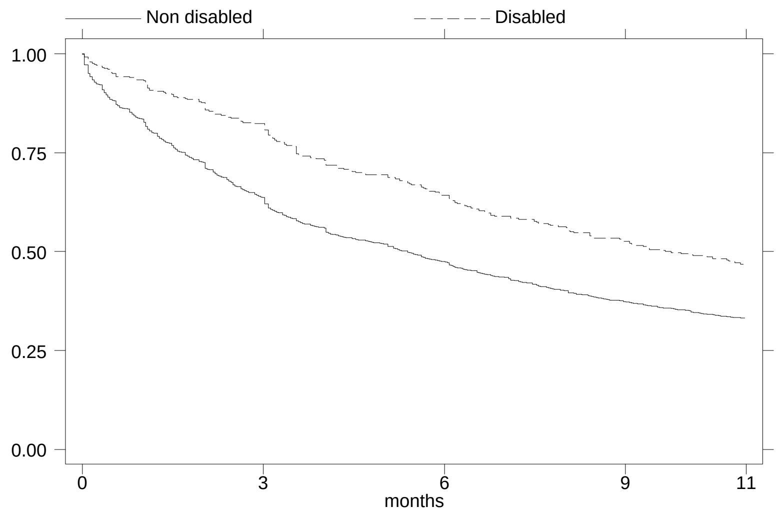

Figure 7: Existence of disability

A variable related to the possibility of receiving some kind of benefit in the moment of control is also disposal (Figure 8).

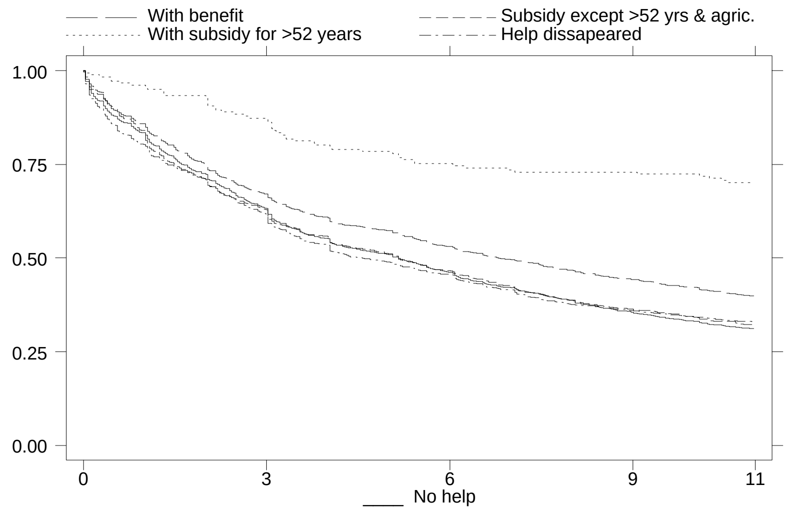

Figure 8: Benefits

The workers with the highest probability to be unemployed are those who receive benefits for workers more than 52 years. This result is consequent with the previous result of the age groups. Those workers with benefits do not have the same incentives to leave unemployment as those without help, but the differences are not highly significant among them.

It is difficult to obtain a figure for all the provinces of Spain. Figure 9 presents a classification depending on the position of each province in the figure. Although there is not a clear relationship between rich provinces and an increase of the probability of leaving unemployment, there are differences between the behaviour of Islas Baleares and *Group* 4[7] and the poor evolution of Melilla. Apart from these extreme cases, it is possible to distinguish among provinces belonging to *Group* 1, *Group* 2 and *Group* 3[8].

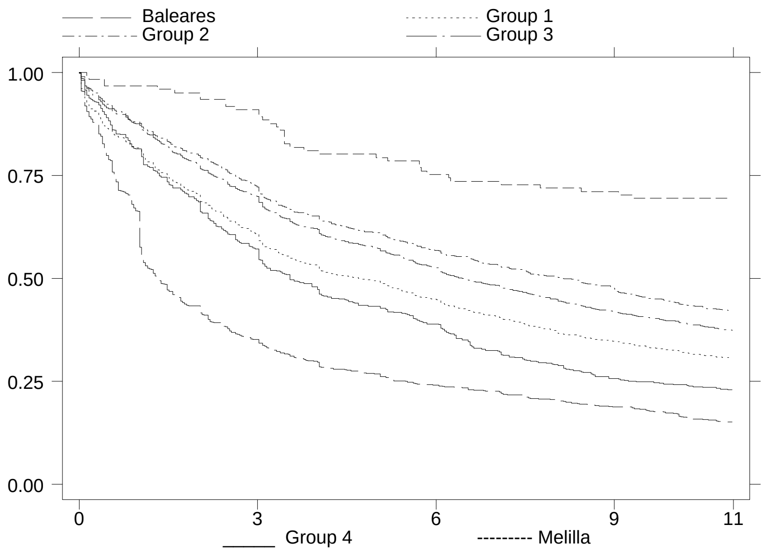

Figure 9: Provinces of Spain

Dividing the days of active labour demand into four time intervals, the positive relationship between active labour demand and the probability to be unemployed is clearly observed. The differences are higher when the active labour demand period increases (Figure 10).

[7] *Group* 4 is composed by Almería, Castellón, Guadalajara, Huesca, La Rioja, Lérida and Lugo (see map in Appendix A).

[8] *Group* 1 is composed by *provX* when X =1, 3, 5, 8, 9, 16, 17, 20, 28, 30, 31, 36, 38, 39, 43, 44 and 45 (see index in Appendix A).

Group 2 is composed by Albacete, Cádiz, Málaga, Palencia, Salamanca and Vizcaya.

Group 3 is composed by *provX* when X = 6, 10, 13, 14, 15, 18, 21, 23, 24, 32, 33, 35, 40, 41, 42, 46, 47, 49, 50 and 51 (see index in Appendix A).

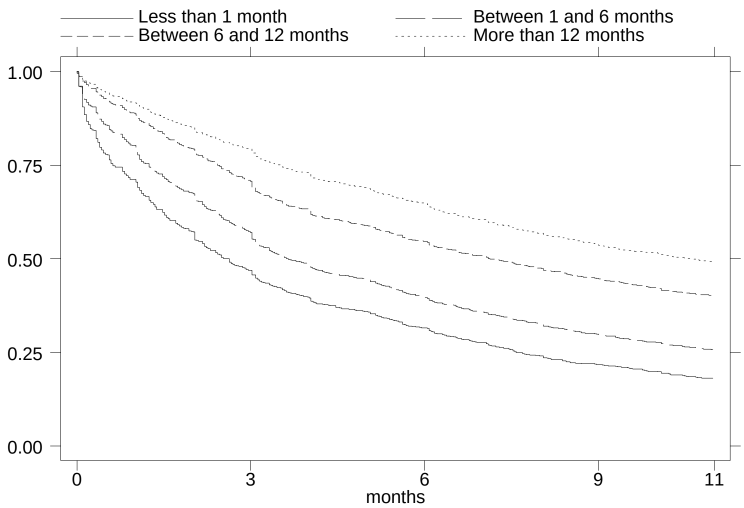

Figure 10: Duration of active labour demand

Figure 11 presents the economic activity of the job the worker applied for as first option. A previous classification of economic activities is established depending on the similarities of the survivor functions. The worst option is Management and Public Administrations (*Group* 1), because the labour supply is very small compared to the (qualified) labour demand in this economic activity. The rest of groups do not show important differences, and Skilled Workers in Farming and Fishing (*Group* 6) and Workers in Factory Industry, Construction and Mining (*Group* 7) are the best groups versus White-collar Workers (*Group* 4).

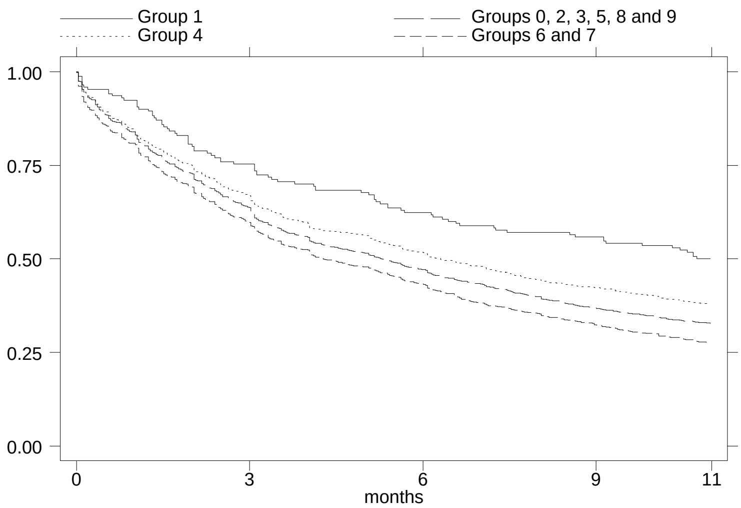

Figure 11: Economic activity of the first option

The worker's civil status can reflect how employment probability is affected by familiar topics like children, specially in the case of women. There exists an important difference between single workers and the rest of groups, increasing as the unemployment period finishes. This effect can be explained, at least partially, taking into account the relationship between civil status and age (Figure 12).

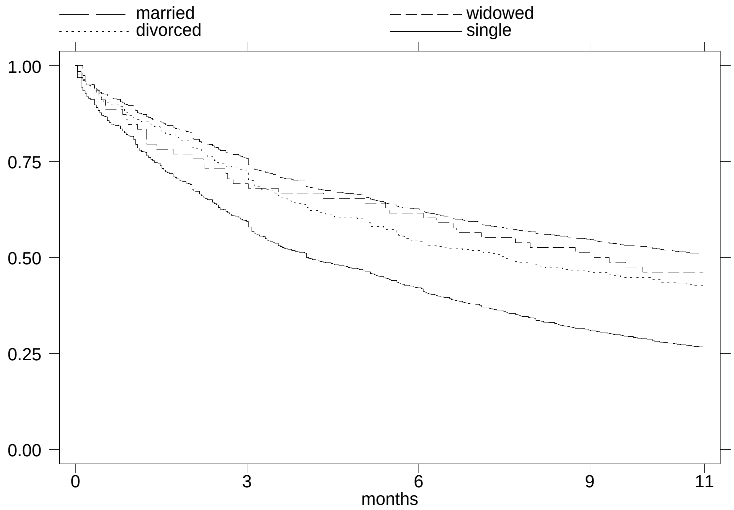

Figure 12: Civil Status

With respect to the satisfaction of the military service, only men are considered in Figure 13. The conclusion is also related to age, because workers who do not satisfy the military service (or the alternative to this compulsory condition) are usually young, so they have a greater probability to leave unemployment before.

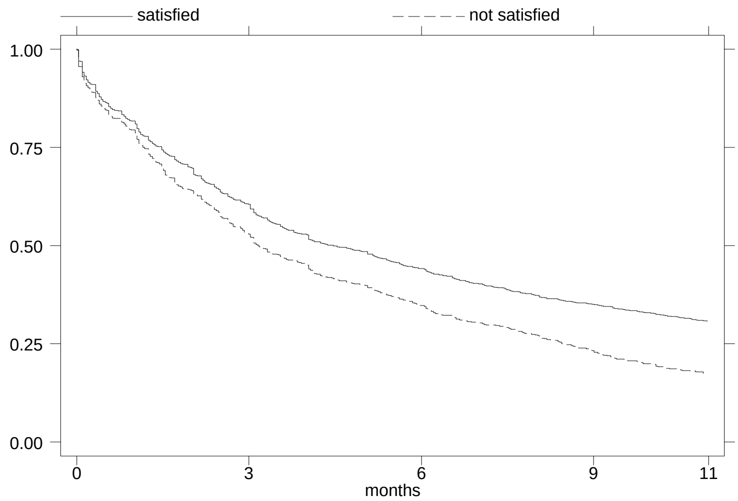

Figure 13: Military Service

I have information about the reason to leave the previous job (Figure 14). Because of the similar behaviour of the workers without reason and those who finish the temporary job duration, only one of them is included in Figure 14. The higher the worker´s involvement in the job loss is, the lower the probability to leave unemployment is. Voluntary dismissal and contract ending show better results than firing. This fact is also related to age. The mean age is 32 years, except for fired workers (mean age in this case is 39 years) and workers included in job adjustment plans, whose mean year is 45 years.

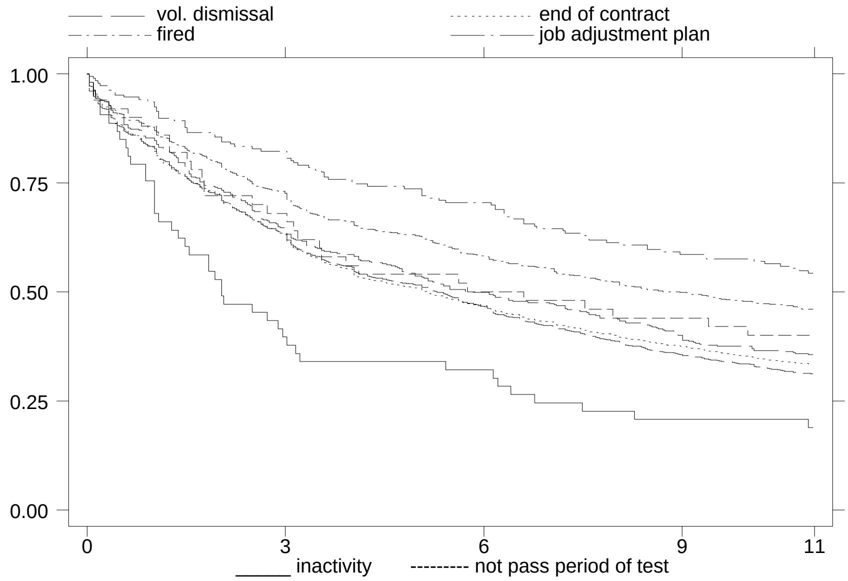

Figure 14: Causes of leaving the last job

As it can be observed in Figure 15, the highest differences are between category C and categories D and E. The difference of survivor functions between no driving license and the basic level (*Level B*1) is not important, but a high level of driving license increases the probability of getting a job before, especially for the groups of workers in the Road Transport Sector.

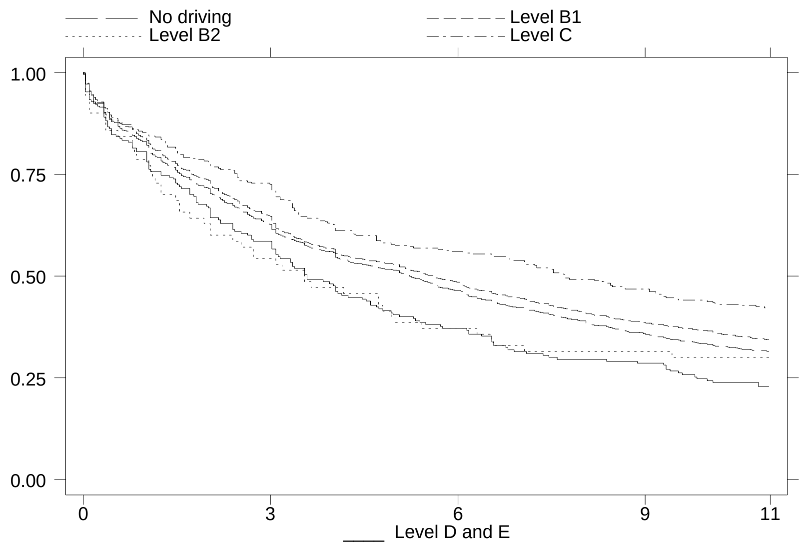

Figure 15: Driving License

Figure 16 shows idioms do not seem to be useful to get a job. They do not generate negative effects, except for French. One possible explanation is associated to the degree of knowledge of the idiom, specially in the case of French. The best result is obtained by German, because a majority of workers who know this language live in Islas Baleares. This region presents the highest unemployment exit rate (Figure 9) and it is characterized by an important German tourism.

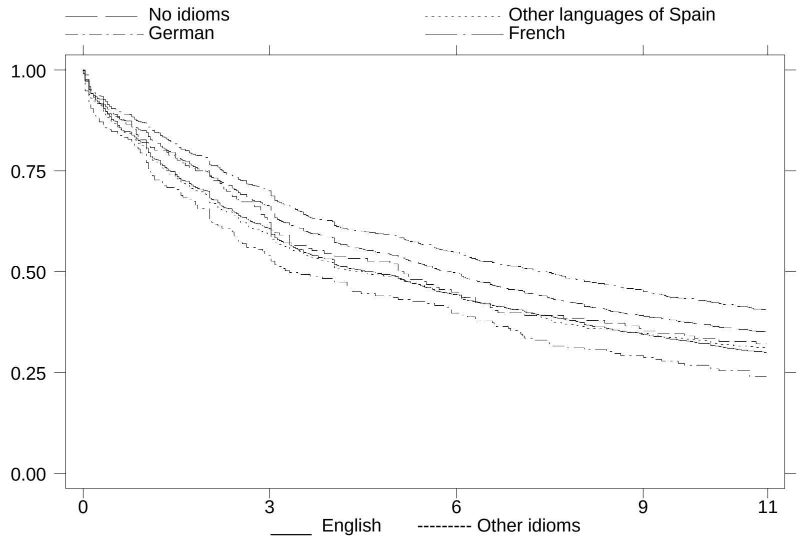

Figure 16: Knowledge of other languages

Although the competence of management of the courses does not introduce a clear and direct effect in the probability of leaving unemployment, the result is slightly better for the courses managed by INEM than those managed by regional governments. Only workers who belong to the treatment group appear in Figure 17.

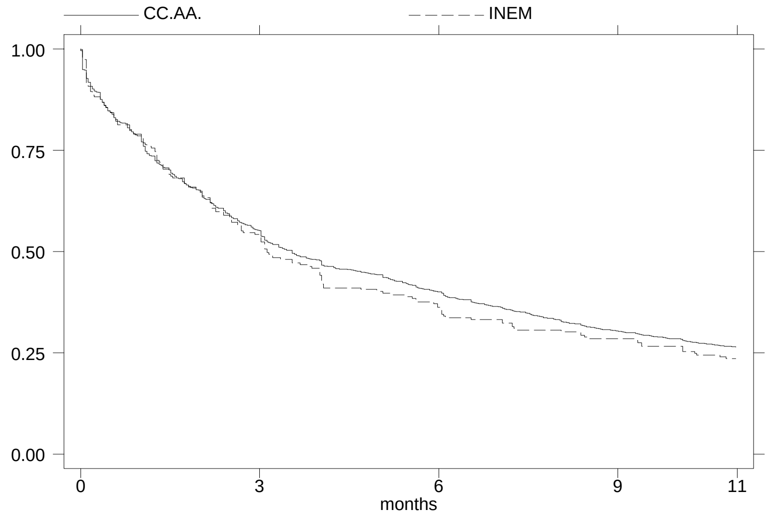

Figure 17: Management of the courses

The duration of the work experience is other important variable which shows information of the labour history of the worker. There is a quadratic effect on the probability to be employed. A period below two years produces is preferred to other option. A possible explanation is that firms do not show interest in highly experienced workers, but young workers who have some experience to carry out the new job.

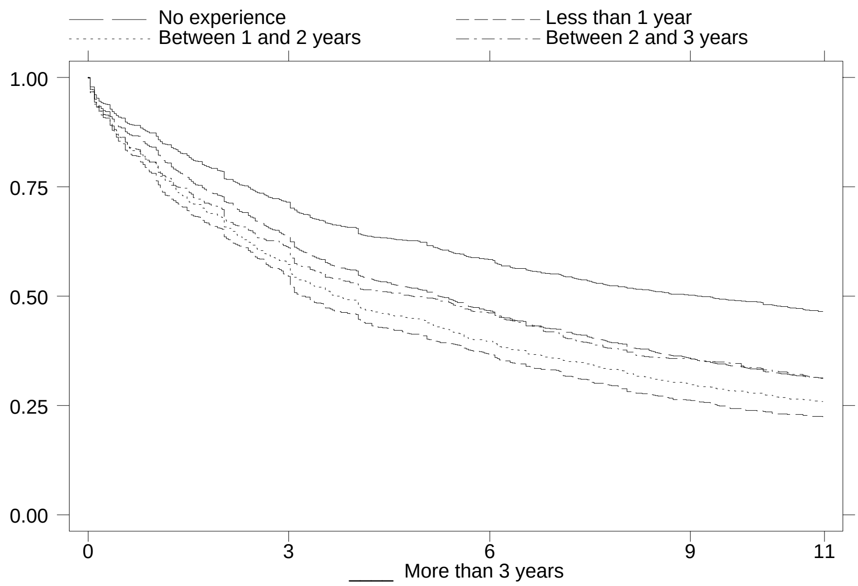

Figure 18: Duration of work experience

Adjustments and comparison of survivor functions are realized, controlling the estimates of the survivor functions for each variable. The reasons are the same as those presented in the previous figures, lack of independence and censorship[9].

Considering all possible combinations, the number of significant adjustments is modest. Whether the differences among values of a variable are slighter than other, then these adjustments[10] of survivor function estimates are more probable to occur in the former variable than in the latter.

An alternative to adjustments and graphic solutions is testing the similarity of groups for each variable. The tests belong to a family of test statistics which are extensions of non-parametric rank tests to compare two or more distributions for censored data[11]. The null hypothesis of these tests is:

$$H_0 : h_1(t) = h_2(t) = ... = h_k(t) \quad \text{for all } t.$$

where $h(t)$ is the hazard function at time t, against the alternative hypothesis that at least one of the functions is different for some t.

Results are consistent with conclusions from figures. The tests confirm the statistical inexistence of similarity among groups for each variable, regardless of the most usual confidence intervals (at 99%, 95% and 90%).The exception is the competence regarding management of the courses, where the null hypothesis is not rejected. In the case of curves adjusted for all the combinations of variables, the conclusions are similar to the case without adjustments, and compatible with those derived from graphic study.

[9] When censoring time is not independent of survival time, the estimates of the survivor function will be overestimated if individuals who disappear have a great probability to be hired, or underestimated if individuals who disappear have a small possibility to find a job.

[10] Significant adjustments are considered not only movements of curves, but the differences among groups are clearly reduced. This fact indicates the elimination of differences is justified by the adjustment variable. Given this result and the relatively high number of groups for each variable, conclusions are reasonable.

[11] The Mantel-Haenszel test (also known as **log-rank test**), the Breslow test (also known as **generalized Wilcoxon test**), the **Tarone-Ware test** and the **Peto-Peto-Prentice test**. More information in Miller (1981).

5 Parametric and semi-parametric studies

Data have been used to calculate the probability that a worker gets a job before time t conditional on each worker's characteristic in previous section. The goal of following sections lies in fitting the survivor function from data base using all together variables. There exist two basic models to implement it:

- Parametric models: Accelerated Failure-Time (AFT) models.
- Semi-parametric models: Proportional Hazard rate (PH) models.

Both models can be considered as particular cases of so-called Mixed Proportional Hazard rate (MPH) models. Although parametric survival distributions are included in statistical packages, they are not necessary for specification and identification, as van den Berg (2000) points out.

5.1 AFT models

In spite of the difficulties and limitations of AFT models, van den Berg (2000) points out the utility of these models: "From an econometric point of view, the AFT approach is unsatisfactory, because it does not focus on the parameters of the individual hazard as the parameters of interest. However, if one is only interested in the sign or significance of a covariate effect on the individual durations then the AFT approach may be useful."

Let T be the unemployment duration, a non-negative continuous random variable, x be a finite-dimensional vector of observed explanatory variables (or covariates), and b be a vector of regression coefficients. As van den Berg (2000) and van Ours (2001) point out, several assumptions are considered, as inexistence of unobserved heterogeneity among individuals who affect the hazard function, elimination of anticipation of workers' decisions, and all individuals are mutually independent.

Linear regression models are connected to hazard models through AFT models. Suppose that $Y = \ln T$ is related to x via a linear regression model:

$$Y = x'b + z \tag{1}$$

where z is the error term with density function $f(\cdot)$.

Exponentiation of (1) gives

$$T = \exp(x'b)T'$$

where $T' = \exp(z)$ has hazard function $h_0(t')$ independent of b. Hence, the hazard function for T can be written in terms of this baseline hazard function $h_0(\cdot)$

$$h(t,x) = h_0(te^{-x'b})e^{-x'b}$$

This model specifies that the effect of the covariates is multiplicative on t, in stead of the hazard function as in PH models. The role of the covariates is to change (accelerate or decelerate) the time to failure.

The distributional form of the error determines the properties of the regression model. Six parametric survival distributions are used, Exponential, Log-logistic, Log-normal, Weibull, Generalized Gamma and Gompertz[12].

The models proposed in the paper are defined as

$$\begin{aligned} y &= \ln t = x'b + w'c + af + fx'd + z && \text{in the general case} \\ y_i &= \ln t_i = x_i'b_i + w_i'c_i + a_i f_i + f_i w_i' d_i + z_i && \text{for each course } \ i = 1,2,3,4 \end{aligned}$$

The set of covariates is constituted by the variable *treatment*, identified by f (for f_i , $i = 1,2,3,4$ when each course is considered independently) and the rest of covariates in the vectors x and w.

Let $f(\cdot)$ and $S(\cdot)$ be the appropriate distributions for the desired parametric regression model. Assume that there are N individuals, U of whom have uncensored times. The full log-likelihood function that should be maximized to estimate the parameter vector v and the covariate coefficients vector b is:

$$\ln L = \sum_{j=1}^{U} \ln \{f(t_j, v \mid t_{0j})\} + \sum_{j=U+1}^{N} \ln \{S(t_j, v \mid t_{0j})\} \tag{2}$$

where a subject known to fail at t contributes to the likelihood function $f(t_j, v \mid t_{0j})$, the value of the density at t conditional on the entry time t_0, whereas a censored observation, only known to be unemployed up to t, contributes $S(t_j, v \mid t_{0j})$, the probability of being unemployed beyond t conditional on the entry time t_0. Censorship is considered as non-informative[13].

[12] More information about these distributions in Kabfleisch and Prentice (1980).

[13] Miller (1981) and Kalbfleisch and Prentice (1980) discuss the conditions for censored data to do a valid study.

5.1.1 Results from estimation

The final models include the variables defined in Appendix A and products of them, which capture the effect of the interaction between covariates. Using the individual significance tests, 90%, 95% and 99% confidence levels are taken into account. Basic significant estimates with their associated standard deviations appear in Appendix B[14]. Several regressions are shown depending on the set of workers and the treatment. In Table B1, the first column of estimates considers the complete sample and the second and third columns present estimates for men and women, for general treatment. In Table B2 and Table B3, males and females are analyzed taking into account the four course levels.

The discrete variables which are not dichotomous were transformed into multiple dummy variables. The first value of each covariate is disappeared, except for provinces where the variable eliminated was Madrid, to avoid problems of collinearity. Therefore, the values of the estimates of the dummy variables have as reference the eliminated group.

The interpretation of the coefficients is complex as almost all covariates are discrete. However, the signs of the coefficients may be interpreted and the quantities among covariates of the same group can be compared among them. If the sign is positive, the corresponding covariate increases the duration in unemployment. Otherwise, duration is reduced.

An interesting aspect is the sign consistency of estimates in the AFT models, at least in the case of those statistically different from cero, as Appendix B shows partially.

The fact to be a woman affects positively to maintain the unemployment situation. This conclusion is related to the reality of Spanish labour market and the result from Kaplan-Meier estimates, where female unemployment rate is significantly higher than male one.

With respect to age, there is a non-linear effect for the full sample. This result is a combination of effects between men and women. The increase of age is profitable to get a job for men, but this effect decrease with further increases of age. This conclusion is similar to Kaplan-Meier estimates. However, women have more problems to get a job when age increases and this effect is linear. Experience is not rejected as a main factor to be employed for men. It would be less effective for women, especially whether job offers may not need a high level of experience. A young person (with a high labour life expectancy) is more profitable to be trained than an old worker. Taking into account the labour demand, young women may accept any job offer before than adult women.

The education level does not generate important effects on unemployment exit rate. The same conclusion can be established when education is combined

[14] All estimates of models are available on request.

with gender or treatment. A possible justification is the specific and limited know-how of job offers, so education cannot be interpreted neither as a signal, nor as a substitute for experience.

The existence of disability increases the probability of being unemployed for the full sample, but this effect is only significant for men. When the different course levels are considered, the effect is negative regardless of gender but women show less destructive results than men.

The effect of benefits is restricted to two groups, workers who receive benefits and those whose help has been disappeared by any reason. The latter group gets a job before than the former. This result is consistent with the negative effect of benefits on searching for a job, and it appears in all the estimates.

The economic activity of the first option selected by the worker does not usually affect the unemployment period. Something similar occurs with the driving license and the military service for men except for the full sample. Military service may reflects other effect coming from the interaction of age and gender.

The results are modest in the case of idioms. Only English, Galician and Basque reduce the unemployment duration. Considering gender and the course levels, English and Galician are good instruments to leave unemployment situation for men, and Basque for women.

The civil status reflects a different effect between men and women. Married women get a job latter than other group. Firms prefer women without family charges and with total dedication to the job. Married women are very demanding with job search because of familiar (e.g. other family members generate income) or labour questions. Finally, a possible depreciation of human capital when women who were married and left the labour market makes incorporation into a new job more difficult.

The local factor is more important for women. For this reason, the interaction between provinces and women is incorporated in the full sample. Considering the regional level, the provinces with higher difficulty to get a job belong to the regions of Andalucía (except for Almería), Asturias, Castilla-La Mancha (except for Guadalajara), Extremadura, Murcia, and especially Ceuta and Melilla. Alicante, La Coruña, León, Las Palmas, Pontevedra, Valladolid, Valencia and Vizcaya are also affected by this situation. Those with good results are the provinces of Aragón, Islas Baleares, Cataluña and Navarra, and moreover, Alava, Guadalajara and Zamora. The conclusions are usually similar to Kaplan-Meier estimates. A clear economic relationship between the probability ti leave unemployment situation and GDP per capita cannot be confirmed.

The effect of the period of active labour demand on employment search is negative, significant and non-linear. An increase of the days is harmful to get a job, but the growth rate of this damage decreases with the active labour demand period.

The effect of training courses is significant and helps to get a job, except for Level 1 (Broad basis). The training courses managed by INEM reduces women's unemployment period for the medium level courses (Occupation and Specialization). However, the interaction of training and gender does not obtain a significant estimate. Finally, the combination of treatment and disability does not appear to generate any positive result.

5.1.2 Model selection and unobserved heterogeneity

There exist several selection criteria among the different models. When parametric models are not nested, an appropriate approach is the Akaike Information Criterion (AIC)[15]. Akaike (1974) suggested penalizing each log-likelihood function to reflect the number of parameters being estimated in a particular model and then comparing among them. The AIC can be defined as

$$AIC = -2(\log L) + 2\,(c + p + 1)$$

where *log L* is the value of the log-likelihood function, p is the number of model-specific ancillary parameters and c is the number of model covariates.

There are slight differences in the value of the log-likelihood function between Weibull and Generalized Gamma models, normally in favour of the latter. Following the AIC, the best option is the Generalized Gamma distribution in the majority of cases. Fortunately, both models are nested. The estimate of one of the ancillary parameters $\hat{\kappa}$ for Generalized Gamma distribution (with standard deviation), allows to reject the hypothesis that $\kappa = 0$ (test for the appropriateness of the lognormal) and $\kappa = 1$ (strong support against rejecting the Weibull model)[16].

Other method is to calculate an empirical estimate of the cumulative hazard function based on the Kaplan-Meier survival estimates, taking the Cox-Snell residuals[17] as the time variable. If the estimated model fits the data, then

[15] Kalbfleisch and Prentice (1980) justify the selection of a model basing on the value of the log-likelihood function.

[16] The density function of the Generalized Gamma distribution is

$$f(t) = \begin{cases} \frac{\gamma^{\gamma}}{\sigma t\sqrt{\gamma}\Gamma(\gamma)} \exp\left(z\sqrt{\gamma} - u\right), & \text{if } \kappa \neq 0 \\ \frac{1}{\sigma t\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right), & \text{if } \kappa = 0 \end{cases}$$

where $\gamma = \mid \kappa \mid^{-2}$, $z = sign(\kappa)\frac{\ln(t)-\mu}{\sigma}$, $u = \gamma \exp\left(\mid \kappa \mid z\right)$.

[17] The Cox-Snell residuals can be derived from the expression

the Cox-Snell residuals have a standard censored exponential distribution with hazard ratio 1. These figures are usually consistent with the choice of the Generalized Gamma model. Despite the selection methods between the Weibull and Generalized Gamma distribution, their estimates show small differences, so the selection effect is limited.

Inexistence of unobserved heterogeneity was assumed in the previous subsections. The use of frailty models or survival models with unobservable heterogeneity is proposed to analyze the importance of this assumption. Frailty is introduced as an unobservable multiplicative effect α on the hazard function such that[18]

$$h(t \mid a) = ah(t)$$

where $h(t)$ is a non-frailty hazard function, a is a random positive quantity (for purposes of model identification is assumed to have mean one and variance θ finite) with density function $g(a)$. For purposes of mathematical tractability, the choice is limited to one of either the Gamma distribution $\mathrm{G}\left(\frac{1}{\theta}, \theta\right)$ or the Inverse-Gaussian distribution $\mathrm{IG}\left(1, \frac{1}{\theta}\right)$[19]. Results indicate negligible unobserved heterogeneity in AFT models, specially for models distinguishing among course levels and gender.

$$\hat{r}_{C_i} = -\log\left[\hat{R}\left(y_i\right)\right]$$

where $\hat{R}\left(y_i\right) = 1 - G\left(\frac{y_i - x_i'\hat{b}}{\hat{s}}\right)$. $G(\cdot)$ is the distribution function of the model. The Cox-Snell residual for a subject at time t is defined as $\hat{H}(t)$, the estimated cumulative hazard function obtained from the fitted model.

[18] Taking into account this idea, Lancaster (1990) presents parameter a as the total effect of unmeasured systematic differences on the hazard function. He mentions several reasons for a mixture model, as omitted variables and errors in the equation.

[19] Gamma distribution G(a,b) has the density function

$$g(x) = \frac{x^{a-1}e^{-\frac{x}{b}}}{\Gamma(a)b^a}$$

For the case of Inverse-Gaussian with parameters a and b

$$g(x) = \left(\frac{b}{2\pi x^3}\right)^{\frac{1}{2}} \exp\left[-\frac{b}{2a}\left(\frac{x}{a} - 2 + \frac{a}{x}\right)\right]$$

5.2 PH models

Although there is not any economic principle justifying hazard functions should be proportional, following Lancaster (1990), Proportional Hazard rate models (PH models) are used often. In AFT models, distribution function is assumed known, except for a few scalar parameters. The PH model is non-parametric in the sense that it involves an unspecified function whose structure is an arbitrary baseline hazard function. Therefore, this model is more flexible and it presents further complex conditions.

Let $h(t, x)$ be the hazard function of an individual with a vector of observed variables x at time t. The PH model proposed by Cox (1972) is specified by the hazard relationship:

$$h(t, x) = h_0(t) \cdot e^{x'b} \tag{3}$$

where $h_0(t)$ is an arbitrary and unspecified baseline hazard function. In this setting, variables included in x act multiplicatively in the hazard function, unlike AFT models. This model provides estimates of the vector b, but it does not provide a direct estimate of $h_0(t)$. Therefore, it is complicated to compare hazard functions of Cox and AFT models.

The most important assumption of the PH model is that the hazard ratio is proportional over time. Once the models are estimated, proportional hazards assumption is evaluated using a test of proportional hazards based on the generalization by Grambsch and Therneau (1994). The null hypothesis of a zero slope in a generalized linear regression of the scaled Schoenfeld residuals[20] on functions of time. The null hypothesis is accepted, specially when more restrictive subsamples are considered.

The PH models estimated are defined as

$$\begin{aligned} h(t, f, x, w) &= h_0(t) \cdot \exp\left[x'b + w'c + af + fx'd\right] \\ h_i(t_i, f_i, x_i, w_i) &= h_{0i}(t_i) \cdot \exp\left[x_i'b_i + w_i'c_i + a_i f_i + f_i x_i' d_i\right] \quad \text{for course } i = 1, 2, 3, 4 \end{aligned}$$

where the definition of each term coincides with the description of the AFT model.

The PH model assumes that the hazard function is continuous, so there should not be tied survival times. However, tied events do occur in survival

[20]The Schoenfeld residual for an individual i and a variable k consists of the difference between the value of this variable for individual i and its estimated expected value conditional on the hazard group when i fails (R_i):

$$\hat{r}_{ik} = X_{ik} - \hat{E}\left(X_{ik} \mid R_i\right)$$

More information in Schoenfeld (1982).

data. Efron method is used to solve this problem. This procedure assigns the same probability of failure to observations that fail at the same time inside the subset of risk observations.

The process of eliminating variables to avoid collinearity is similar to the AFT models. Estimates of the most outstanding variables appear in Appendix C. The PH models present the same sign of the significant estimates for all the variables in Appendix C, as it occurred in AFT models. In PH models, the estimates are interpreted as the effect on the unemployment exit rate (AFT models analyse the effect on the unemployment period). Therefore, the results of the PH models are consistent with those of the AFT models.

6 Conclusions and extensions

The goal of the paper lies in the study and evaluation of one of the active labour market policies elaborated in Spain: the National Plan of Training and Professional Insertion. It is carried out by INEM or the region with the corresponding competence. The training courses are divided in four levels, Broad Basis, Occupation, Specialization and Adaptation-Occupation. A subsample of workers who did a course in the first quarter of 2000 (treatment group) is compared to a group of workers who did not any course (control group).

Using the unemployment period as dependent variable, non-parametric methods of survivor functions and duration models are used to evaluate the effectiveness of the different types of courses. Conclusions from estimates of AFT and PH models are similar in this experiment.

The existence of disabilities for men, the rise of age (especially for women) and the days of active labour demand increase the unemployment period. These characteristics are similar to the conclusions from non-parametric estimates.

However the condition of being disabled for women, the driving license, the education level and the economic activity of the first option do not seem to affect the unemployment duration clearly. These facts do not appear in the Kaplan-Meier estimates

The opportunity of women to get a job improves with respect to men when the workers do the training courses. This effect is particularly important for medium level courses, Occupation and Specialization. The last conclusions suggest the labour market measures have to encourage an interest in medium level training courses specially in women.

References

[1] Abbring, J. H. and G. J. van den Berg (2000), The non-parametric identification of treatment effects in duration models (under revision).

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

[3] Arellano, A. (2005), Evaluating the effects of labour market reforms "at the margin" on unemployment and employment stability: the Spanish case, *Universidad Carlos III de Madrid, Economics series 5, Working paper 05-12*, February 2005.

[4] Cox, D. R. (1972), Regression models and life-tables (with discussion), *Journal of the Royal Statistical Society*, Series B, 34: 187-220.

[5] Dolado, J. J., C. García-Serrano and J. F. Jimeno (2001), Drawing lessons from the boom of temporary jobs in Spain, *The Economic Journal*, vol. 112, 480.

[6] Grambsch, P. M. and T. M. Therneau (1994), Proportional hazards tests and diagnostics based on weighted residuals, *Biometrika*, 81: 515-526.

[7] Ham, J. C. and R. J. Lalonde (1996), The effect of sample selection and initial conditions in duration models: evidence from experimental data on training, *Econometrica*, 64: 175-205.

[8] Holland, P.W. (1986), Statistics and Causal Inference, *Journal of the American Statistical Association*, *81*: 945-970.

[9] Kalbfleisch, J. D. and R. L. Prentice (1980), *The Statistical Analysis of Failure Time Data*, New York: John Wiley & Sons.

[10] Kaplan, E. L. and P. Meier (1958), Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, 53: 457-481.

[11] Kugler, A., Jimeno, J. F. and V. Hernanz (2002), Employment consequences of restrictive permanent contracts: evidence from Spanish labor market reforms, *IZA Discussion Paper 657.*

[12] Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Cambridge: Cambridge University Press.

[13] Lechner, M. (1999), Earnings and employment effects of continuous off-the-job training in East Germany after unification, *Journal of Business & Economic Statistics*, 17: 74-90.

[14] Martin, J. P. (2000), What works among active labour market policies: evidence from OECD countries' experiences, *OECD Economic Studies*, 30, 2000/1.

[15] Miller, R. G. (1981), *Survival Analysis*, New York: John Wiley & Sons.

[16] OECD (2002), *Employment perspectives.*

[17] Ridder, G. (1986), An event history approach to the evaluation of training, recruitment and employment programs, *Journal of Applied Econometrics*, 1: 109-126.

[18] Rubin, D.B. (1974), Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of Educational Psychology*, 66: 688-701.

[19] Schoenfeld, D. (1982), Partial residuals for the proportional hazards regression model, *Biometrika*, 69: 239-241.

[20] Van den Berg, G. J. (2000), Duration models: specification, identification, and multiple durations, in: Heckman, James J., and Leamer, Edward (eds.) *Handbook of Econometrics*, Volume V, North-Holland.

[21] Van Ours, J. C. (2001), Do active labor market policies help unemployed workers to find and keep regular jobs?, in Lechner, M., and Pfeiffer, F. (eds.) *Econometric Evaluation of Labour Market Policies*, ZEW Economic Studies, Volume 13, Heidelberg New York: Physica-Verlag.

Appendix A: Variables and descriptive statistics.

In order to distinguish between original variables and transformed ones, we used the letter X. For example, *levestu* is a discrete variable that takes values between 0 and 9, and *levestuX* is a transformed dummy variable that takes value one if $levestu = X$ and cero otherwise.

1. *woman* is a dummy variable equals 1 if female.

2. *age* is a variable which takes values from 16 to 60 years old. The product by itself is called *age*2.

3. *levestuX* is a group of dummy variables that adopts value one if the worker's education level is X, where X is:

$$X = \begin{cases} 0 \rightarrow \text{ No education} \\ 1 \rightarrow \text{ Pre-Primary education without certificate} \\ 2 \rightarrow \text{ Pre-Primary education with certificate} \\ 3 \rightarrow \text{ Basic Technical College} \\ 4 \rightarrow \text{ Primary education} \\ 5 \rightarrow \text{ Medium Technical College} \\ 6 \rightarrow \text{ Secondary education} \\ 7 \rightarrow \text{ Superior Technical College} \\ 8 \rightarrow \text{ University education (3 years)} \\ 9 \rightarrow \text{ University education (more than 3 years)} \end{cases}$$

4. *disabled* is a dummy variable with value 1 if the worker is disabled.

5. *benefitX* is a set of dummy variables indicating if a worker has some benefit. X indicates the type of benefit:

$$X = \begin{cases} 0 \text{ if the individual does not receive any help} \\ 1 \text{ if the individual receives benefits} \\ 2 \text{ if the individual receives any subsidy except for workers above 52 years and agriculture} \\ 3 \text{ if the individual receives a subsidy for workers above 52 years} \\ 4 \text{ if the help has been disappeared} \end{cases}$$

6. *treatment* is a dummy variable equals to 1 if the worker did a training course in the first quarter of 2000.

7. *provX* is a set of dummy variables indicating in which province the worker lives. The value of X appears in the index of this Appendix.

8. *groupX* is a set of dummy variables for each economic activity in which workers desire to work as their first preference. These variables follow the ten Big Groups of the National Classification of Occupations (CNO-94):

$$X = \begin{cases} 0 \rightarrow & \text{Armed Forces} \\ 1 \rightarrow & \text{Management and Public Administrations} \\ 2 \rightarrow & \text{Technicians, professionals, scientists and intellectuals} \\ 3 \rightarrow & \text{Support technicians and professionals} \\ 4 \rightarrow & \text{White-collar workers} \\ 5 \rightarrow & \text{Restaurant workers, protection and sellers} \\ 6 \rightarrow & \text{Skilled workers in farming and fishing} \\ 7 \rightarrow & \text{Workers in factory industry, construction and mining} \\ 8 \rightarrow & \text{Operators of installations and machinery, and assemblers} \\ 9 \rightarrow & \text{Non-skilled workers} \end{cases}$$

9. *dmonth* defines the number of months of the last worker´s active labour demand period. The product by itself is defined as *dmonth*2.

10. Several dummy variables related to civil status are introduced: *single*, *married*, *widowed*, *divorced*.

11. *milser* indicates whether a man did military service (value equals 1) or not.

12. *drliX* is a set of dummy variables which show different levels of driving license:

$$X = \begin{cases} 0 \rightarrow & \text{No driving license} \\ b1 \rightarrow & \text{Cars for private use and lorries (weight } \leq \text{ 3,500 Kg.)} \\ b2 \rightarrow & \text{Cars for public use (weight } \leq \text{ 3,500 Kg.)} \\ c1 \rightarrow & \text{Cars for private use and lorries (weight } \leq \text{ 16,000 Kg.)} \\ c2 \rightarrow & \text{Trucks} \\ d \rightarrow & \text{Buses and coaches} \\ e \rightarrow & \text{Vehicles with trailers} \end{cases}$$

13. $firedX$ is a group of dummy variables which identify the reason to loose the previous job:

$$X = \begin{cases} 0 \rightarrow & \text{No previous job} \\ 1 \rightarrow & \text{Voluntary dismissal} \\ 2 \rightarrow & \text{Contract ending} \\ 3 \rightarrow & \text{Firing} \\ 4 \rightarrow & \text{Job adjustment plan} \\ 5 \rightarrow & \text{Inactivity for a discontinuous permanent contract} \\ 6 \rightarrow & \text{Trial period is not passed} \end{cases}$$

14. There are several dummy variables indicating whether the workers knows other languages: English, French, German, Catalan, Galician, Basque, Valencian and other European (*European*) and non-European (*othlan*) languages.

15. *expocu* is a continuous variable defining the duration of the work experience in months.

16. In the case of management of the courses, there are two dummy variables: *treatinem* and *treatccaa*.

17. There is a set of dummy variables about the economic activity of the previous job:

    agrfish→ Agriculture and Fishing
    indust→ Industry and Energy
    constr→ Construction
    commer → Commerce
    cater→ Catering services
    transp→ Transports
    finan→ Financial services
    estate→ Real estate activity
    pubadm → Public Administration
    educat→ Education
    health→ Health
    othser→ Other Services

Index of provinces:

Alava	1	León	24
Albacete	2	Lérida	25
Alicante	3	Lugo	27
Almería	4	Madrid	28
Asturias	33	Málaga	29
Avila	5	Melilla	52
Badajoz	6	Murcia	30
Baleares	7	Navarra	31
Barcelona	8	Ourense	32
Burgos	9	Palencia	34
Cáceres	10	Las Palmas	35
Cádiz	11	Pontevedra	36
Cantabria	39	La Rioja	26
Castellón	12	Salamanca	37
Ceuta	51	S. C. Tenerife	38
Ciudad Real	13	Segovia	40
Cordoba	14	Sevilla	41
Coruña	15	Soria	42
Cuenca	16	Tarragona	43
Gerona	17	Teruel	44
Granada	18	Toledo	45
Guadalajara	19	Valencia	46
Guipuzcoa	20	Valladolid	47
Huelva	21	Vizcaya	48
Huesca	22	Zamora	49
Jaen	23	Zaragoza	50

Figure A1: Provinces of Spain

Variables	Total	Treatment	Control
Woman	59.93	63.72	58.04
	(0.49)	(0.48)	(0.49)
Age	30.36	30.02	30.52
	(10.21)	(9.03)	(10.75)
Primary Education	40.41	36.16	42.53
Disabled	2.04	1.89	2.17
	(0.14)	(0.14)	(0.14)
No Benefits	63.35	65.83	62.11
Residence in Madrid	24.64	43.34	15.32
Residence in provinces > 1,000,000[21]	54.34	58.85	52.10
White-collar workers (as first option)*	21.34	24.38	19.83
Restaurant workers, protection and sellers *	22.51	20.18	23.67
Months of active labour demand	9.13	10.88	8.26
	(11.24)	(11.73)	(10.89)
Single	71.25	75.07	69.34
Driving license B1	63.88	67.48	62.08
No idioms	46.25	44.93	46.91
English	33.07	37.66	30.78
No previous job	22.74	21.10	23.56
Size	18,617	6,190	12,427

Table A1: Descriptive statistics[22]

[21] The provinces with a population higher than one million people in 2001 were (in alphabetical order) Alicante, Asturias, Barcelona, Cádiz, La Coruña, Madrid, Málaga, Murcia, Sevilla, Valencia and Vizcaya (see map in Appendix A).

[22] The table reports means and percentages for the indicated group. Standard deviations are in parenthesis where appropriate.

MEN			
Variables	Total	Treatment	Control
Age	31.90	29.96	32.74
	(11.38)	(9.84)	(11.88)
Primary Education	43.02	40.69	44.02
Disabled	3.16	3.03	3.22
	(0.18)	(0.17)	(0.18)
No Benefits	54.08	60.95	51.11
Residence in Madrid	27.04	50.31	17.01
Residence in provinces > 1,000,000	58.73	65.32	55.88
White-collar workers (as first option)*	13.22	13.12	13.45
Restaurant workers, protection and sellers*	10.71	10.53	11.13
Months of active labour demand	8.48	9.48	8.05
	(10.73)	(10.65)	(10.74)
Single	72.09	80.94	68.28
Compulsory military service	82.57	80.63	83.41
Driving license B1	64.84	66.07	64.31
No idioms	48.23	47.82	48.41
English	31.64	36.29	29.63
No previous job	17.14	16.03	17.63
Size	7,460	2,246	5,214

WOMEN			
Variables	Total	Treatment	Control
Age	29.32	30.05	28.92
	(9.21)	(8.54)	(9.53)
Primary Education	38.67	33.57	41.45
Disabled	1.29	1.24	1.32
	(0.11)	(0.11)	(0.11)
No Benefits	69.55	68.61	70.07
Residence in Madrid	23.03	39.38	14.10
Residence in provinces > 1,000,000	51.39	55.17	49.34
White-collar workers (as first option)*	26.77	30.60	24.68
Restaurant workers, protection and sellers*	30.39	25.33	33.16
Months of active labour demand	9.56	11.67	8.41
	(11.56)	(12.24)	(11.00)
Single	70.68	71.73	70.11
Driving license B1	63.23	68.28	60.47
No idioms	44.93	43.28	45.83
English	34.02	38.44	31.61
No previous job	26.49	23.99	27.85
Size	11,157	3,944	7,213

Table A2: Descriptive statistics by gender

Appendix B: Tables of estimates - AFT models.

Table B1: AFT models for the full sample

	Total	Men	Women
Woman	.324*** (.065)		
Age	-.005 (.01)	-.055*** (.014)	.039*** (.014)
Age×age	.0007*** (.0001)	.001*** (.0002)	.00006 (.0002)
Military service	-.18*** (.053)	-.09 (.058)	
Married	-.065 (.058)	-.01 (.065)	.308*** (.051)
Treatment	-.458*** (.048)	-.487*** (.05)	-.52*** (.044)
Treatment×woman	-.099 (.062)		
Treatment×disabled	.252 (.202)	.335 (.264)	.161 (.314)
Treatment managed by INEM	-.354*** (.112)	-.045 (.187)	-.591*** (.139)
Married×woman	.44*** (.066)		
Active labour demand in months (dmonth)	.081*** (.003)	.081*** (.005)	.08*** (.004)
dmonth2	-.0009*** (.00007)	-.0008*** (.0001)	-.001*** (.00008)
English	-.068** (.028)	-.087** (.044)	-.065* (.036)
Galician	-.573*** (.194)	-1.009*** (.293)	-.242 (.257)
Basque	-.22** (.108)	-.219 (.171)	-.204 (.138)
Disabled	.783*** (.136)	.757*** (.147)	.296 (.186)
Disabled×woman	-.501*** (.194)		
Worker receives benefits	.289*** (.036)	.265*** (.053)	.332*** (.051)
Help has been disappeared	-.389*** (.047)	-.27*** (.069)	-.481*** (.063)
Size	18,617	7,460	11,157

Table B2: AFT models for men and course levels

MEN	Course 1	Course 2	Course 3	Course 4
Age	-.06*** (.016)	-.061*** (.015)	-.056*** (.015)	-.058*** (.016)
Age×age	.001*** (.0002)	.001*** (.0002)	.001*** (.0002)	.001*** (.0002)
Military service	.0003 (.067)	-.069 (.062)	-.039 (.063)	-.00009 (.065)
Married	.003 (.073)	-.008 (.069)	.009 (.07)	-.006 (.072)
Treatment	-.662 (.783)	-.508*** (.058)	-.449*** (.084)	-.395*** (.104)
Treatment×disabled		.325 (.296)	.317 (.504)	.51 (.901)
Treatment managed by INEM		.038 (.213)	-.382 (.409)	-1.268 (.941)
Active labour demand in months (dmonth)	.075*** (.006)	.083*** (.006)	.072*** (.006)	.077*** (.006)
dmonth2	-.0007*** (.0001)	-.0008*** (.0001)	-.0006*** (.0001)	-.0007*** (.0001)
English	-.131** (.052)	-.099** (.048)	-.13*** (.049)	-.115** (.05)
Galician	-.786*** (.294)	-1.002*** (.304)	-.776*** (.292)	-.828*** (.288)
Basque	-.285 (.188)	-.21 (.184)	-.337* (.18)	-.235 (.181)
Disabled	.735*** (.141)	.757*** (.149)	.734*** (.139)	.743*** (.14)
Worker receives benefits	.378*** (.06)	.306*** (.056)	.315*** (.057)	.379*** (.059)
Help has been disappeared	-.328*** (.08)	-.262*** (.075)	-.34*** (.077)	-.322*** (.078)
Size	5,217	6,779	5,629	5,477

$***$ 99% significant level $**$ 95% significant level $*$ 90% significant level , standard deviations in parenthesis.

Table B3: AFT models for women and course levels

WOMEN	Course 1	Course 2	Course 3	Course 4
Age	.032** (.016)	.047*** (.015)	.025 (.016)	.032** (.016)
Age×age	.0001 (.0002)	-.00005 (.0002)	.0003 (.0002)	.0001 (.0002)
Married	.288*** (.062)	.31*** (.055)	.284*** (.059)	.318*** (.061)
Treatment	.004 (.621)	-.525*** (.049)	-.53*** (.063)	-.384*** (.09)
Treatment×disabled		-.012 (.348)	.688 (.545)	-2.044 (1.266)
Treatment managed by INEM		-.534*** (.158)	-.711*** (.274)	-.071 (1.253)
Active labour demand in months (dmonth)	.079*** (.005)	.082*** (.005)	.08*** (.005)	.077*** (.005)
dmonth2	-.0009*** (.0001)	-.001*** (.00009)	-.0009*** (.0001)	-.0008*** (.0001)
English	-.03 (.045)	-.039 (.04)	-.052 (.042)	-.042 (.043)
Galician	-.373 (.286)	-.064 (.275)	-.521* (.275)	-.434 (.28)
Basque	-.26* (.15)	-.176 (.143)	-.283* (.146)	-.274* (.148)
Disabled	.322* (.177)	.305* (.186)	.305* (.181)	.311* (.176)
Worker receives benefits	.401*** (.062)	.355*** (.056)	.383*** (.058)	.373*** (.06)
Help has been disappeared	-.472*** (.081)	-.492*** (.067)	-.466*** (.076)	-.492*** (.078)
Size	7,221	9,694	8,309	7,572

$***$ 99% significant level $**$ 95% significant level $*$ 90% significant level , standard deviations in parenthesis.

Appendix B: Tables of estimates of PH models.

Table C1: PH models for the full sample

	Total	Men	Women
Woman	-.237*** (.048)		
Age	.005 (.007)	.042*** (.01)	-.029*** (.011)
Age×age	-.0004*** (.0001)	-.001*** (.0001)	-.00005 (.0002)
Military service	.137*** (.039)	.067 (.042)	
Married	.049 (.043)	.004 (.048)	-.237*** (.039)
Treatment	.33*** (.035)	.345*** (.036)	.391*** (.033)
Treatment×woman	.08* (.046)		
Treatment×disabled	-.167 (.153)	-.237 (.198)	-.088 (.241)
Treatment managed by INEM	.27*** (.083)	.044 (.137)	.459*** (.105)
Married×woman	-.335*** (.05)		
Active labour demand in months (dmonth)	-.06*** (.002)	-.058*** (.004)	-.06*** (.003)
dmonth2	.0007*** (.00005)	.0006*** (.00008)	.0008*** (.00006)
English	.052** (.021)	.062* (.032)	.053* (.028)
Galician	.381*** (.144)	.674*** (.214)	.149 (.195)
Basque	.156* (.08)	.177 (.125)	.137 (.105)
Disabled	-.584*** (.104)	-.556*** (.11)	-.242* (.144)
Disabled×woman	.36** (.147)		
Worker receives benefits	-.208*** (.027)	-.185*** (.039)	-.249*** (.039)
Help has been disappeared	.295*** (.035)	.204*** (.051)	.373*** (.048)
Size	18,617	7,460	11,157

Table C2: PH models for men and course levels

MEN	Course 1	Course 2	Course 3	Course 4
Age	.048*** (.013)	.046*** (.011)	.045*** (.012)	.046*** (.012)
Age×age	-.001*** (.0002)	-.001*** (.0002)	-.001*** (.0002)	-.001*** (.0002)
Military service	-.003 (.052)	.049 (.044)	.029 (.049)	-.001 (.05)
Married	-.006 (.057)	.003 (.05)	-.01 (.055)	.0003 (.056)
Treatment	.519 (.595)	.347*** (.041)	.354*** (.066)	.31*** (.08)
Treatment×disabled		-.228 (.219)	-.236 (.4)	-.373 (.725)
Treatment managed by INEM		-.011 (.153)	.303 (.319)	1.107 (.713)
Active labour demand in months (dmonth)	-.057*** (.005)	-.059*** (.004)	-.055*** (.004)	-.059*** (.005)
dmonth2	.0006*** (.0001)	.0006*** (.00009)	.0006*** (.0001)	.0006*** (.0001)
English	.102** (.041)	.069** (.034)	.103*** (.039)	.088** (.039)
Galician	.565** (.228)	.647*** (.218)	.566** (.227)	.596*** (.224)
Basque	.23 (.147)	.169 (.133)	.274* (.142)	.189 (.141)
Disabled	-.568*** (.111)	-.547*** (.111)	-.572*** (.111)	-.575*** (.111)
Worker receives benefits	-.281*** (.047)	-.208*** (.041)	-.237*** (.045)	-.283*** (.046)
Help has been disappeared	.26*** (.063)	.193*** (.054)	.273*** (.06)	.256*** (.061)
Size	5,217	6,779	5,629	5,477

$***$ 99% significant level $**$ 95% significant level $*$ 90% significant level , standard deviations in parenthesis.

Table C3: PH models for women and course levels

WOMEN	Course 1	Course 2	Course 3	Course 4
Age	-.024* (.013)	-.035*** (.011)	-.017 (.012)	-.025** (.013)
Age×age	-.0001 (.0002)	.00004 (.0002)	-.00009 (.0002)	-.00008 (.0002)
Married	-.234*** (.051)	-.24*** (.042)	-.224*** (.047)	-.259*** (.05)
Treatment	-.002 (.509)	.396*** (.037)	.418*** (.049)	.312*** (.073)
Treatment×disabled		.029 (.268)	-.487 (.436)	1.76* (1.016)
Treatment managed by INEM		.421*** (.12)	.553*** (.213)	.093 (1.004)
Active labour demand in months (dmonth)	-.063*** (.004)	-.062*** (.003)	-.062*** (.004)	-.062*** (.004)
dmonth2	.0008*** (.00009)	.0008*** (.00007)	.0008*** (.00008)	.0008*** (.00009)
English	.025 (.036)	.033 (.03)	.042 (.033)	.035 (.035)
Galician	.289 (.232)	.018 (.211)	.392* (.217)	.341 (.228)
Basque	.203* (.121)	.121 (.11)	.204* (.115)	.216* (.12)
Disabled	-.27* (.144)	-.248* (.144)	-.258* (.144)	-.261* (.144)
Worker receives benefits	-.321*** (.051)	-.27*** (.043)	-.296*** (.046)	-.3*** (.049)
Help has been disappeared	.387*** (.065)	.381*** (.052)	.374*** (.06)	.405*** (.064)
Size	7,221	9,694	8,309	7,572

$***$ 99% significant level $**$ 95% significant level $*$ 90% significant level , standard deviations in parenthesis.