Modelling Short Unemployment in Europe*

Antonella D'Agostino
Dip.Metodi Quantitativi, Università di Siena
e-mail: anto@turing.dmq.unisi.it
Fabrizia Mealli
Dip. di Statistica "G. Parenti", Università di Firenze
e-mail: mealli@ds.unifi.it

1 Introduction

In contemporary society the growing interest in longitudinal data comes as no surprise, principally because "stability" and "change" represent ubiquitous experiences in modern time. Changes occur within the individuals' life course; in particular, changes in the individuals' work history are interesting both from a socio-economic and a political point of view. In economic and social sciences a lot of attention has been paid to the duration of unemployment; recently, the necessity to have an overview of European labour markets has been recognized .

The discussion on the reasons for unemployment in Western countries has recently focused on the institutional regulation of labour markets and how this regulation obstructs market clearing (Grubb and Wells 1993; Sibert 1997). In spite of these efforts, no conclusive and coherent results have emerged and it seems that there is no simple relationship between the institutional regulation of the labour market and the *level* of unemployment, i.e.

*This paper is based on work carried out during a visit at the European Centre for Analysis in the Social Sciences (ECASS) at the University of Essex supported by The European Commission Training and Mobility of Researchers- Access to Large Scale Facilities programme. The work was also co-financed by Murst funds for the projects "Lavoro e disoccupazione: questione di misura e di analisi" and "Qualità ed efficacia dell'istruzione universitaria e della formazione professionale: metodi e modelli di valutazione".

how many unemployed people there are (see among others Scarpetta, 1996; Nickell, 1997) as well as the characteristics of the unemployment, i.e. who is unemployed.

Despite the unification of most of Western Europe in terms of circulation of goods and people, there are still considerable differences among labour markets in the various countries. For this reason, statistical comparisons are often very difficult to carry out, both because of the differences among the ways data are collected and because of intrinsic differences in educational systems, economic environments, institutional regulations and so on. At least the first problem can be overcome thanks to a longitudinal survey, named ECHP (European Community Household Panel), conducted by Eurostat in many European countries since 1994. Indeed, the European Panel was developed to collect comparable information, at a European level, on households and individuals in order to monitor the household/individual socio-economic situation and to help identify economic and social policy measures at a community level.

The aim of this paper is to analyse and make comparisons of unemployment experience in various European countries. We focus on the process of *exit* from unemployment using survival analysis: the transition from unemployment to employment is studied by means of flexible duration models.

Three particular features characterise our study. First, the analysis concentrates on the experience of short spells of unemployment in Europe over two years; second we pay attention to the specification of different models by analysing generalized residuals and, third, we will try to simplify the interpretation of the results using some synthetic measures computed for individuals with different characteristics in order to show how much the differences across countries are due to different labour markets and how much they depend on individual heterogeneity.

The paper is organised in the following way. Section 2 presents the ECHP data. In section 3, after a short introduction of the Cox model, some validation measures are outlined and in section 4 the empirical results are shown. Some final remarks end the paper.

2 Data

Since 1994 Eurostat coordinates a longitudinal survey, named ECHP (European Community Panel Survey). This survey is conducted every year in some European countries in order to monitor the household/individual socio-economic situation. Longitudinal data have become a useful tool for the analysis of many aspects, such as the transition from one state of activity to another, which would be not feasible with cross-sectional data. Such longitudinal aspects are of great policy relevance, especially in times characterized by high unemployment and by the rise of insecure forms of employment. Indeed, one of the main aims of this complex project was also to build up a regular European panel for territorial comparisons, that has not been possible until now.

As far as work histories are concerned, one part of the questionnaire reconstructs, retrospectively,

individual work histories during the last year. Each individual interviewed had to report all changes in his/her condition and also the time when they occurred, following a-monthly calendar of activity status. The employment classification follows the main activity criterion: an individual working at least 15 hours per week is classified as employed.

At present, only the first two waves, 1994-1995, for each country are available for the user's data base. The reference period for the retrospective data covers two years, from January 1993 until December 1994. No information is available about job history before January 1993; some information could be drawn from the panel data but it is only partial and relates to people that had a job at the interview time.

Focussing on unemployment spells, we found a lot of left censored observations in the data; explicit treatment of left censoring inside a duration model usually relies on very strong assumptions concerning the dynamic data generating process such as, for example, the assumption of an exponential hazard function, that usually lead to misleading interpretation of phenomena. In order to avoid the presence of left censored observations a flow sample of unemployment spells has been drawn. Flow sampling consists on randomly sampling the population of people entering or leaving a given state over an interval of time. This kind of sampling has the advantage that does not bias the population distribution (Lancaster, 1990) as for example stock sampling does. In this study a flow sample was drawn, taking the first unemployment spell of each individual that started in the reference period.

As the reference period covers two years, our analysis is necessarily concentrated on short unemployment spells; for this reason the conclusions we can draw do not refer to the problem of long unemployment. Nevertheless the aim of the paper is an interesting one, because the analysis of short spells of unemployment can help understanding differences among countries; it is useful to show how much these differences are due to the different labour markets or/and how much they depend on individual heterogeneity.

3 Modelling unemployment duration: the Cox model

As already mentioned in section 1, the main aim of the paper is to analyse, and make comparisons of, unemployment experience in various European countries. For this reason, from a methodological point of view, it is useful to specify a robust and flexible model w.r.t. to mispecification errors concerning, for example, the presence of unobserved heterogeneity and the form of duration dependence.

Even if the Cox model (Cox, 1972) preserves the proportionality assumption is certainly the one that matches such requirements. In the following sections we briefly introduce the model and the estimation method, together with some validation measures based on generalized residuals.

3.1 Basic framework

In the proportional hazard, or Cox, model, the effect of the exogenous variable is specified by a function that depends on exogenous variables only and multiplies a baseline hazard function of mispecified form.

Let T_i be a positive random variable, representing the length of time (duration) spent by an individual i ($i = 1...N$) in a given state. The Cox model defines the hazard function of T_i as:

$$\lambda\left(t|\mathbf{x}_i,\theta\right) = \lambda_0\left(t|\alpha\right) f\left(\mathbf{x}_i,\boldsymbol{\beta}\right),\ \boldsymbol{\theta} = \left(\alpha,\boldsymbol{\beta}\right), \tag{1}$$

where $\lambda_0\left(t|\alpha\right)$ is the baseline hazard function and f is a known function and $\mathbf{x}_i$ is a vector of exogenous variables.

When interest is focused on the role of the exogenous variables, α is often a nuisance parameter and only $\boldsymbol{\beta}$ is of interest. In such a case, $\lambda_0\left(t|\alpha\right)$ is specified in the most general way as $\lambda_0\left(t|\alpha\right) = \alpha\left(t\right) = \lambda_0\left(t\right)$, i.e. as a functional parameter. The model is known as the semiparametric Cox model (Cox, 1972, 1975; Tsiatis, 1981), with $\boldsymbol{\theta} = (\alpha,\boldsymbol{\beta})$, where α takes its value in a functional space, whereas $\boldsymbol{\beta}$ takes its value in a (finite dimensional)

Euclidean space.Obviously the function $f(\mathbf{x}, \boldsymbol{\beta})$ should be non-negative, so an easy way to obtain that property without restriction on $\boldsymbol{\beta}$ is the log-linear specification

$$f(\mathbf{x}, \boldsymbol{\beta}) = \exp\left(\mathbf{x}'\boldsymbol{\beta}\right), \ \boldsymbol{\beta} \in \Re^k. \tag{2}$$

It is well known that a direct approach to computing the marginal likelihood (Kalbfleisch and Prentice, 1973) for the proportional hazard model when ties are present creates a difficult problem (Cox and Oakes, 1984). Breslow (1974) and Efron (1977) have suggested approximations to such a computation, which are satisfactory except when the data exhibit heavy ties. In this case a computationally feasible expression for the likelihood has been introduced by DeLong, Guirguis and So (1994). The marginal likelihood[1] for $\boldsymbol{\beta}$ is

$$L_B(\boldsymbol{\beta}) = \prod_{i=1}^{N} \left\{ \int_0^{\infty} \prod_{j=1}^{d_i} \left[1 - \exp\left(\frac{\exp\left(\mathbf{x}_j'\boldsymbol{\beta}\right)}{\sum_{l \in R_i^*} \exp\left(\mathbf{x}_l'\boldsymbol{\beta}\right)} \times t \right) \right] \exp(-t)\, dt \right\}, \tag{3}$$

where $\mathbf{x}_j$ is the vector of explanatory variables for the jth individual, R_i^* is the set of individuals whose events or censored times exceed t_i or whose censored times are equal to t_i and d_i is the multiplicity of failures at t_i, that is, d_i is the size of the set D_i of individuals that fail at t_i. In other words, the exact conditional probability is computed under the proportional hazard assumption that all tied event times occur before censored times of the same value or before larger values. This is equivalent to summing all terms of the marginal likelihood for β that are consistent with the observed data (Kalbfleisch and Prentice 1980; DeLong, Guirguis, So, 1994).

3.2 Validation measures

Referring to Fleming and Harrington (1991) and Andersen *et al.* (1992), the semiparametric Cox model is a special case of a multiplicative hazards model. Consider a set of n subjects such that the counting process $N_i \equiv \{N_i(t), t \geq 0\}$ for the ith subject represents the number of observed events experienced over time t. Step functions with jumps of size $+1$ are the sample

[1] Kalbfleisch and Prentice (1980) have derived an alternative expression for $L_B(\boldsymbol{\beta})$, which is computationally more difficult to evaluate than this integral representation.

paths of the process N_i, where $N_i(0) = 0$. Again, let $\boldsymbol{\beta}$ denote the vector of unknown parameter, the multiplicative hazards function $\lambda(t|\mathbf{x}, \boldsymbol{\beta})$ for N_i is given by

$$Y_i(t)\, d\lambda(t|\mathbf{x}_i, \boldsymbol{\beta}) = Y_i(t) \exp\left(\mathbf{x}_i^{'}\boldsymbol{\beta}\right) d\lambda_0(t), \tag{4}$$

where $Y_i(t)$ indicates whether the ith subject is at risk at time t (specifically, $Y_i(t) = 1$ if at risk and $Y_i(t) = 0$ otherwise); $\mathbf{x}_i$ is a vector of exogenous variables; $\lambda_0(t)$ is an unspecified baseline hazard function.With this notation it is easy to define the martingale residual at t as

$$\widehat{Mart}_i(t) = N_i(t) - \int_0^t Y_i(\delta) \exp\left(\mathbf{x}_i^{'}\widehat{\boldsymbol{\beta}}\right) d\widehat{\lambda}_0(\delta). \tag{5}$$

The martingale residual in (5) estimates the difference over $(0, t]$ between the observed number of events for the ith subject and the conditional expected number of events. $\widehat{Mart}_i \equiv \widehat{Mart}_i(\infty)$ is referred to as the martingale residual for the ith subject. For the Cox model, with no time-dependent exogenous variables, the martingale residual for the ith subject with observation time t_i and event status d_i ($d_i = 0$ if t$_i$ is a censored time and $d_i = 1$ otherwise) is given by

$$\widehat{Mart}_i = d_i - \widehat{\lambda}_0(t_i) \exp\left(\mathbf{x}_i^{'}\widehat{\boldsymbol{\beta}}\right). \tag{6}$$

The martingale residuals have the drawback of assuming values between 1 and $-\infty$; a characteristic that often make their plot not easy to interpret. An useful transformation of the martingale residuals are the deviance residuals: they have the same properties but the advantage to assume values between $-\infty$ and $+\infty$. For the Cox model the deviance residual reduces to

$$\widehat{dev}_i = sign\left(\widehat{Mart}_i\right) \sqrt{2\left[-\widehat{Mart}_i - d_i \log\left(d_i - \widehat{Mart}_i\right)\right]}. \tag{7}$$

These residuals are very useful as a validation measure of the model. Like the plot of residuals in linear regression analysis, if the model is well specified the plot of such residuals versus the linear predictor has to show no particular behavior, but they must be distributed at random around zero. In practice, a smoothed plot obtained by a non parametric estimator, like Kernel-estimator or spline (see Hardle, 1990; Wand and Jones, 1995), can be a useful tool to show their behavior.

Table 1: Some characteristic of the sample (u.=unemployment, c.=censored)

Contries	sample size	% u. spells	% c. spells	mean duration (s.e.)
UK	672	7.73	44.64	5.41 (4.76)
BE	480	5.52	55.00	6.98 (5.23)
IE	697	8.03	46.20	5.76 (4.40)
GR	973	11.20	43.47	6.30 (4.28)
PT	674	7.76	51.48	7.22 (4.56)
DK	727	8.37	40.44	6.01 (5.01)
FR	1152	13.27	52.86	6.05 (5.01)
ES	2132	24.55	51.74	7.13 (5.00)
IT	1179	13.58	60.39	7.87 (4.62)
TOT.	8686	100	50.37	6.66 (4.82)

4 The experience of short unemployment in Europe

4.1 Sample characteristics

Before presenting the duration models, it is interesting to dwell on the results of the flow sampling performed.

As we can see from Table 1, the incidence of unemployment spells is concentrated in southern European countries; it is very high in Spain, followed by Italy, France and Greece. Even if the empirical mean durations are biased by the high percentage of censored observations, we note that Italy shows the highest mean duration of unemployment. On the contrary, the shortest mean duration is in UK: this could suggest that the British labour market is a dynamic one that prevents people to experience long unemployment spells and gives them more opportunities to find a job. Table 6 in Appendix shows the empirical hazards, estimated with a Life-Table method.

We note that they have different shapes and are not monotone. For Belgium, Portugal, Denmark and UK we observe a negative duration dependence: the chance of leaving unemployment is high for short duration and decreases with the time spent in the state. For the other countries the hazards show a different shape, with peaks around very short durations and around the longest ones. It is worth noting that some of the peaks maybe presumably due to a heaping effect, because people are inclined to smooth the real duration around some particular length. In Italy and France this effect appears stronger than in the other countries.

4.2 Estimation results

In the previous section we made some remarks about the data and the shapes of the empirical hazards; we are now interested in understanding if the differences across countries are due to different individual observed characteristics or if there is a significant "country effect", i.e. relevant differences among different labour markets.

In order to give an answer to this question, the analysis has to consider some individual factors that can influence the duration of unemployment.

The irregular shape of the empirical hazards suggests to specify a semi-parametric Cox model. One of the main problem related to the specification of a duration model is the presence of unobserved heterogeneity that can bias the duration dependence towards negative duration dependence; the Cox model is known to be more robust with respect to such mispecification.. Identification of the distribution of unobservables with single spell data relies on not testable assumptions (such as proportionality); for this reason we prefer not to deal with this problem (Lancaster, 1990; Elbers and Ridder, 1982).

Some preliminary analysis suggested to consider the following variables as relevant determinants of the length of unemployment: gender (SEX=1 if the individual is male); marriage (MARRY =1 if the individual is married); previous experiences of training (TRAINING=1 if the individual had an experience of training before 1993); area (NUTS1=1 if the area where the individual lives is an area with low unemployment rate); educational level (D1DEGR=1 if the individual has a university degree, D2DEGR =1 if the individual has a second level degree, D3DEGR =1 if the individual has a first level degree); start date of the current spell (months from January 1, 1993), date of birth (months from first January 1900) and the date of birth square.

In Tables 2, 3, 4 estimation results are reported. Smoothed plots of the deviance residuals are shown in Appendix (Table 7): there is no suggestion for bad specification of the models[2], smoothed lines do not have a particular

[2] The kernel smoothing was performed under a normal distribution assumption and using the following "rule of thumb" to determine the bandwidth of the kernel density: $\widehat{h}_0 = 1.06 \min\left(\widehat{\sigma}, \frac{\widehat{R}}{1.34}\right) n^{-1/5}$ where $\widehat{R} = X_{[0.75n]} - X_{[0.25n]}$ (Hardle, 1990; Wand and

Table 2: Cox model estimates for UK, BE, IE

Variables	Estimate (s.e.) UK	Estimate (s.e.) BE	Estimate (s.e.) IE
SEX	-0.059 (0.11)	0.356 (0.14)	-0.019 (0.10)
MINIZ	-0.0224 (0.009)	-0.030 (0.01)	-0.001 (0.01)
TRANING	0.086 (0.10)	0.614 (0.16)	0.275 (0.11)
ANNO	0.082 (0.039)	0.124 (0.09)	0.066 (0.05)
ANNO2	-0.0007 (0.0003)	-0.0007 (0.0007)	-0.0004 (0.0004)
MARRY	-0.036 (0.12)	0.264 (0.17)	0.007 (0.166)
D1DEGR	0.789 (0.33)	0.729 (0.22)	0.986 (0.18)
D2DEGR	0.554 (0.33)	0.310 (0.22)	0.374 (0.15)
D3DEGR	0.487 (0.31)	-0.208 (0.25)	0.141 (0.16)
D1NUTS	0.109 (0.10)	-	-

behavior, but are arranged randomly around zero. Note that the presence of ties produces a particular scatter plot of the residuals with parallel lines; the reason of which should be clear from equations (6) and (7). Interpretation and implications of the estimates are not easy to handle because of the number of estimated models; for this reason in the next section some synthetic measures can help to comment the results.

Nevertheless, gender appears to have a significant effect in Belgium, Greece, France, Spain, Denmark and Portugal. The positive sign of the coefficient means that men have better chances to leave unemployment than women. In Italy gender does not seem to have an influence on the chance to leave unemployment. On the contrary the variable NUTS1 shows a very strong effect: people living in Southern Italy have less chances to leave unemployment. The estimated risk ratio is very big: living in Northern/Central Italy increases the probability to leave unemployment about 2 times (risk ratio 1.775).

On the contrary, the start date of the spell shows a significant effect in almost every country and the negative sign suggests that the later is the start of the spell more difficult it is to leave unemployment. This result is quite interesting; the entrance in the labour market was indeed worse in 1994 than during 1993. People having an experience of unemployment at the beginning of the reference period seem to have more chances to leave unemployment. This finding reflects the big recession in EC countries during 1993 (see among

Jones, 1995).

Table 3: Cox Model Estimates for GR, FR, ES

Variables	Estimate (s.e.) GR	Estimate (s.e.) FR	Estimate (s.e.) ES
SEX	0.439 (0.09)	0.294 (0.08)	0.473 (0.06)
MINIZ	-0.065 (0.01)	-0.026 (0.008)	-0.034 (0.006)
TRANING	-0.191 (0.11)	0.187 (0.09)	-0.143 (0.08)
ANNO	0.048 (0.03)	0.109 (0.04)	0.134 (0.02)
ANNO2	-0.0005 (0.0003)	-0.0007 (0.0004)	-0.0011 (0.0002)
MARRY	0.246 (0.11)	-0.220 (0.11)	-0.050 (0.08)
D1DEGR	-0.099 (0.14)	0.072 (0.13)	-0.037 (0.10)
D2DEGR	-0.233 (0.11)	0.064 (0.14)	-0.125 (0.11)
D3DEGR	-0.247 (0.13)	0.104 (0.10)	-0.131 (0.07)
D1NUTS	-	0.0007 (0.08)	-0.175 (0.07)

Table 4: Cox Model Estimates for IT, PT, DK

Variables	Estimate (s.e.) IT	Estimate (s.e.) PT	Estimate (s.e.) DK
SEX	0.139 (0.09)	0.194 (0.11)	0.264 (0.098)
MINIZ	-0.070 (0.01)	-0.059 (0.01)	-0.002 (0.009)
TRANING	0.081 (0.11)	0.097 (0.21)	0.112 (0.11)
ANNO	0.112 (0.04)	0.114 (0.04)	0.0790 (0.04)
ANNO2	-0.001 (0.0004)	-0.0008 (0.0004)	-0.0004 (0.00003)
MARRY	0.362 (0.13)	-0.164 (0.15)	0.121 (0.11)
D1DEGR+D2DEGR	0.181 (0.15)	0.002 (0.19)	0.167 (0.12)
D3DEGR	0.100 (0.14)	-0.238 (0.18)	0.053 (0.13)
D1NUTS	0.592 (0.09)	-	-

the others Michie and Smith, 1994; Amin and Tomaney, 1995). In Italy, for example, the political events of the early nineties caused a big economic crisis that had repercussions on the labour market, especially at the end of 1993.

Having an experience of training before 1993 increases the probability to leave unemployment in Belgium, Ireland, and France and reduces it in Spain and Greece. Training systems are very different across countries, thus results are very hard to comment. Systems of training in modern societies can not be understood in isolation, but must be considered in relation to the general organization of educational systems and in connection with the nation-specific employment system. In fact, the negative effect of training on the length of unemployment, found in some countries, could be due to the fact that in such countries training is offered to disadvantaged people. As we do not have information on these characteristics, training could capture also these effects. In addition, it is well known that training effects are affected also by a process of self-selection that is impossible to address with our data. Multiple spells data should be required; some examples, that could explain our results, are given in Gritz (1993) and Mealli and Pudney (1999).

The effect of age is also interesting, as will be shown in the next section.

4.3 Some synthetic measures

To simplify and summarize the results some synthetic measures were computed. We define a reference individual that is modal with respect to the characteristics used in the analysis and for him/her we compare the estimated survival functions for the different countries.

The reference individual is a man, he started an unemployment spell on the first of January 1993, he had no previous experience of training, he was born in 1969, he is not married, he has a first level degree, he lives in an area with low unemployment rate (for UK, Italy, France and Spain).

In Table 8 (in Appendix) the empirical survival functions and the estimated survival functions in the different countries are reported. The figures for individuals with the same characteristics present substantial differences across countries: thus the different shapes shown in the empirical survival functions are due not only to differences among individuals (i.e. different sample composition) but also to differences across labour markets.

The first figures in the Table 8 (in Appendix) refer to UK, France, Spain and Italy because of the presence of the variable D1NUTS in the models for these 4 countries. We note that the empirical survival functions are almost identical for very short durations; differences become more evident: for durations longer than 4 months. For example, an individual in Spain has a probability of surviving longer than 12 months which is higher than in the other countries (UK, Italy, France), while the same probability for France is the lowest. It is also interesting to note the effect of observed heterogeneity; for example in Italy, without controlling for heterogeneity, the probability of remaining unemployed after 12 months is 0.6; after controlling for it, the reference individual has a survival probability less than 0.3. On the contrary the UK empirical survival function is below that of the other three countries, but when we compare the survival functions for the reference individual the British labour market shows the worst performance.

The second two pictures in Table 8 (in Appendix) show the behavior of the same individual in the others 5 countries; here differences are observed also for short durations. The effect of observed heterogeneity seems to mark the differences across different labour markets. In particular, in Belgium the survival probability is always higher than in the other countries, while the reference individual has highest chance to leave unemployment in Portugal.

In order to examine the effect of age on the probability to leave unemployment, the expected durations for different date of birth were computed as

$$\widehat{\mu}_u = \sum_{t=\mathbf{0}}^{\infty} \widehat{S}\left(t|\mathbf{x},\theta\right), \tag{8}$$

where $\widehat{S}\left(t|\mathbf{x},\theta\right)$ is the estimated survival function of the model specified in (1).

In Table 9 (in Appendix) expected durations are plotted against date of birth [3].

In UK, Italy and Spain, a parabola seems to describe quite well the behavior of the expected durations (with a very symmetrical shape for Italy), while in Portugal, France, Denmark the shape is almost a straight line. That means that in the first three countries there are obvious difficulties to find a job for younger and older people, while in the other countries such difficulties interest mainly older people.

In particular in Italy an individual born in 1950 (e.g. 43 years old in 1993) has an expected duration of unemployment of 6 months, that is shorter than

[3] Even if in table 9 we reporte the graphics for each country, comparisons are justified only among countries where the variables ANNO and ANNO2 are significant.

Table 5: Expected duration of unemployment for individuals with different educational levels

	UK	BE	IE	GR	FR	ES	IT	DK	PT
University degree	5.87	5.46	5.32	5.34	7.18	8.70	6.68	6.22	5.59
Second level degree	7.16	7.75	8.29	5.95	7.22	9.32			
First level degree	7.55	10.38	9.45	6.02	6.99	9.36	7.11	6.84	6.73
Less than first level degree	10.29	9.36	10.13	4.91	7.60	8.44	7.65	7.14	5.60

that of a young individual born in 1970. As expected, unemployment in Italy is a problem that interests both young and old people. The minimum expected duration in England is that of people born in 1955. In Spain, very young people have more chances to experience shorter unemployment spells with respect to old people, but the overall duration level is higher than in Italy and England. For the other three countries (France, Denmark, Portugal) we can observe a decreasing shape, with shorter expected durations for young people.

The effects of educational levels on the expected duration of unemployment differ across countries as shown in Table 5: here, the expected durations for the reference individual are reported, changing one at a time the value of the three dummies, D1DEGR, D2DEGR and D3DEGR.

These effects are not significant in all of estimated models, so we limit our discussion to the significant ones. In England, Belgium and Irland people with a university degree have shorter expected duration in unemployment, while in Greece and Spain educational levels seem to have very little effect.

This result suggests that, as expected, in the countries placed in Northern Europe educational qualification have a positive effect on the chance to leave short unemployment spells; this effect is not so evident in countries placed in southern Europe.

5 Final remarks

In the paper we estimate semiparametric proportional hazard models for the duration of short unemployment spells in different European countries, using data drawn from the European Community Household Panel (ECHP). The analysis has shown a substantial heterogeneity among labour markets: the behavior of a reference individual is, in fact, strongly influenced by the labour market he belongs to.

In particular, estimation results show that in countries, such as Belgium, Greece, France, Spain, Denmark and Portugal women have, as expected, less chances to leave unemployment. Living area effects are particularly strong for Italy: people living in Southern Italy find it very hard to leave unemployment. Particularly interesting is the effect of age: in Italy, UK and Spain unemployment seems to interest both the younger and the older generations, while in Portugal, France and Denmark difficulties in leaving unemployment are encountered mainly by older people. Higher educational levels shorten unemployment duration in UK, Belgium and Ireland, while in Greece and Spain educational level does not seem to have strong effects on expected duration.

It is important to stress that the analysis is based on the specification of a simple model, the aim of the paper being that of exploring the potential information of the ECHP and make the interpretation of the results as easy as possible. Nevertheless, the analysis could be deepened by employing more complex model specifications that might include possibly endogenous variables representing past work history, such as previous experience of employment and/or unemployment.

References

AMIN A., TOMANEY J., (1995), *Behind the Myth of European Union: prospects for cohesion*, Routledge, London and New York.

ANDERSEN P.K., BORGAN Ø ., GILL R.D., AND KEIDING N., (1992), *Statistical Models Based on Counting Processes*, New York: Springer-Verlag.

BRESLOW N. (1974), Covariance Analysis of Censored Survival Data, *Biometrics*, **30**, 89-99.

COX D.R. (1972), Regression models and life tables (with discussion), *Journal of the Royal Statistical Society*, **B 34**, 187-208.

COX D.R. (1975), Partial Likelihood, *Biometrika*, **62**, 269-276.

COX D.R. , OAKES D.(1984), *Analysis of Survival Data*, London: Chapman and Hall.

DELONG D.M., GUIRGUIS G.H., SO Y.C. (1994), Efficient Computation of Sunset Selection probabilities with Application to Cox Regression, *Biometrika*, **81**, 607-611.

EFRON B. (1977), The Efficiency of Cox's Likelihood Function for Censored data, *Journal of American Statistical Association*, **72**, 555-565.

ELBERS C., RIDDER G (1982), True and Spurious duration dependence the identifiability of the proportional hazard model, *Review of Economic Studies*, **49**: 403-410.

FLEMING T.R., HARRINGTON D.P. (1991), *Counting Processes and Survival Analysis*, New York: John Wiley & Sons, Inc.

GRITZ M.R. (1993), The Impact of Training on the Frequency and Duration of Employement, *Journal of Econometrics*, **57**, 21-51.

GRUBB D., WELLS W. (1993), Employment regulation and patterns of work in EC countries, OECD *Economic Studies*, **21**.

HARDLE W. (1990), *Smoothing Tecniques with Implementation in S*, Springer series in Statitsics, Sringer Verlag.

LANCASTER T. (1990), *The Econometric Analysis of Transition Data*, Cambridge University Press, Cambridge.

KALBFLEISCH T., R. PRENTICE (1973), Marginal likelihoods based on Cox's regression and life model, *Biometrika*, **60**, 267-278.

KALBFLEISCH T., R. PRENTICE (1980), *The Statistical Analysis of Failure Time Data*, Wiley, New York.

MEALLI F., S. PUDNEY (1999), Specification Tests for Random-effects Transition Models: An Application to a Model of the British Youth Training Scheme, *Lifetime Data Analysis*, **5**, 213-237.

MICHIE J., SMITH J.G., (1994), *Unemployment in Europe*, Academic Press.

NICKELL S. (1997), Unemployment and Labour Market Rigidities: Europe versus North America, *Journal of Economic Perspective*, **3**.

SCARPETTA S. (1996), Assessing the role of labour market policies and institutional settings on unemployment: a cross-country study, in OECD *Economic Studies*, **26**, 1, 44-97.

SIEBERT H. (1997), Labour market rigidietes: at root of unemployment in Europe, *Journal of Economic Perspective*, **3**.

TSIATIS A. (1981), A Large Sample Study of Cox's Regression Model, *Annals of Statistics*, **9**, 93-108.

WAND M.P. , M.L. JONES (1995), *Kernel Smoothing*, Monographs on Statistics and Applied probability 60, Chapman & Hall.

Appendix

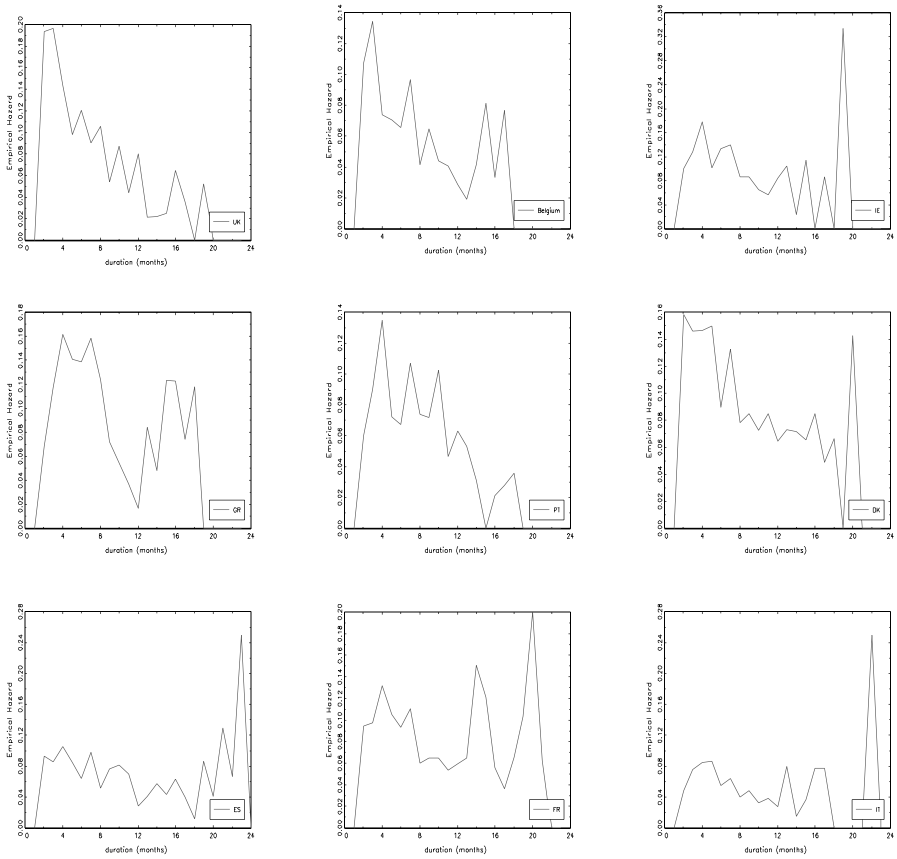

Table 6: Empirical Hazard (Life Table Method)

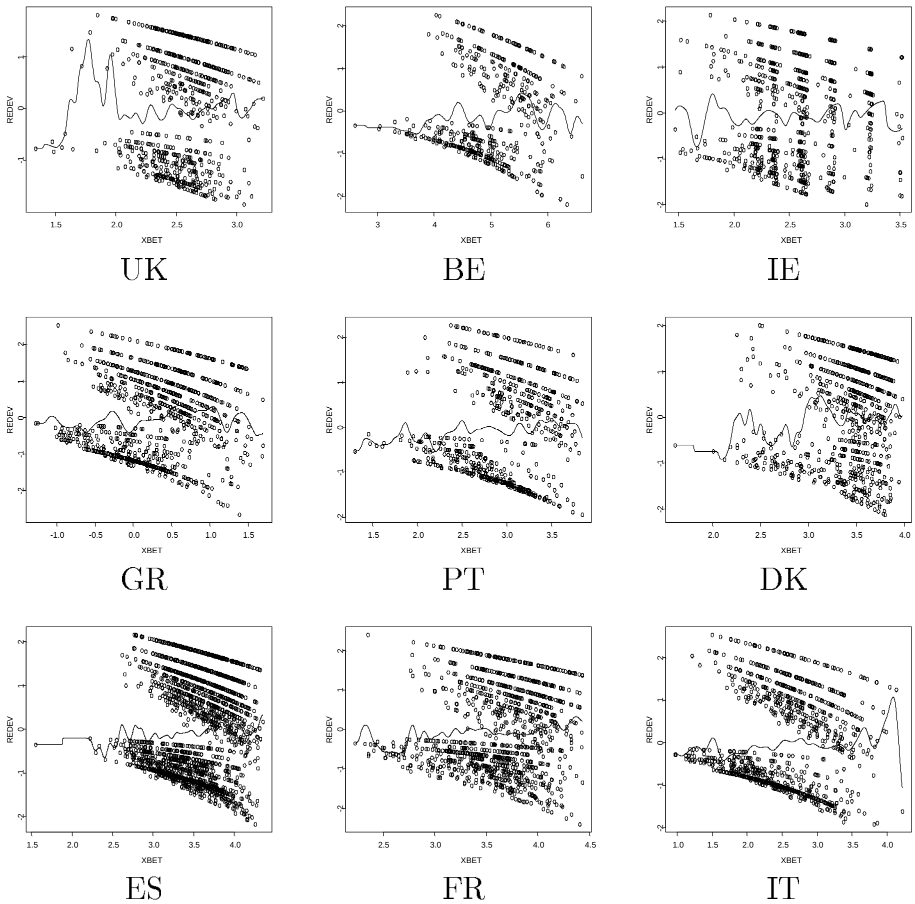

Table 7: Deviance Residuals Smoothed Plot

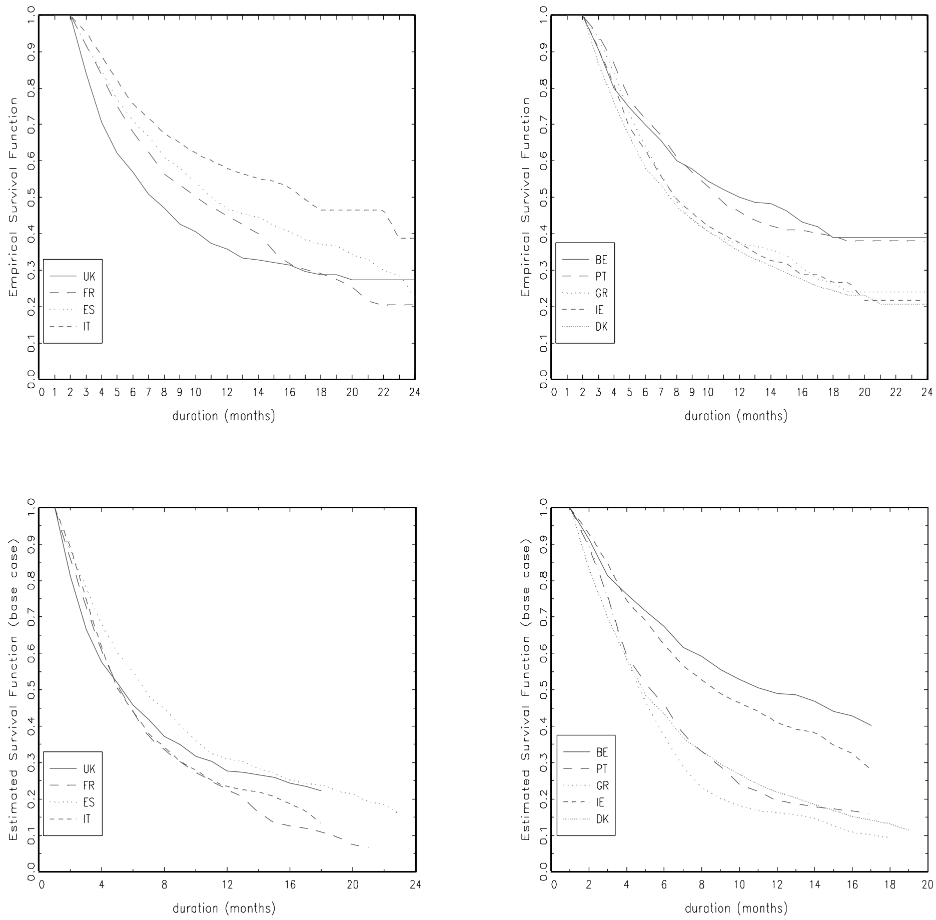

Table 8: Estimated Survival Functions

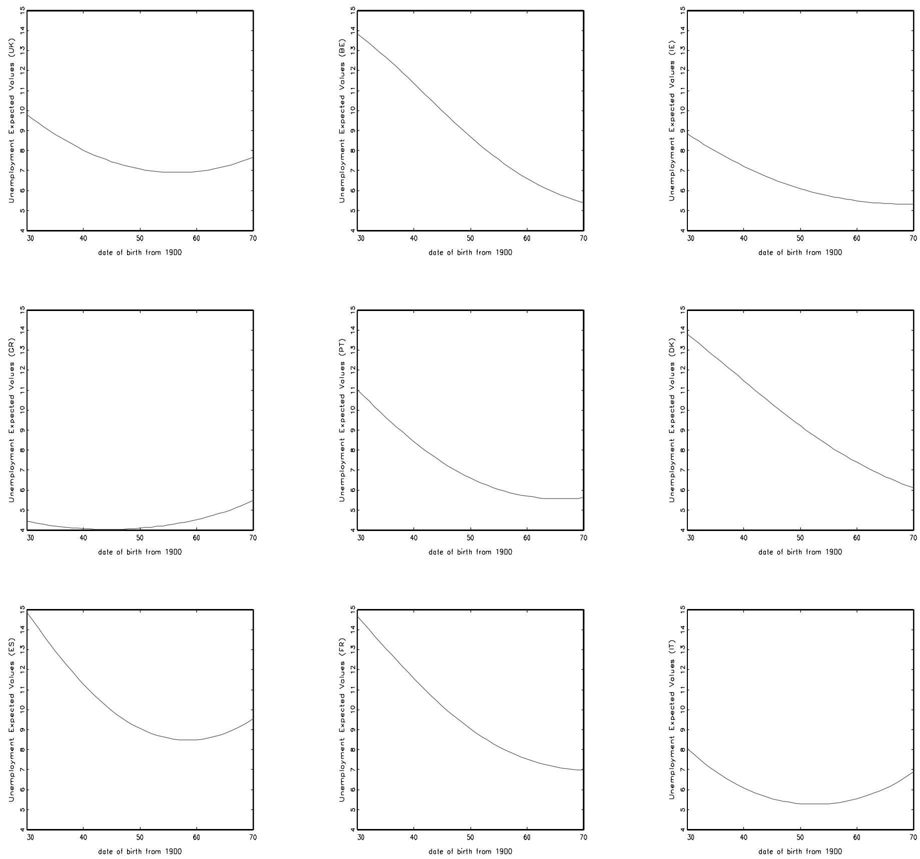

Table 9: Unemployment Expected Durations for different generetions and different countries: UK, BE, IE, GR, PT, DK, ES, FR, IT