

Distribution of Unemployment Spells and Survival Analysis: The Case of Bosnia and Herzegovina

Narela Bajram

International University of Sarajevo, Faculty of Business and Administration, Hrasnicka Cesta 15, Ilidža 71210
Sarajevo, Bosnia and Herzegovina

Article Info

Article history:

Received 17 Sep. 2013

Received in revised form 17 Oct 2013

Keywords:

Unemployment durations, survival analysis and frailty model.

Abstract

It is not enough to study the labour market only by analysing the static variables like rates of employment and unemployment. For the decision making on the labour market it is essential to see the movement of people into and out of jobs, the extent to which they can or cannot quickly find alternative employment and to which extent different groups of the labour force are more affected than others.

Using the frailty models, as a specific area in survival analysis, we show that the durations of unemployment are sensitive to the educational level and gender. First, females are experiencing significantly longer durations of finding job than man. Second, better-educated individuals appear to find job more quickly than the less-educated. Finally, there is evidence of duration dependence in unemployment in Bosnia and Herzegovina. The re-employment hazard exhibits positive duration dependence in the first 59 months, and then declines approximately 60 months.

1. INTRODUCTION

It is not enough to study the labour market only by analysing the static variables like rates of employment and unemployment. For the decision making on the labour market it is essential to see the movement of people into and out of jobs, the extent to which they can or cannot quickly find alternative employment and to which extent different groups of the labour force are more affected than others.

Unemployment duration refers to the amount of time that an individual remains unemployed. It is an important variable which can explain the changes in labour markets and it is widely used in the job destruction and job creation models when analysing the flows between employment, unemployment and out the labour force.

In the "Employment in Europe 2009" a example illustrates the importance of estimating the unemployment duration: "a 10% unemployment rate can represent two entirely different realities in terms of their implications for the welfare of those affected: one where every individual in the labour force experiences unemployment during 5

weeks per year and another where 10% of the population are unemployed during the ole year." A labour market in which few people become unemployed, but where those who do are likely to remain unemployed for a very long time is likely to more damaging than one in which are many more who become unemployed but remain in that position for only a short period of time.

The welfare of unemployed depends on the probability of leaving unemployment and obtaining a job (hazard rate). Duration of unemployment is sensitive to the level on unemployment benefits and to the entitlement period to benefits. In this paper we try to determine other factors (gender and education level) which may produce variation of unemployment duration, using the frailty models as a specific area in survival analysis.

The paper is organized as follows. Section 2. discusses the frailty model as special area in survival analysis. Section 3. A case study involving time being unemployed in Bosnian labour market is introduced, as well as the factors that determine the unemployment duration and their effect. Section 4. Gives the population (unconditional) predictions. Section 5 concludes.

2. FRAILTY MODELS AS AN EXTENSION OF A SURVIVAL REGRESSION MODEL

The statistical analysis of lifetime data (or more exactly, time-to-event, event-history or duration data) plays an important role in medicine, epidemiology, biology, demography, economics, engineering and other fields. It has expanded rapidly in the last three decades, with works having been published in various disciplines in addition to statistics. But what distinguished survival analysis from other fields of statistics? Why does survival data need a special statistical theory? The main reason is that we are observing something that develops dynamically over time. Survival analysis is generally defined as a set of methods for analyzing data where the outcome variable is the time until the occurrence of an event of interest. The event can be deaths, occurrence of a disease, marriage, divorce, unemployment and etc. The time to event of survival time can be measured in days, weeks, years, etc.

Unlike ordinary regression models, survival methods correctly incorporate information from both censored and uncensored observations in estimating important model parameters. The dependent variable in survival analysis is composed of two parts: one is the time to event and the other is the event status, which records if the event of interest occurred or not. One can then estimate two functions that are dependent on time, the survival and hazard functions. The survival function gives, for every time, the probability of surviving (not experiencing the event) and the hazard function gives the potential that the event will occur. While these are often of direct interest, many other quantities of interest may subsequently be estimated from knowing either the hazard or survival functions. It is generally of interest in survival studies to describe the relationship of a factor of interest to the time to event, in the presence of several covariates', such as gender, education level, etc. A number of models are available to analyze the relationship of a set of predictor variables with the survival time. Methods include parametric, nonparametric and semi-parametric.

Parametric survival models are one of available models which analyze the relationship of a set of predictor variables with the survival time. The distribution of the response in these models are chosen to be consistent with what one would see if the response is time-to-failure. Examples of such distributions are the exponential, Weibull, log-normal, log-logistic, Gompertz, and the generalized gamma, among others. The distribution of the time-to-failure is characterized by the hazard function, which is the ratio of the probability density function to one minus the cumulative density function.

This paper focuses on frailty models, a specific area in survival analysis. The frailty approach is a statistical modeling concept which aims to account for heterogeneity, caused by unmeasured covariates. In statistical terms, a frailty model is a random effect model for time-to event data, where the random effect (frailty) has a multiplicative effect on the baseline hazard function.

In essence, the frailty concept goes back to work of Greenwood and Yule (1920) on "accident proneness". But the term frailty was first suggested by Vaupel et al. (1979) in the context of mortality studies, and Lancaster (1979) incorporated the frailty concept into a study of duration of unemployment. Hougaard (1984) discusses the ramifications of the assumed distribution of the frailty, whether gamma or inverse Gaussian.

3. APPLICATION

3.1. Data

The data used is obtained on the basis of the Labour Force Survey (ARS 2011) carried out by the statistical institutions in Bosnia and Herzegovina (the Agency from Statistics of Bosnia and Herzegovina-BHAS, the Federal Institute of Statistics-FZS, and the Republic Institute of Statistics of the Republic of Srpska-RZSPS). We consider a sample of 311 people in the period between January 2002 and November 2011.

In our case all observations are not complete. To be more specific, the subjects are not observed at the beginning ($t=0$) but instead initially come under observation at some time $t_0>0$. This is known as left-truncated observation.

Under observation at time $t=1$ we consider 311 people, because by the time we initially observe each person, he/she has already been unemployed. Otherwise we would not have observed him/her. Therefore, we represent this assumption by specifying an entry time equal to 1 when we set our dataset.

3.2. Effect of Gender and Education Level on Unemployment Duration

If the dynamics between the states is slow enough then the transition rates from unemployed to employ do not follow an exponential distribution. Han and Hausman (1990) and Meyer (1990) propose an estimator that assumes that the baseline hazard is piecewise-constant, to permit flexibility, and that the heterogeneity has a gamma distribution.

Because Bosnia and Herzegovina is characterized as a country with long time unemployment, we fit a Weibull-gamma model in the accelerated failure time metric (AFT) and estimate p , β , and θ . When $\theta=0$, then to measure the effects of the covariates we should use the standard survival model.

Table 1. Weibull regression -- accelerated failure-time form- Gamma frailty

| | | | | | |
|-----------------------------|------------------------|-----------|-------|-------|----------------------|
| No. of subjects = 279 | Number of obs = 279 | | | | |
| No. of failures = 279 | | | | | |
| Time at risk = 947 | | | | | |
| | Wald chi2(2) = 1003.03 | | | | |
| Log likelihood = -225.21004 | Prob > chi2 = 0.0000 | | | | |
| <hr/> | | | | | |
| _tl | Coef | Std. Err. | z | P> z | [95% Conf. Interval] |
| <hr/> | | | | | |
| gender | .1456818 | .0684568 | 2.13 | 0.033 | .0115088 .2798547 |
| educational-1 | .6527588 | .0246272 | 26.51 | 0.000 | .6044904 .7010271 |
| <hr/> | | | | | |
| /ln_p | 1.444067 | .1241995 | 11.63 | 0.000 | 1.20064 1.687493 |
| /ln_the | .6762199 | .1936626 | 3.49 | 0.000 | .2966482 1.055792 |
| <hr/> | | | | | |
| p | 4.237895 | .5263443 | | | 3.322243 5.405912 |
| 1/p | .2359662 | .0293069 | | | .1849827 .3010014 |
| theta | 1.96643 | .3808241 | | | 1.345342 2.87425 |

Likelihood-ratio test of $\theta=0$: $\chi^2(01)=44.94$ Prob>= χ^2 = 0.000

In Table 1 we estimate $\theta=1.966$ and a significant likelihood ratio test for the presence of heterogeneity. In standard models, a hazard ratio is interpreted as a proportional shift in the hazard function due to a unit change in the associated covariates. When using a frailty model, however, reported “hazard ratios” carry this usual interpretation only if comparing two hazards conditional on a given α . In our case the hazard ratio reported for the covariate gender as “all other things equal, including the frailty α , the hazard function for females is 0.14 times the hazard for males”. There is a significant difference by gender, males exit the unemployment state sooner than females.

Education is also one of the factors affecting the duration of unemployment (the p-value=0.000). In Figure 1 we present the survival curves on educational levels. We have considered 3 education levels: primary, secondary and university.

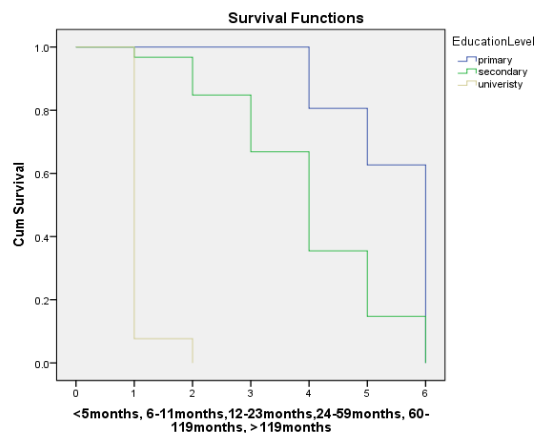
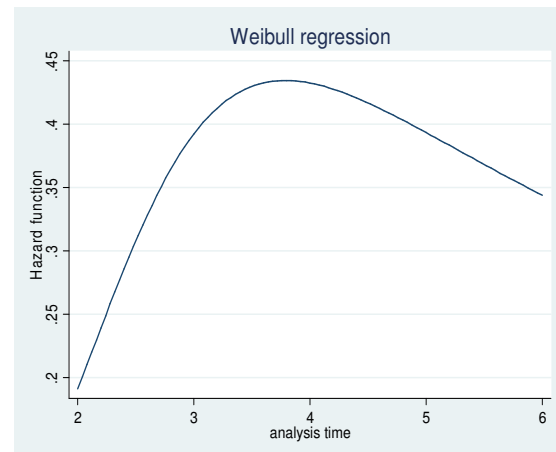


Figure 1 : Survival function by education level

We can notice that higher educated individuals appear to search for a job shorter (max 11 months) than persons with primary and secondary education.

4. OBTAINING UNCONDITIONAL PREDICTION

Prediction used after a univariate frailty model will give population (unconditional) prediction. The predictions are adjusted for the estimated frailty variance θ and for the assumed frailty distribution.



Note: 2→ 5-11; 3→11-23; 4→23-59; 5→59-119 6→above 119;

Figure 2: Estimated population hazard functions

Figure 2 shows the estimated hazard function. We estimate it as monotone increasing ($p=4,23$ which is greater than one), yet the hazard function $h_0(t)$ start decreasing past a certain point. Thus, at the beginning the probability that the person will exit the unemployment state is increasing but after a point (after 59 months being unemployed) this probability starts decreasing.

The initial period of positive duration dependence may result from greater search activity soon after becoming unemployed, or from individuals being recalled to their previous job after a shorter spell. But the individuals after a long period of job search and individuals with the worst employment prospects became discouraged or are entering in voluntary and governmental retraining programs which decrease the re-employment hazard among job searchers.

5. CONCLUSIONS

Using information contained in the Labour Force Survey (ARS 2011) of Bosnia and Herzegovina we analyzed the gender and education level as factors which may produce variation of unemployment durations.

Our results confirm that duration of unemployment is sensitive to educational level and gender. First, females are experiencing significantly longer durations than man before finding job. Second, better-educated individuals appear to find job more quickly than the less-educated. Finally, there is evidence of duration dependence in unemployment in Bosnia and Herzegovina. The re-employment hazard exhibits positive duration dependence in the first 59 months, and then declines approximately 60 months.

This study may be considered as a picture of the factors shaping unemployment duration, which may be helping policy makers. Because it is not enough to study the labour market, only by analyzing the static variables, like rates of employment and unemployment. Remaining unemployed for a very long time is likely to be more damaging than

having a higher unemployment rate and individuals being in that position for only a short period of time.

REFERENCES

Cox, D.R. (1971) Regression Models and Life-Tables. *Journal of the Royal Statistical Society B* 34, 187-200.

Employment in Europe 2009, European Commission, Directorate-General for Employment, Social Affairs and Equal Opportunities, Belgium, 2009.

Greenwood, M., Yule, G.U. (1920) An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society* 83, 255-79.

Han, A.K. and Hausman, J.A. (1990) Flexible Parametric Estimation of Duration and Competing Risk Models. *Journal of Applied Econometrics*.

Hougaard, P.(1984) Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika* 71: 75-83.

Lancaster, T. (1979) Econometric methods for the duration of unemployment. *Econometrica* 47: 397-96.

Meyer, B.D. (1990) Unemployment Insurance and Unemployment Spells", *Econometrica*, 58, 757-782.

The Labour Force Survey(ARS 2011) ,The Agency from Statistics of Bosnia and Herzegovina-BHAS, the Federal Institute of Statistics-FZS, and the Republic Institute of Statistics of the Republic of Srpska-RZSPS, Bosnia and Herzegovina,2011,availableonline at:
http://www.bhas.ba/tematskibilteni/LFS_2011_001_01_bh.pdf

Vaupel, J.W., Manton, K.G., Stallard, E. (1979) The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. *Demography* 16, 439-454.