Finite mixture modeling of unemployment duration

Lorenzo Corsini[1] and Paolo Frumento[2]

Abstract

We analyze the determinants of unemployment duration adopting a finite mixture modeling. There are several factors and characteristics that affect unemployment duration and our approach allows identifying clusters of individuals which differs in the way that those factors affect unemployment duration. The idea we are pursuing in our analysis is that similar characteristics and variables can have different effects depending on the specific individual. Our methodology, allows to directly estimating the existence of such clusters, without an a priori definition of them. We apply our methodology to German data on unemployed workers from 2002 and we identify three groups of unemployed individuals that react differently to the same inputs. The groups differ in terms of baseline hazard and on the effect that unemployment benefits have on duration. In particular, some groups have mildly increasing baseline hazard suggesting they put a lot of effort in the search from the very start while others appear to have increasing hazards suggesting that during the first part of the spell they are less active. Moreover benefits have different effect on these groups.

JEL CODES: C41; C51; J64; J65.

KEYWORDS: unemployment insurance; unemployment duration, survival analysis, finite mixture, wealth.

1. Introduction

There are several factors that determine unemployment duration of unemployed individuals. Some of them are related to the macroeconomic situation while others are more directly related to the individual characteristics of the unemployed workers or to the effects that labour market institutions have on the specific worker. From a theoretical point of view, job search theory has developed a framework that helps in understanding the process of job search and the resulting unemployment duration. According to this theory, the duration for a given individual depends on the matching technology (i.e. on the mechanisms that allow the demand to meet the offer), on the demand for his/her skill (that is, how many suitable vacant positions exist), on his/her search effort and on his/her reservation wage. The basic mechanism behind this theory is that, for a given matching technology, the demand for skills and the search effort determine the offers arrival rate

[1] Dipartimento di Scienze per l'Economia e l'Impresa, Università di Firenze. Email: lorenzo.corsini@unifi.it

[2] Karolinska Institutet, Institute for Environmental Medicine (IMM),Unit of Biostatistics, Stockholm. Email:paolo.frumento@ki.se

while the reservation wage determines the number of acceptable job offers. As it easy to understand, lower search effort and higher reservation wages increase unemployment duration.

From an empirical point of view, the analysis of the unemployment duration was mostly developed using survival analysis. Within this framework, it is assumed that the probability that individuals leave unemployment is made of two components: the first is the "pure" effect of time (usually called baseline hazard) which represents how the time component affects the search for a job and the second is given by all the characteristics of the individual that enhance or impair the pure effect of time. Standard analyses assume that baseline hazard is the same for all individuals and, similarly, that the effect of the covariates is the same for all individuals. While the analysis can be augmented to allow to introduce an priori clustering of the baseline hazard and of the effect of the covariates, in our work we aim to make a step further, modelling unemployment duration with a finite mixture approach. Using this approach we allow for the baseline hazard and for the effect of the covariates to be different for different groups (clusters) of the population and, moreover, we develop a procedure that identifies statistically those groups without resorting on an a priori definition of them.

Even within this approach, job search theory is still crucial in highlighting the mechanism behind duration and in selecting and interpreting the variables of interest. As a matter of fact, within the job search framework there is a strict connection between unemployment duration and unemployment insurance (UI) schemes and this relationship will be one of the topic to which we will apply our estimation methodology. According to search theory, unemployment benefits raise the reservation wages of workers thus reducing the number of acceptable job offers and increasing unemployment duration. Similarly, individuals receiving benefits are likely to be less financially strained and can afford to search less actively: this mitigation of financial stress ends up increasing duration. However, things are more complicated than this simple depiction because benefits themselves are a complex thing. UI schemes usually entail search requirements and re-employment counselling: these aspects can off-set the other drawbacks and benefits scheme could actually decrease unemployment duration.

In any case, this line of reasoning indicates that search effort and reservation wages depend on the degree of economic stress of individuals and, consequentially, personal and household wellbeing in general and economic situation in particular, can also play an important role in the determination of unemployment duration. Unemployed workers living in households that are richer or with less financial stress should feel less pressure to search actively for a job or to accept the very first offer they receive: therefore, over things being equal, unemployed from these kind of households should experience longer duration. Moreover, the actual effect of unemployment benefits may be different for individuals from different economic backgrounds. In fact, benefits may be less relevant in mitigating financial stress if individuals, or their households, can count on a more solid economic situation.

From a broader point of view, the effect of benefits may be different on different groups of individual and the finite mixture approach we are pursuing is thus particularly useful within this

context. In fact, the essence of our work will be exactly to develop a method of analysis and estimation able: i) to identify those groups, ii) to determine the characteristics that determine the belonging to a given group iii) to find the determinants of unemployment duration within each group. From what we have discussed above, the economic situation of the individual may be a key determinant for the belonging of a given group but other characteristics (education, gender, household composition and so on) can also be relevant and should be explored in an empirical analysis.

This far we have mentioned how unemployment benefits should be an important determinant of unemployment duration. However the same reasoning should be applied to all other characteristics that explain unemployment duration and thus we investigate whether the effects of all explanatory variables (age, education, family composition) can be different in different groups. In practice, it is possible that a single relationship (or model) is not enough to univocally link unemployment duration to the variables that explain it: on the contrary, it is possible that several relationships co-exist, each one belonging to a different cluster/group of the population. Our work exactly tries to develop and apply an estimation methodology that is able to highlight the existence of these clusters and determine which kind of individuals are more likely to belong to each cluster. While some emphasis will be put on the role of benefits, wealth and the economic background, we want to stress that the methodology we propose is suitable to highlight all the determinants of unemployment duration and all the determinants of clusters that may exist within this context.

Given these premises about the possibility that the determinants of unemployment duration are different for different groups/cluster of individuals, we develop a methodology to estimate survival regressions with a mixture of distributions. In details, we estimate a selection model which determines how the individuals' characteristics affect the belonging to a given group and then we estimate, through a parametric cox-regression, the effect of the characteristics, in each group, on the unemployment duration. Therefore, the groups are identified through estimation and not through some arbitrary a-priori and the effect of the covariates and the very baseline hazard are allowed to differ from cluster to cluster.

We make an application of our methodology using German data on employment, income, wealth and household composition for the year 2002 from the SOEP survey. This dataset contains detailed information that allows us to determine the exact unemployment duration, the receipt of unemployment insurance and several other socio-economic data. The choice of the year is due to the specific topic that was surveyed in that year. In fact, the 2003 edition of the SOEP contains information on the precise structure of the financial situation in terms of assets and liabilities, allowing us to assess if and how wealth and financial stress affect unemployment duration. To obtain a more homogenous group of observations, we focus only on workers that have just become unemployed so that the duration of unemployment before the period of observation is null for all the individuals.

The work is organized as follows: in section two we present a brief review of the empirical literature on unemployment duration; in section three we discuss how, from a theoretical point of view, UI schemes, wealth and financial condition could affect unemployment duration; in section four we describe the methodology we use for the analysis; in section five we apply this methodology to perform the empirical analysis on German data and discuss possible interpretations of the results and in section six we conclude.

2. A Literature review of empirical analyses

Empirical literature is rich of analyses on unemployment duration and unemployment benefits: a good review of the first empirical studies on the subject of unemployment duration and UI is contained in Atkinson and Micklewright (1991) where they clearly indicate the duration models as the main tool to estimate the effect of unemployment benefits. A group of relatively more recent studies focused on the role of eligibility criteria on the search effort and unemployment duration. Those studies are based on field experiments and perform causality analysis distinguishing between treated/non treated groups, trying to assess whether the criteria imposed to be eligible for benefits affect or not search behaviour. The results of the conclusion of these studies are mixed: Klepinger et al. (2002) performs a causality analysis using the data from Maryland UI work-search demonstration (a plan that randomly assigned benefits recipients to different search criteria) and show that stricter criteria improved search efforts and reduced unemployment duration. Somehow differently, Ashenfelter et al. (2005) exploit differences in the eligibility criteria of different American states (Connecticut, Massachusetts, Virginia and Tennessee) to conduct an experiment and finds out that stricter search criteria do not affect sensibly the access to benefits. Manning (2009) uses a difference in differences estimations to capture the treatment effect on unemployment duration, using the change of unemployment benefits regulation that happened in UK in 1996: his results indicates that criteria affect the access to claims but stricter criteria discourage workers to effectively meet the search requirement and thus do not facilitate the transition to a employment.

Several works have also tried to assess the role of unemployment benefits maximum duration on unemployment duration: Nickell and Layard (1999) give an assessment of this issue and suggest that benefits duration and unemployment duration are positively correlated. Among the more recent analyses of this aspect, Lalive (2007) performs a regression discontinuity analysis trying to assess the role that maximum duration of the unemployment benefits has on unemployment length. In this case, the discontinuity is given by the differences between benefits duration for workers older than 50 years that is inherent in Austrian legislation: the results show that a large extension of benefits duration increase the length of unemployment spells but this very effect is not present when the extension is small.

While the analyses of unemployment duration and UI are abundant only a few of them cover the role of wealth, financial stress and household background and these aspects are still to be fully

assessed. An interesting analysis on how wealth affects the transition from unemployment to employment is contained in Bloemen (2002) where a proxy variable for wealth is computed and its effect on the probability of obtaining a job is tested; however, no analysis on duration is performed nor is the role of benefits explored. The interaction between wealth and benefits in determining unemployment duration has been tackled more in details in the works by Chetty (2005 and 2008, with the latter being an extend version of the former): in them the author tries to disentangle the moral hazard and liquidity constraint effects that benefits have on unemployed workers. To achieve this, a cox hazard models is used to perform estimations on US data for different groups of people, defined by their being liquidity constrained or not. In addition, estimation with data on lump sum severance payment is performed: since no moral hazard effect can be derived from these kinds of payments those data allow explicitly focusing on liquidity constrain effects. Card, Chetty and Weber (2007) use Austrian data and a regression discontinuity approach to evaluate the effect of a lump sum benefits and of maximal potential duration of benefits on the search behaviour and unemployment duration, where the discontinuity stems from the eligibility criteria for the unemployment benefit scheme in Austria: since the effect of the severance payment and of maximum duration seems to be similar, they conclude that liquidity constraints motives seems to be more relevant than moral hazard one (which should be absent in the case of the severance payment).

As for the mixture distribution approach we are to use, we are aware of only very few contributions that uses similar methodologies within the unemployment duration analysis. In particular, McDonald and Butler (1987) stress how the use of mixture distribution could enrich the reliability of duration analysis. Their contribution, however focus only on the possibility that the baseline hazard may be different in different groups of the population and do not include any selection model for the participation to those groups nor allow for the fact that the rest of explanatory variables can have different effects in different groups. A package to estimate discrete duration regressions with mixture of distribution in unobserved individual heterogeneity is provided in Jenkins (1998): even in this case there is no selection model for the identification of the different group.

Our analysis improves on this line of research, trying to detect and estimate clusters within the determinants of unemployment duration and delivers an application that focus on the role that wealth and financial stress has on unemployment duration and how these aspects can interact with unemployment benefits and on other characteristics.

3. On the effect of benefits and economic well-being on unemployment duration

Economic theory, and in particular job search theory, has stressed how UI schemes can affect unemployment duration. At a basic level, there is a connection between unemployment benefits and duration that is obtained through the reservation wage, which is increasing in the level of benefits, and through the search effort, which is decreasing in benefits (see Rogerson *et al.* 2005

for an analytical discussion of these effects). Clearly, higher reservation wage and lower effort produce longer duration so that a basic positive relationship between benefits and unemployed duration is obtained. However, these relationships, though logically consistent, are also too simple and forego some important aspects of the problem as well some key features of the real working of the UI schemes. If we go more in details, there are at least two different reasons that explain the positive relationships between benefits, reservation wages and duration: the first is a moral hazard argument and the second is related to the drop in income and to the role of liquidity constraints. According to the first, unemployed workers are basically paid for being idle and only as long as they are idle (even if some eligibility criteria may force the workers to actively search for a job). Consequently, workers may prefer to postpone their search and to turn down job offers to extract the most they can from this system. The second argument is instead related to the drop in income faced by the unemployed, something that is further exacerbated by the fact that unemployed workers usually face credit constraints. In these circumstances, consumption largely falls and, in the absence of a proper consumption smoothing, utility is greatly reduced: as a consequence, workers are eager to prevent their fall in consumption and are willing to search actively and to accept any job offer they receive. It is clear that unemployment benefits mitigate the drop in consumption and they allow the workers to stay closer to their optimal consumption path without having to accept the very first job offer they receive. The distinction between moral hazard and liquidity constraint is particularly relevant: in fact, if duration is increased because of the former, we move away from a social optimal situation (with benefits recipients acting as parasites) but if it is the latter argument that is relevant then unemployment benefits can increase social welfare allowing a better consumption smoothing.

These considerations can be extended to include the role that savings, wealth and economic well-being in general have in the determination of unemployment duration. In fact, unemployed workers that can count on some financial resources or other economic support probably obtain a better consumption smoothing during their unemployment spell. This should prevent an immediate fall in consumption so that the unemployed worker should feel less pressure to search for a job or to accept the very first offer received: therefore, other things being equal, wealthier unemployed should experience longer duration. More in general, unemployment duration should be affected by the financial pressure and thus, any factor that affects financial pressure could also affect duration.

Given these arguments, there might be some interaction between the effects of unemployment benefits and of wealth (or, in a broader sense, of economic situation). In fact, for wealthier recipients, the benefits should not be so important in mitigating the liquidity constraints and the drop in consumption: unemployment duration should thus be only mildly affected by benefits through this channel. Therefore, for recipients that are better off, benefits should affect unemployment duration weakly or not at all and any positive effect of benefits on unemployment duration should be attributed mostly to moral hazard motives.

This argument can be further extended so that the effect of benefits on duration can vary not only on the base of wealth and also other characteristics can define clusters where the effect of

benefits is different. Moral hazard is a possible example of this, as benefits should produce larger increase in duration for individuals that are more prone to moral hazard behavior but several other characteristics can contribute to the definition of these clusters.

This far, we have discussed how benefits could increase unemployment duration but, on different grounds, it has also been argued that, given the presence of eligibility criteria, benefits schemes offer incentives to search actively for a job. In fact, among the eligibility criteria is often included the necessity to provide proof of being actively searching, so that workers on benefits are somehow forced to search for a job. Moreover, another common eligibility criterion is the necessity to have been in employment in the period before claiming the benefits: this provides further incentives to workers to search for and to accept jobs so that, in the future, they can be covered again by unemployment insurance. Finally, together with income support, UI schemes usually provide some forms of job counseling and training: these actions should enhance re-employment probabilities.

4. Finite Mixture Models

We present now the methodology we adopt for the estimation of unemployment duration, of its determinants and of clusters (components in technical terms). As already mentioned, our methodology is an application of finite mixture to survival analysis.
Finite mixtures arise when the distribution of some outcome Y is assumed to depend on a latent categorical variable Z, playing the role of an unobserved covariate. The marginal distribution is said to be a k-components mixture, where k is the number of modalities of Z. The mixture is said to be ”finite” or ”nonparametric”, in contrast with parametric mixtures (e.g., mixed effects models) where Z is supposed to be a continuous random variable.
In some case, Z has a clear interpretation and can be seen as a group membership based on a well-defined partition of the population. For example, this is the case in the analysis of randomized experiments suffering from noncompliance with the treatment assignment, where Z represents the compliance behavior and is partially observed. In other settings, Z is interpreted as an unobserved source of heterogeneity. Eventually, this class of models can be thought of as a flexible alternative to fully parametric approaches.

The density of a finite mixture model is the following:

$$f(y \mid \boldsymbol{\vartheta}, \boldsymbol{p}) = p_1 f_1(y \mid \boldsymbol{\vartheta}_1) + \ldots + p_k f_k(y \mid \boldsymbol{\vartheta}_k)$$

where $\boldsymbol{p} = (p_1, \ldots, p_k)$ with $p_1 + \ldots + p_k = 1$, and $\boldsymbol{\vartheta} = (\boldsymbol{\vartheta}_1, \ldots, \boldsymbol{\vartheta}_k)$, while $f_1, \ldots, f_k$ are properly defined probability density functions. The weight of group j is $p_j = P(Z = j)$ and, conditional on $Z = j$, the outcome distribution is f_j and is governed by a q-dimensional parameter vector $\boldsymbol{\vartheta}_j$, $j = 1 \ldots, k$. Usually $f_j = f$, meaning that all the components of the mixture are assumed to belong to the same parametric family and only differ for the parameter $\boldsymbol{\vartheta}_j$. Optionally, both $\boldsymbol{\vartheta}$ and $\boldsymbol{p}$ may be assumed to depend on an observed set of covariates $\boldsymbol{x}$. In most cases, a multinomial model (e.g. logit, or

probit) is fitted to the mixing proportions $\boldsymbol{p}$, while a GLM-like model is specified for $\boldsymbol{\vartheta}$. For example, if f is the normal distribution, we could set $\boldsymbol{\vartheta}_j = (\boldsymbol{x\beta}_j, \sigma_j)$.

The model can be fitted by optimizing the corresponding likelihood function with respect to $\boldsymbol{\vartheta}$ and $\boldsymbol{p}$, and extension to censored outcomes is straightforward. Usually the Expectation-Maximization (EM) algorithm (Dempster et al., 1977) is used. A good reference for finite mixture modeling is the book by McLachlan and Peel (2000).

In the present study, we assume *f* to be the *Weibull* distribution with shape a and scale *b, (a,b) > 0*. We allow *b* to depend on covariates, setting $b = exp(\boldsymbol{x\beta})$. We use a multinomial logistic model for the mixing proportions, letting

$$p_j \propto exp\{\mathbf{x}\gamma_j\}$$

with $\boldsymbol{\gamma}_1 = 0$ to make the coefficient uniquely identifiable. The model parameters are thus

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_1, \ldots \boldsymbol{\alpha}_k$$
$$\boldsymbol{\beta} = \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k$$
$$\boldsymbol{\gamma} = \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_k$$

We fit the model using the EM algorithm. We fit the multinomial logistic model using the nnet R package, and the Newton-Rapshon algorithm to optimize the Weibull likelihood. The inference is based on nonparametric bootstrap.

5. Empirical Analysis

We present now an application of our estimation methodology to unemployment duration. We produce thus a survival analysis using mixture distribution and we try to identify: i) the groups (components) within the population, ii) the determinants of the belonging to a given group and iii) the determinants of unemployment duration within each group.

We apply our methodology to German evidence, using data on employment, income and wealth for the year 2002 from the SOEP survey. The dependent variable is given by unemployment duration measured in months: to obtain a more homogeneous sample we only consider individuals that became unemployed during 2002. We track these individuals for 12 months after they became unemployed: in some cases, they are still unemployed at the end of this period and, therefore, they are considered censored. Among the independent variables, we focus on the role of unemployment income support and of household wealth. In particular, we introduce unemployment income support as a dummy variable, that is, a variable that is one if the individual received any amount of money in the form of unemployment benefits or unemployment assistance: an accurate description of the mechanism governing UI in Germany is contained in Appendix A. Wealth is measured as three distinct variables computed at the household level. The first one is a dummy variable which is one if the dwelling is owned by a member of the household. The second is a dummy variable which is one if there is a mortgage on the dwelling. The third

variable represents net wealth and is given by the sum of financial asset value plus real estate value (apart dwelling) less total debt (apart mortgage on dwelling). Basically we consider all source of wealth apart the value of dwelling and mortgage on dwelling, that are treated separately as dummy variables. Finally, some socio-demographic variables are used: age, age squared, gender, education (expressed as having or not a college degree) and number of children within the family[3]. To capture unobserved heterogeneity we also add a variable measuring the share of past working life spent in unemployment.

We use a finite mixture distribution approach and we assume the existence of *k* components (groups) to which individuals belong on the base of the above variables and according to a multinomial logit model that acts as a selection model. Moreover, unemployment duration is determined according to survival models to which the above variables contribute: the specific model is a parametric cox model with Weibull distribution. While the described variables enter the Cox model of all groups, their effect can be different in different groups and the effect of time (the baseline hazard) may differ as well. Within the parametric cox model with Weibul distribution the baseline hazard is governed by the ancillary parameter: when the value of this parameter is equal to zero the baseline hazard is flat (time has no effect on re-employment probabilities), for values greater (lower) than zero the baseline is increasing (decreasing).

The estimation technique we use is computationally very intensive: to partly simplify the degree of complexity of computation we normalized all variables imposing mean zero and standard deviation equals to 1.[4]

To select the appropriate number *k* of components, we estimate the model assuming different numbers of components and we compute the value of the Aike Information Criterion (AIC) for each case: such value is given by twice the number of parameters less twice the log-likelihood of the estimation. According to this method, the best model to be estimated is the one with the lowest AIC value.

Results for the AIC are contained in Table 1 and they shows that the best model is the one with *k=3*.

[TABLE 1]

Following this indication, we proceed in estimating the model allowing for the presence of three components. In Table 2 we report the results for the selection model for the for groups: group 1 is considered the baseline group and the value of the coefficients describe whether that given variable increases or decreases the probability of belonging to group 2 or to group 3. The standard errors of the coefficients and the related z-statistics can be found in Appendix B. It must be

[3] We also tried to include gender and marriage status however these two variables turned to be not significant and we omitted them from the regressions.

[4] The only exception to this is the variable age that was normalized to the interval (1,2): this was done to avoid problems when using the square of age in the regression.

stressed that the assignment to given group is of probabilistic nature and does not, on the contrary, implies a strict or deterministic membership.

[TABLE 2]

The results contained in the above table shows that group 1 is the most common, with about half of the sample belonging to it. Group 2 is still quite large (slightly more than one third of the sample), group 3 is small but not irrelevant. Some characteristics affect the probability to belong to a given groups. In particular, group 3 seems to be quite well defined by the number of children in the household, which decreases the probability to be in group 3, and by the college degree which increases it. Also group 1 and group 2 are quite neatly different: selection to group 2 is more likely for individuals that are not married and from less wealthy households.

The selection model suggest interesting insights. First, individuals that have children and do not belong to the highly educated class are more likely to belong to group 1 and group 2. From an economical point of view, we can imagine that individuals in these groups are probably more in need of work (having children) and have less job opportunities (given the lower education): then we should expect them to search more intensively and to be less picky in accepting jobs.

We turn now to better explore the determinants of unemployment duration for these three groups. The results, for each group, are presented in Table 3 below: a positive (negative) coefficient implies an increase (decrease) in unemployment duration. The standard errors of the coefficients and the related z-statistics can be found in Appendix B.

[TABLE 3]

We start discussing the results for group 1. First of all, the ancillary parameter is significantly greater than one (i.e. the log of the parameter is greater than zero) but not particularly large: this implies that baseline hazard is very slightly increasing, suggesting that re-employment probability increase only slightly with time. Probably, individuals from this group start searching immediately and with quite ample effort. This confirms the idea that individuals in this group are steadly searching for and accepting jobs. This aspect is also confirmed by the fact that benefits have non-significant effect on duration for these individuals. Moreover, being married provides them more financial security and increases duration while the contrary is true if they have a mortgage.

Group 2 has an ancillary parameter statistically larger than one (the log is positive) and its magnitude is quite large. In other words, time put quite a lot of pressure on them and they increase their effort or reduce their reservation wage. As a matter of fact, financial conditions of this individuals are less secure than those in group 1 and they react searching more with time. It appears that these individuals want a job at all costs and receiving benefits does not alter their behavior. College degree, on the contrary, probably opens up more opportunities, reducing unemployment duration. The only thing that is slightly out of picture is the positive effect of owning the dwelling: of all things this is the all aspect that reduce pressure and probably exerts a negative effect on their search effort.

Group 3 has an ancillary parameter that is statistically larger than one (the log is positive) and greatly so. Probably, individuals from this group start searching immediately; however, as financial stress increases through time, they apply even more effort. This kind of individuals are in need of a job but want also to take their time to find one and this is confirmed by the fact that, for them, benefits have a significant positive effect on duration: the presence of an unemployment income relieves some of the financial pressure and refrain them to rush in accepting jobs. Even in this group the middle aged have the highest re-employment probability. This could suggest that individuals take their time in the search process, being picky and reluctant during the initial spell of unemployment: only at later stages they appear more active and willing to take into consideration any offer they receive. We could explain this with the fact that individuals with fewer or no children feel less urge to search or to accept any offer they receive. In addition, people with more education are more likely to belong to this group and, therefore, they may expect more from a job and take their time when choosing one. Moreover, unemployment benefits reduce unemployment duration. This is a very important result and it highlights how the search requirement is effective for the individuals of this group: as a matter of fact, it seems that these individuals are not putting all the effort that they could in the search process and thus imposing a search requirement is largely effective. In this respect, it appears that the moral hazard behavior is not present or, at any rate, it is off-set by the search requirement.

Finally, it appears that in all the groups, individuals that are in their middle age have higher re-employment probability; this is in line with the expectation that younger and older individuals are those facing more problems in finding a job.

A specification was also tried where we included some “ethical” variables in the selection model. In particular we included three dummy variables: the first describes whether individuals were very concerned about immigration to Germany, the second describes if individuals are strongly environmental conscious and the third describes whether they belongs to some religion. While the results we obtain for the determinants of unemployment duration within the three groups remains the same, the estimation of the selection model do not deliver significant results and in particular, the three ethical variables fail to explain the process of selection to the groups.

6. Conclusions

Our analysis was aimed at the identification of different patterns and clusters in the way in which several individuals’ characteristics affect unemployment duration. We develop a methodology which estimated duration with mixed distribution Cox regression. Particular emphasis was given to the role of unemployment benefits, of household wealth, of financial pressure and of the interaction of these aspects. As a matter of fact, both benefits and financial solidity should mitigate the financial pressure on the unemployed workers and, therefore, they could reduce the job search effort and the probability that a job offer is accepted: as a consequence, these components could increase unemployment duration. Our analysis examines three different groups of unemployed individuals that directly emerge from the estimation. The group 1 and group 2 are

the most numerous and represents individuals that appear to be steadily searching for a job. They are probably individuals that, having children, feel more pressure on finding a job. Moreover individuals that have lower education are more likely to belong to this group furthering implying less opportunities and more pressure. Unemployment benefits bestowed to them do not increase duration as they appear to be strongly focusing on finding a job. The difference between the two groups is that the second group appears to be made up by individuals with less financial solidity and, as a matter of fact, time has a stronger, positive, effect on re-employment on group 2 than on group 1. The third group is instead made up of individuals that clearly take their time in the search process. Time has a strong, positive, effect on their re-employment probabilities, indicating that at the beginning of the spell they are searching less actively or disregarding more job opportunities. For this group, benefits reduce unemployment duration: the search requirements seem to work well and force them to increase their search effort.

Our analysis shows that different clusters exist and that they strongly differ in terms of the baseline hazard and of the role that explanatory variables have on unemployment duration. This strongly suggests that our approach of using mixture distribution in duration models helps the understanding the mechanics and the determinant of unemployment duration.

From a policy point of view, our analysis clearly highlights how different groups of individual react differently to unemployment benefits and would suggest a deep tailoring of the eligibility and requirement. Clearly, the actual design process is extremely complex; however, the selection model we use could help in identifying the different categories of unemployed workers and could be useful in an actual tailoring of the UI schemes.

References

Bloemen, H. (2002), “The relation between wealth and labour market transitions: an empirical study for the Netherlands”, *Journal of Applied Econometrics,* Vol. 17 No. 3, pp. 249-268.

Bloemen, H. and Stancanelli, E. G. F (2001), “Individual Wealth, Reservation Wages, and Transitions into Employment”, *Journal of Labor Economics*, Vol. 19 No. 2, pp. 400-439.

Bloemen, H. and Stancanelli, E. G. F. (2005), “Financial Wealth, Consumption Smoothing and Income Shocks Arising from Job Loss”, *Economica*, Vol. 72 No. 3, pp. 431-452.

Browning, M. and Crossley, T. F. (2001a), "The Life-Cycle Model of Consumption and Saving", *Journal of Economic Perspectives*, Vol. 15 No. 3, pp. 3-22.

Browning, M. and Crossley, T. F. (2001b), "Unemployment insurance benefit levels and consumption changes", *Journal of Public Economics*, Elsevier, Vol. 80 No.1, pp 1-23.

Browning, M. and Crossley, T. F. (2008), "The long-run cost of job loss as measured by consumption changes", *Journal of Econometrics*, Vol. 145 No. 1-2, pp. 109-120.

Card, D., Chetty, R. and Weber, A. (2007), “Cash-On-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market”, *The Quarterly Journal of Economics*, Vol. 122 No. 4, pp. 1511-1560.

Cahuc, P. and Zylberberg, A. (2004), *Labor Economics*, MIT Press, Cambridge MA.

Chetty R. (2005), “Why do Unemployment Benefits Raise Unemployment Durations? Moral Hazard vs. Liquidity”, NBER Working Papers 11760, National Bureau of Economic Research, Inc.

Chetty R. (2008), “Moral Hazard vs. Liquidity and Optimal Unemployment Insurance”, *Journal of Political Economy*, Vol. 116 No. 2, pp. 173-234.

Corsini, L. (2012), “Unemployment Insurance Schemes, Liquidity Constraints and Re-employment: A Three Country Comparison", *Comparative Economic Studies*, Vol. 54 No. 2, pp. 321-340.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society. Series B (Methodological), Vol. 39, No. 1, pp. 1-38.

Eurostat (2007), “Description of SILC user Database Variables: Cross-Sectional and Longitudinal. Version 2007.1 from 01-03-09”, Eurostat, Unit F-3, Luxembourg.

Stephen P. Jenkins (1998), Discrete time proportional hazards regression, *Stata Technical Bulletin*, Vol. 7, No. 39, pp. 22-32.

McDonald, J. B. and Butler, R. J., 1987. Some Generalized Mixture Distributions with an Application to Unemployment Duration, *The Review of Economics and Statistics*, 69 (2), pp. 232-240.

McLachlan, G. and Peel, D. (2000). Finite Mixture Models. *Wiley Series in Probability and Statistics*, New York.

Pellizzari, M. (2006), “Unemployment duration and the interactions between unemployment insurance and social assistance”, *Labour Economics*, Vol. 13 No.6, pp. 773-798.

Petrongolo, B. (2001), “Reemployment Probabilities and Returns to Matching”, *Journal of Labor Economics*, Vol. 19 No. 3, pp. 716-41.

Rogerson, R., Shimer, R. and Wright, R. (2005), “Search-Theoretic Models of the Labor Market: A Survey”, *Journal of Economic Literature*, Vol. 43, pp. 959-988.

Appendix A. Income support schemes in Germany

The German system for income support (in the year 2002) is made of three tiers of support: the unemployment insurance (UI), the unemployment assistance (UA) and the social assistance (SA).

The UI tier entitles unemployed workers to receive unemployment benefits that amount to 60% of their previous net earnings. Workers with at least one dependent child receive 67% of their previous earnings. In any case, the maximum benefit is set at 60% or 67% of 4700 Euros of monthly gross earnings. Benefits from unemployment insurance are tax exempt.

Eligibility criteria dictate that workers should: a) be younger than 65, b) be a registered unemployed that is looking and available for work, c) have paid social security contributions for at least 12 months during the last three years. Contributions (and insurance) are compulsory.

The maximum duration of these benefits is equal to half of the contribution periods (measured in months) with a ceiling of 12 months for individuals up to 45 years old. For individuals older than this threshold, the maximum duration is still given by half of the contribution periods but the ceiling is extended depending on the actual age, reaching 32 months for individuals older than 57.

All UI claimants have to register with the Public Employment Service (PES), search for jobs and be available for participation in labour market programs. The claimants have to report periodically on their job-search activities: in particular, they have to report once every 3 months. Failure to comply with benefit conditions can lead to suspension of benefit receipt for a 12 weeks period and after two offences benefits cease completely.

The UA scheme is reserved to workers that have exhausted the UI. The eligibility criteria remain the same as the UI but the requirements in terms of past contributions are not relevant anymore and means tested criteria are added. In particular, UA is granted only to individuals whose households' asset is below the following threshold: 520 Euros per person in the household multiplied by his/her age up to 33800 per person; the value of main dwelling is not computed within this threshold and assets in the form of private pensions are treated differently. Workers that are eligible for the UA benefits receive 50% of their previous net earnings (57% if they have at least one dependent child) but the actual amount is reduced depending on the presence of other sources of income within the household. Duration is indefinite but must be renewed each year.

All UA claimants have to register with the PES and have exactly the same search requirements as the UI claimants.

Finally, all individuals in need can claim the SA. There is no requisite in terms of past contributions or employment so that this scheme acts independently from the entitlement to UI or UA and as long as the individuals are in need. In this case the SA is granted only if total household assets are below a very strict threshold: 1278 Euro for the claimant plus 528 for his/her spouse and plus 276 for each child; even in this case, the value of the main dwelling is disregarded.

The actual amount paid by SA is given by the difference between a basic amount computed at the household level and the total income of the household. The basic amount is computed starting from the household composition[5] and adding housing and heating allowances, medical insurance

[5] For the first adult within the household the amount given is 292 Euro (282 in former East Germany). Each extra adult adds another 80% of that amount and each child (below 18) adds between 50% and 90% depending on the child age.

and some specific extra allowances (for lone parents, disabled people, pregnant women, school fees and so on).

Recipients of SA have to register at PES, and if they are capable to work, they have to be looking for employment, be available for work and to participate in activation programs, however, formal requirements, reporting activity and proofs are not required.

Appendix B. Standard Errors and Z-statistics

[TABLE A1]

[TABLE A2]

Table 1
Aike Information Criterion (AIC)

Number of Components (*k*)	Number of Parameters	AIC
1	10	2474.67
2	29	2380.34
3	48	2365.54
4	67	2375.28

Table 2
Selection Model for Groups

	Group 1 (baseline)	Group 2	Group 3
Age	-	27.273	32.322
Age Squared	-	-8.853	-12.180
Unemployment Benefits	-	1.030	-0.030
College Degree	-	1.316	**2.434 *****
Female	-	0.442	-0.728
Married	-	**-1.856 *****	-0.509
Number of Children	-	-0.024	**-1.112 ****
Ownership of Dwelling	-	-1.472	0.187
Mortgage on Dwelling	-	0.684	0.976
Net Wealth	-	**-1.926 ****	-0.452
Constant	-	-21.382	-23.615
Share of individuals	53.37%	38.74%	7.89%

Significance Levels: *1%, **5%. ***10%

Table 3
Cox Regression Estimation for Groups

	Group 1	Group 2	Group 3
Age	**-14.94 *****	**-9.45 *****	**-4.92 *****
Age Squared	**5.38 *****	**3.66 *****	**1.88 *****
Unemployment Benefits	0.10	-0.97	**-0.69 *****
College Degree	-0.51	**-3.65 ***	0.02
Female	-0.03	0.19	-0.01
Married	**0.64 ****	-0.08	-0.01
Number of Children	-0.04	0.13	0.04
Ownership of Dwelling	0.05	**2.57 *****	-0.05
Mortgage on Dwelling	**-0.46 ****	-0.57	-0.04
Net Wealth	0.33	1.42	-0.01
Constant	**12.03 *****	15.33	**3.91 *****
Log of ancillary parameter	**0.39 *****	**1.84 ****	**3.83 *****

Significance Levels: *1%, **5%. ***10%

Table A.1
Standard Errors and Z-statistics of Selection Model for Groups

	Group 1 (baseline)	Group 2	Group 3
Standard Errors			
Age	-	18.26	26.14
Age Squared	-	6.33	8.99
Unemployment Benefits	-	1.11	0.78
College Degree	-	1.16	0.83
Female	-	0.64	0.60
Married	-	0.68	0.73
Number of Children	-	0.33	0.54
Ownership of Dwelling	-	5.66	3.04
Mortgage on Dwelling	-	0.77	0.78
Net Wealth	-	0.97	0.93
Constant	-	13.65	11.10
Z- Statistics			
Age	-	1.49	1.24
Age Squared	-	-1.40	-1.36
Unemployment Benefits	-	0.93	-0.04
College Degree	-	1.14	2.94
Female	-	0.70	-1.22
Married	-	-2.72	-0.70
Number of Children	-	-0.07	-2.06
Ownership of Dwelling	-	-0.26	0.06
Mortgage on Dwelling	-	0.89	2.55
Net Wealth	-	-1.99	-1.56
Constant	-	-1.56	-1.25

Table A.2
Standard Errors and Z-statistics of Cox Regression Estimation for Groups

	Group 1	Group 2	Group 3
Standard Errors			
Age	4.66	2.67	0.86
Age Squared	1.63	1.01	0.27
Unemployment Benefits	0.22	2.47	0.02
College Degree	0.50	2.08	0.03
Female	0.21	0.56	0.02
Married	0.29	0.57	0.03
Number of Children	0.14	0.16	0.03
Ownership of Dwelling	0.32	0.75	0.06
Mortgage on Dwelling	0.20	0.54	0.04
Net Wealth	0.20	1.04	0.12
Constant	3.30	13.81	0.68
Log of ancillary parameter	0.11	0.76	0.53
Z- Statistics			
Age	-3.20	-3.54	-5.69
Age Squared	3.30	3.64	6.84
Unemployment Benefits	0.48	-1.20	-30.79
College Degree	-1.02	-1.75	0.82
Female	-0.12	0.35	-0.56
Married	2.24	-0.14	-0.38
Number of Children	-0.30	0.78	1.07
Ownership of Dwelling	0.17	3.41	-0.88
Mortgage on Dwelling	-2.28	-1.06	-1.01
Net Wealth	1.67	1.37	-0.10
Constant	3.65	1.11	5.73
Log of ancillary parameter	3.69	2.41	7.27