Duration Models and Differential Evolution in the Analysis of Large Data Sets

by

Michele Lalla¹
Sandra Paterlini²

April 2005

¹ e-mail: michele.lalla@unimore.it
² e-mail: sandra.paterlini@unimore.it
Abstract

Two measures of unemployment duration were obtained for each individual from data collected by Istat through the Quarterly Labour Force Survey. One measure was the length of on-going spells of unemployment, declared by the interviewed individual in a first wave and suitably adjusted considering information in a second wave (Spell Answers). The other was the time elapsed between the dates of the end of the previous job and the beginning of the current job in the second wave (Spell Reconstructions). They were modelled by traditional duration models via a mixture distribution of the answered and reconstructed spells, instead of using their minimum. The likelihood of the mixture is highly complex with a large number of covariates. Hence, classical optimisation methods do not provide robust parameter estimations. A stochastic search algorithm, called Differential Evolution, is proposed to tackle the numerical optimisation problem. In fact, even if DE is still rather unknown, it proves to be consistently and clearly superior to other search heuristic and classical optimisation approaches in many applications. Moreover, it is very simple to implement and requires little or no parameter tuning. A comparison between the ordinary duration models based on a sole unemployment spell for individuals (spell answers or spell reconstructions) is also reported.

JEL classification: C41, J64
Keywords: Unemployment duration, unobserved heterogeneity, heaping effect, Differential evolution, proportional hazard model

1. Introduction

The average length of time that workers spend seeking a job is an important indication of economic welfare, but, despite this role, its measurement still appears elusive, as is the question: How long do spells of unemployment last? One reason is almost an ontological issue, as it lies in the addition of the fourth

* Corresponding author.
E-mail address: lalla.michele@unimore.it
dimension to the sampling strategy: survey populations constantly change over time in the composition and the characteristics of their members, giving rise to many problems and different objectives for analysis (Duncan and Kalton, 1987). There are a variety of sample designs for collecting data at several points of time. Each of them varies in the amount and kind of overlaps between periods, i.e., the percentages of the same statistical units to retain in the sample for each wave, or round, of data collection and the number of waves. The Italian Quarterly Labour Force Survey (QLFS) is conducted by Istat and uses a rotation scheme 2-2-2 for the sampling units, which are families. Each rotation group remains in the survey for two consecutive quarters, drops out for two subsequent quarters, and then returns for two more consecutive quarters, as described in Table 1. Each survey consists of a point-in-time data collection, as is the Current Population Survey (CPS), and it cannot measure important aspects of phenomena enduring over time (Kiefer et al., 1985). However, the rotation scheme acts firstly to maintain an up-to-date sample of a changing population, secondly to reduce panel conditioning and panel loss with respect to the nonrotating panel survey, and thirdly to detect the transitions between the labour market states of the individuals interviewed in two different waves, \( t_1 \) and \( t_2 \). For example, in Table 1, group C was interviewed in 1998:1 and 1998:2, or group E in 1997:4 and 1998:3 (Torelli and Trivellato, 1988).

Table 1
Rotating panel design of the Italian Quarterly Labour Force Survey

<table>
<thead>
<tr>
<th>Year</th>
<th>Q</th>
<th>Month</th>
<th>Rotation groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>4</td>
<td>October</td>
<td>A B E F</td>
</tr>
<tr>
<td>1998</td>
<td>1</td>
<td>January</td>
<td>B C F G</td>
</tr>
<tr>
<td>1998</td>
<td>2</td>
<td>April</td>
<td>C D G H</td>
</tr>
<tr>
<td>1998</td>
<td>3</td>
<td>July</td>
<td>D E H I</td>
</tr>
<tr>
<td>1998</td>
<td>4</td>
<td>October</td>
<td>E F I L</td>
</tr>
<tr>
<td>1999</td>
<td>1</td>
<td>January</td>
<td>F G L M</td>
</tr>
</tbody>
</table>

The ordinary transitions in the labour market considered referred to three states (professional conditions) —Employed (E), Unemployed (U), Out-of-the-Labour-Force (OLF)— and the standard definition of unemployment duration is given by the minimum between two spells: (i) job search duration, hereafter referred to as spell answer (SA) and denoted by the random variable \( T_A \), and (ii) length of the period since the last job held, hereafter referred to as spell reconstruction (SR), and denoted by the random variable \( T_R \) (EUROSTAT,
The job-seeker is subject to some specific conditions: (a) at least fifteen years old, (b) without work during the reference week, (c) available to start work or to accept a job offer before the end of the next two weeks, (d) «actively» searching for an occupation. He/she begins the action at a certain point in time, \( t_{B,i} \), and continues until the end, at another point in time, \( t_{E,i} \), because he/she finds a job or he/she reduces the «intensity» of the search, thus leaving the labour force. Therefore, generally speaking, the unemployment spell for the \( i \)-th individual would be the random variable \( T_i = t_{E,i} - t_{B,i} \). However, it not convenient to detect the intensity of the search over time and condition (d) is collected at the date of survey, ascertaining that the interviewee carried out at least one job search attempt in the previous four weeks, with respect to the reference week. The situation is illustrated in Fig. 1, where calendar time is the abscissa, the lower bound of a horizontal segment indicates the time when the spell started, the upper bound shows when it concluded. The length of each segment is an observation of a random variable, termed completed spell of unemployment. However, almost all spells will be in progress when each survey is conducted, i.e., the durations collected by the Italian QLFS are incomplete spells of unemployment, as in an ordinary CPS. In Fig. 1, case \( c_1 \) represents the standard situation, where all data about employment duration are available. The other cases illustrate only some situations with missing data or nonsampling variable errors, such as processing and field errors in coding and tabulating components. For example, \( c_2 \) denotes a direct transition into the employment state, which is generally not collectable, \( c_3 \) refers to the unavailable date of the beginning of an unemployment spell indicated by a question mark (but limited by the time interval elapsing between two consecutive waves at \( t_1 \) and \( t_2 \)), \( c_4 \) concerns a right-censored spell, and so on.

---

1 A new definition of unemployment (EC Commission Regulation No. 1897/2000) may cause a break between 2000 and 2001, but it is fully compatible with ILO standards and increases comparability in the EU, not yet fully applied in Italy (EUROSTAT, 2002).
Fig. 1 – Different types of spells collectable between two generic waves conducted at $t_1$ and $t_2$, respectively.

The analysis of duration data through the hazard function has become a standard technique in the last two decades and it is currently used in many scientific and social fields, although the terminology changes from one discipline to another, thus known as reliability theory, survival analysis, and duration data (Kalbfleisch and Prentice, 1980; Cox and Oakes, 1984; Lancaster, 1990). Analyses of the labour market and turnover through the life table method date back to Silcock (1954), but this method received more attention after the applications of Nickell (1979a,b) and Lancaster (1979) and has been extensively developed in recent decades. However, model building and parameter estimation can be difficult tasks in the analysis of labour market transitions, when the distribution function of duration is not easy to handle, all spell lengths are censored, and the number of explanatory variables (covariates) is extremely high.

In the absence of these problems, unemployment spells can be handled through highly standardized methods encompassing almost routine procedures (Kleinbaum, 1996; Lee, 1993), but the data modelled below present two specific complications, namely a complex distribution function and a large set of covariates.

The aim of the paper is to analyse unemployment spells using traditional duration models, attempting to use both of these two unemployment spells simultaneously, combined in a mixture of distributions. Generally, researchers should choose only one measure, which could be one of them, or their minimum. To account for unobserved heterogeneity and heaping effect, the likelihood of the mixture model becomes highly complex with a large number of explanatory variables. Hence, a robust estimation of the parameters in duration models is
difficult to achieve using classical optimisation methods. We propose to use a
stochastic search algorithm, called Differential Evolution (DE), to tackle the
numerical optimisation problem. In fact, even if DE is still rather unknown, it
proves to be consistently and clearly superior to other search heuristic and
classical optimisation approaches in many applications (Storn and Price, 1995).
Moreover, DE is very simple to implement and requires little or no parameter
tuning. Comparison with duration models on the single measures, SA and SR, is
also reported in the following.

For the sake of brevity, the transitions considered in the following refer to the
movement of individuals from the unemployment state to the employment state.
The paper proceeds as follows. Section 2 illustrates the available data and the
procedure used to determine the two measures of unemployment duration.
Section 3 outlines Cox’s Proportional Hazards Model (CPHM) with a Weibull
specification of the baseline hazard, the inclusion of the unobserved
heterogeneity, the heaping effect, and the likelihood functions of the models.
Section 4 contains a brief description of the DE stochastic search algorithm.
Section 5 describes the covariates introduced in the model and the estimation of
parameters. Section 6 concludes the paper with some comments and remarks.

2. Sampling data

The unemployment durations modelled below were obtained from six waves
of the Italian QLFS (1997:4–1999:1), as illustrated in Table 1, i.e., six databases
were available, starting from the fourth quarter of 1997 to the first quarter of
1999. They cover a period of 18 months and each file contains more than
200,000 individuals. The choice of the investigation period was arbitrary, but
such that it was long enough to determine the complete unemployment spells and
short enough to avoid nonstationary behaviour and the possible presence of
economic shocks.

For each job seeker, the questionnaire used by the QLFS provides two
elements from which it is possible to evaluate unemployment spell duration. One
concerns the answer to the Question: How many months have you been looking
for a new job? (Q46). The spell answer ranges from 1 to 99 months and it is the
length of the on-going (incomplete) spell of unemployment, as the respondent is
currently unemployed, \( u_{BU} \). Therefore, the completed spell of unemployment
should be determined considering the transition that took place between two
surveys (waves) and the length of current job tenure, \( u_{CJT} \), obtained in the
second wave by the answer to the question: When did you begin your current
job? (Q24, which requires specification of the month and year, \( d_{BCJ} \)). Hence, the
completed spells for transitions from unemployment to employment is given by:
\[ T_A = u_{ISU} + 1_{CJT} (\Delta t - u_{CJT}) + (1 - 1_{CJT}) \Delta t / 2, \]

where \( T_A \) denotes the so-called spell answer (SA), \( \Delta t \) is the time interval between the two linked surveys detecting transition occurrence (see Fig. 1), \( 1_{CJT} \) is an indicator variable equal to 1 when \( u_{CJT} \) is not missing and less than \( \Delta t \). Otherwise, \( 1_{CJT} \) is equal to 0. When \( u_{CJT} \) is greater than \( \Delta t \), presumably there is an error in the data.

The other evaluation of unemployment duration considers the time elapsed between the dates of the end of the previous job and the beginning of the current job. The former is generally obtained in the first wave by the answer to the question: When did your last job end? (Q38, which requires specification of the month and year, \( d_{FLJ} \)). The latter is usually given by \( d_{BCJ} \), defined previously. Therefore, for individuals with changes in the unemployment state between two waves, it is possible to determine the time elapsed between two jobs which could be an approximate measure of an unemployment spell: \( T_R = g(d_{BCJ}, d_{FLJ}) \), where \( T_R \) denotes the so-called spell reconstruction (SR) and \( g(\cdot) \) is a function computing the length (in months) of a period delimited by two dates.

A possible impreciseness of \( T_R \) lies in the practical process of job searching. In fact, job losers/leavers could remain out-of-the-labour-force because they are not looking for work after losing/leaving job. Therefore, in abstract terms, \( T_R \) should be greater than or equal to \( T_A \), but the two types of unemployment spells are not always available (missing data) and \( T_R \) is not always greater than or equal to \( T_A \). However, \( T_R \) may also assume values that are too high — \( t_{R,i} \) was greater than 120 months for 457 out of 15123 individuals with \( T_R \) being available (3.5%), where the linked and selected units amounted to 20617—, while \( T_A \) ranged between \( 1/4 \) to 108 months involving 13309 individuals. For the 457 individuals with \( t_{R,i} > 120 \) months, \( t_{A,i} \) was never missing and only 26.9% of them answered «99 months» to Q46. However, the latter were eliminated from the sample and the number of cases became 20160. For this reduced data set, the distribution of differences between the two durations, \( d_i = t_{A,i} - t_{R,i} \), is reported in Table 2: it showed a positive mean (\( \bar{d} = 4.1 \) months), skewed to the right (\( \gamma_1 = 0.368 \)), and also proved to have a pronounced leptokurtosis (\( \gamma_2 = 3.792 \)). The latter apparently implies the contrary of what was expected, \( i.e., t_{A,i} \) is more frequently greater than \( t_{R,i} \). However, with the inclusion of the 457 individuals eliminated, the conclusions overturned: \( \bar{d} = -4.7 \) months, \( \gamma_1 = -3.368 \), and \( \gamma_2 = 22.220 \), as expected.
There are nine possible transitions between the three professional conditions (E, U, OLF), but some of them generate spell durations that are incomplete and/or over-represent short lengths, such as transitions from employment to employment or from OLF to OLF, while others are not of direct interest here, such as those from employment to OLF for retirement. The selected transitions are reported in Table 3, together with distinctions as to the characteristics of $u_{C,\Gamma}$ used to determine $t_{A,i}$.

Table 2
Absolute frequency and percentage distribution of the differences between the spell answers and spell reconstructions spells ($d_i = t_{A,i} - t_{R,i}$) with the lower (LB) and upper bounds (UB) of the class intervals (the number of missing values is 12,712)

<table>
<thead>
<tr>
<th>LB</th>
<th>-∞</th>
<th>-10.5</th>
<th>-5.5</th>
<th>-2.5</th>
<th>-1.5</th>
<th>-0.5</th>
<th>0.5</th>
<th>1.5</th>
<th>2.5</th>
<th>5.5</th>
<th>10.5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB</td>
<td>-10.5</td>
<td>-5.5</td>
<td>-2.5</td>
<td>-1.5</td>
<td>-0.5</td>
<td>0.5</td>
<td>1.5</td>
<td>2.5</td>
<td>5.5</td>
<td>10.5</td>
<td>+∞</td>
<td></td>
</tr>
<tr>
<td>$n_i$</td>
<td>1019</td>
<td>263</td>
<td>288</td>
<td>164</td>
<td>672</td>
<td>1638</td>
<td>549</td>
<td>230</td>
<td>439</td>
<td>440</td>
<td>1746</td>
<td>7,448</td>
</tr>
<tr>
<td>%</td>
<td>13.7</td>
<td>3.5</td>
<td>3.9</td>
<td>2.2</td>
<td>9.0</td>
<td>22.0</td>
<td>7.4</td>
<td>3.1</td>
<td>5.9</td>
<td>5.9</td>
<td>23.4</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 3
Number of individuals and column percentages for the transitions detected between two waves

<table>
<thead>
<tr>
<th>Transitions</th>
<th>Conditions</th>
<th>Employed</th>
<th>Unemployed</th>
<th>Out-of-the-Labour-Force</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed</td>
<td>$u_{C,\Gamma} &lt; \Delta t$</td>
<td>2221 17.1</td>
<td>2221</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$u_{C,\Gamma} &lt; \Delta t$</td>
<td>1774 13.6</td>
<td>1774</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$u_{C,\Gamma} \geq \Delta t$</td>
<td>1190 9.2</td>
<td>1190</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>$u_{C,\Gamma}$ missing</td>
<td>677 5.2</td>
<td>677</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\Delta t &gt; u_{ISU-2}$</td>
<td>551 100</td>
<td>551</td>
<td></td>
<td>6606</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>13003 100%</td>
<td>551 100%</td>
<td>6606 100%</td>
<td>20160</td>
</tr>
</tbody>
</table>

Legend: $u_{ISU-2}$ denotes an incomplete spell of unemployment in the second wave.
3. Models for unemployment duration

Let \( T_i \) be a random variable denoting the completed unemployment spell of the \( i \)-th individual and \( t_i \), its observed value in months or weeks or days. Let \( \mathbf{x}_i(t) \) be a (row) vector of time-dependent and independent explanatory variables (covariates) representing the individual circumstances of the \( i \)-th job seeker such as, personnel characteristics (gender, age, education), family information (size, number of children, number of employed members) or labour market status (unemployment rate, benefits). A job seeker’s conditional probability of leaving unemployment in a short interval \( (t, t + dt) \) is given by the product of the two elements: the probability of receiving a job offer, \( q[t, \mathbf{x}_i(t)] \), and the probability of accepting it, which implies that the wage offered exceeds the reservation wage, \( \xi[t, \mathbf{x}_i(t)] \) (Lancaster, 1990; Narendranathan and Stewart, 1993). Therefore, the conditional probability of leaving unemployment within the short interval \( (t, t + dt) \) is the hazard

\[
h(t, \mathbf{x}_i(t)) \, dt = q[t, \mathbf{x}_i(t)] \left[ \int_{\frac{\xi[t, \mathbf{x}_i(t)]}{f(w)}} f(w) \, dw \right] = q[t, \mathbf{x}_i(t)] a\{\xi[t, \mathbf{x}_i(t)]\},
\]

where \( f(w) \) is the distribution of wages relative to the vacancies in the labour market. The probability \( q[t, \mathbf{x}_i(t)] \) expresses the job seeker’s possibilities of finding a vacancy and being offered the job, given that he or she is available. The acceptance probability, \( a\{\xi[t, \mathbf{x}_i(t)]\} \), is generally affected by individual factors concerning the stopping strategy and the parameters of the vacancy-wage distribution.

It is possible to model the hazard function alone, ignoring the underlying choice structure, although it loses some distinctive aspects of the job seekers’ behaviour. Different specifications of \( h[t, \mathbf{x}_i(t)] \) allow the model to deal with the problem of nonstationarity, as well as with other features that may arise in empirical contexts (Amemiya, 1985, pp. 433-455). The most frequently reduced form is expressed in a proportional hazards model (Cox and Oakes, 1984)

\[
h(t_i, x_i) = \frac{f(t_i, x_i)}{1 - F(t_i, x_i)} = h_0(t_i; \theta) \varphi(x_i; \beta),
\]

where \( f(t_i, x_i) \) and \( F(t_i, x_i) \) are the density and the distribution functions of \( T_i \), respectively, \( h_0(t_i, \theta) \) is the baseline hazard at time \( t_i \) for the \( i \)-th individual with \( x_i = 0 \). The function \( \varphi(x_i; \beta) \) denotes the impact of covariates \( x_i \) on the baseline hazard, and \( \beta \) is the vector of parameters associated with them. Various
specifications of \( h_0 (t_i, \theta) \) account for different distributions of \( T_i \) and therefore, the shapes of the baseline hazard such as the exponential, which is memoryless because it implies a constant hazard, the Weibull, the gamma, the log-logistic, and so on (Kiefer, 1988; Lancaster, 1979). The log-linear form for \( \varphi(x_i; \beta) = \exp(x_i \beta) \) is often used to express the impact of covariates (Cox, 1972) and the hazard function becomes \( h(t_i, x_i) = h_0 (t_i, \theta) \exp(x_i \beta) \).

Changes occurring over time in the characteristics of individuals and the effects of the initial attributes on the evolution of the professional positions/conditions, or on the transitions between labour force states, represent the main topics of unemployment duration analysis. One aspect concerns the attenuation or accentuation of the impact produced by the starting attributes throughout the subsequent experiences over time, i.e., state dependence (Heckman, 1981). Another aspect concerns the completeness of the selected individual characteristics and/or their persistence over time, i.e., some variations across the sampling units (or over time for the same sampling units) may not have been included in the model due to measurement errors or omission (heterogeneity). Although this aspect is intrinsic in all statistical models, it leads to peculiar effects in the analysis of unemployment spells (Blossfeld and Hamerle, 1992).

In a duration model, as formulated in equation (2), state dependence is expressed by the baseline hazard \( h_0 (t_i, \theta) \), while the heterogeneity is captured by a set of explanatory variables, \( x_i \). Unobserved heterogeneity induced by omitted variables may be introduced on the scale parameter of the distribution assumed for \( T_i \), considered as a continuous random variable (Heckman and Singer, 1984a,b; Heckman, 1991). However, identification of the distribution for the scale parameter is not easy and the choice is conditioned by computational difficulties.

For the available data, \( T_i \) was assumed to be Weibull-distributed. Therefore, \( h_0 (t_i, \theta) = \alpha (\lambda t_i)^{a-1} \) and the hazard is monotonic: increasing if \( \alpha > 1 \), decreasing if \( \alpha < 1 \), and constant if \( \alpha = 1 \). Thus, a value of \( \alpha \) greater (or lower) than 1 implies an increase (or decrease) in the conditional probability that an unemployed individual will find a job shortly, as the unemployment spell increases in length. A value of \( \alpha = 1 \) implies a memoryless job-seeking process, i.e., the hazard function is constant and there is no duration dependence. The density function of \( T_i \) will be

\[
f(t_i, x_i; \lambda, \alpha, \beta) = \alpha \lambda (\lambda t_i)^{a-1} \exp[x_i \beta - (\lambda t_i)^a \exp(x_i \beta)],
\]
but the assumption that $T_1$ is Weibull-distributed, $W(t; \lambda, \alpha)$, is data-dependent, i.e., it will be ascertained by examining the data. If heterogeneity is introduced on the scale parameter, $\lambda$, it becomes a continuous random variable, $\Lambda$. An interesting solution is obtained by assuming again that $\Lambda$ is Weibull-distributed, $W(\lambda; \vartheta, \alpha)$, with the scale parameter $\vartheta$ and the same shape parameter, $\alpha$, of the distribution of $T_i$ (Lancaster, 1990, pp. 65-70; Lalla and Pattarin, 2001). Let $\phi(\lambda; \vartheta, \alpha)$ be the density function of $\Lambda$. Then, the density of the distribution of completed spell lengths is given by

$$f(t_i; x_i; \lambda, \alpha, \vartheta, \beta) = \frac{\alpha t_i^{\alpha-1} \vartheta^a \exp(-x \beta)}{[t_i^a + \vartheta^a \exp(-x \beta)]^2}$$

and its corresponding hazard function (Lalla and Pattarin, 2001) becomes

$$h_i(t_i; x_i; \vartheta, \alpha, \beta) = \frac{\alpha t_i^{\alpha-1}}{t_i^a + \vartheta^a \exp(-x \beta)}.$$  

The inclusion of unobserved heterogeneity, according to the previous procedure, generates a distribution that is more flexible than the Weibull because it has a hazard with a non-monotonic trend: it is zero for $t_i$, increases to a maximum, then decreases and $h_i(t_i, x_i) \rightarrow 0$ for $t_i \rightarrow \infty$. The result seems to be satisfactory because the distribution has a flexible and attractive hazard shape, but it should be noted that the choice of the Weibull is arbitrary and the mean and variance are not defined for some values of $\alpha$.

3.1. Telescoping effect

The measurement of unemployment duration for the $i$-th individual, $t_i$, is strictly based on his/her capacity to evaluate exactly the time interval spent seeking job or to remember the dates determining that spell. However, memory fades and becomes imprecise over time, generating two kinds of errors. One is leaving out an episode: he/she forgets a short period of employment. Another is expansion/contraction of time (telescoping), where an episode is remembered as occurring later or more recently than it actually did. The telescoping effect for unemployment spells would generate pronounced spikes on their empirical distribution (heaping effect), with respect to the expected smoothness of the frequency polygon. In fact, the true unobserved spell length of the $i$-th individual, $t_i$, could have been approximated to a preferred value because it is easier to
communicate, \(t_{HIj}\), such as a quarter, semester, year and their multiples (Sudman and Bradburn, 1973; Sikkel, 1985), i.e., a sort of calendar effect.

The location of the spikes (heaped values) is a matter of empirical evidence. Let \(J\) be the number of heaped values and let \(h_1 < h_2 < \ldots < h_J\) be their corresponding observed abscissas (one for each spike), respectively, sorted in ascending order. Let \(y_i\) be a binary variable: \(y_i = 1\) with a probability of \(H(t_i; \gamma)\) for the observed spell length coinciding with a heaped value; otherwise \(y_i = 0\) with a probability of \([1 - H(t_i; \gamma)]\). Therefore, \(H(t_i; \gamma)\) is the heaping function with a vector of parameters \(\gamma\), denoting the probability of having detected a spell length, \(t_i\), arising from the telescoping effect. A possible model for the stochastic heaping processes becomes

\[
t_{HIj} = t_i + \delta_j y_i,
\]

where \(\delta_j = h_j - t_i\) is such that \(|h_j - t_i|\) represents the minimum for \(j = 1, 2, \ldots, J\) (Torelli and Trivellato, 1993). The exponential distribution, \(H(t_i; \gamma) = 1 - \exp(\gamma)\), is the simplest and most practical assumption, although it may depend on individual characteristics, \(x_i\), which were not considered here in order to avoid over-parameterisation.

The binary variable \(y_i\) was defined for each type of spell: \(y_{A;i}\) for \(T_{A;i}\) and \(y_{R;i}\) for \(T_{R;i}\). Therefore, two sets of heaped values will be obtained. Let only \(T_{A;i}\) be considered. The abscessas of heaped values, \(h_{A;1} < h_{A;2} < \ldots < h_{A;J}\), were determined from the frequency distribution of \(t_{A;i}\) (Fig. 2) and \(T_{R;i}\) (Fig. 3).

Now, it is possible to identify an interval for each \(t_{A;i} = h_{A;i}\), considering two adjacent points \(h_{A;i-1}\) and \(h_{A;i+1}\). The lower bound is \(t_{A;HLij} = (h_{A;i-1} + h_{A;i})/2\), and the upper bound is \(t_{A;HUij} = (h_{A;i} + h_{A;i+1})/2\). Moreover, in general, a spell may be still in progress. The binary variable \(c_i\) could represent censored spells: it is equal to 1 if the observed unemployment spell \(t_i\) is complete and 0 otherwise. Therefore, two dummies are needed for the two types of spells, respectively \(c_{A;i}\) for \(T_{A;i}\) and \(c_{R;i}\) for \(T_{R;i}\). For the \(i\)-th individual, the available data set is given by

\[
\{ t_{A;i}, c_{A;i}, y_{A;i}, t_{A;HLij}, t_{A;HUij}, c_{R;i}, y_{R;i}, t_{R;HLij}, t_{R;HUij}, x_i \},
\]

where \(t_{A;HLij}\) and \(t_{A;HUij}\) assume their corresponding values only when \(y_{A;i}\) is equal to 1 or 0 otherwise.
Fig. 2 – Percentage distributions of spell answers, $T_{R;i}$, for men and women (1997:4-1999:1).

Fig. 3 – Percentage distributions of spell reconstructions, $T_{R;i}$, for men and women (1997:4-1999:1).
The empirical percentage distribution of \( T_{A,i} \) shows many spikes, but their abscissas \( \{3, 5, 13, 14, 19, 20, 25, 26, \ldots \} \) are not immediately interpretable in ordinary calendar time (Fig. 2), as their end points were calculated with a certain precision. On the contrary, the distribution of \( u_{SU} \), collected through Q46 and not reported here, showed more marked and recognizable abscissas values as years or submultiples like quarters and semesters: \( \{3, 6, 12, 15, 18, 24, 36, 48, \ldots \} \).

The empirical percentage distribution of \( T_{R,i} \) in Fig. 3 shows few marked spikes, as the dates are not subject to the calendar effect on the part of respondents. Therefore, the distribution of \( T_{R,i} \) is smoother than that of \( T_{A,i} \), except in some points: \( \{4, 8, 14, \ldots \} \).

### 3.2. Mixture models and parameter estimations

Let \( T_i \) be the unobserved completed spell of the \( i \)-th individual, considered as a random variable. The values of \( T_i \) concern the observation of the random variables \( T_{A,i} \) and \( T_{R,i} \), i.e., spell answers and spell reconstructions, respectively. Therefore, the distribution of \( T_i \) could be expressed by a mixture distribution of \( T_{A,i} \) and \( T_{R,i} \):

\[
f(t_i, x_i; \theta, \beta) = \pi \psi_A(t_{A,i}, x_i; \theta, \beta) + (1 - \pi) \psi_R(t_{R,i}, x_i; \theta, \beta),
\]

where \( \pi \) is the weight of the linear combination of the two distribution functions, \( \psi_A(t_{A,i}, x_i; \theta, \beta) \) and \( \psi_R(t_{R,i}, x_i; \theta, \beta) \) expressed in equation (4), for \( T_{A,i} \) and \( T_{R,i} \), respectively. The general form of the functions in the linear combination would include the heaping effect and the right-censored spells, which are most frequent in duration data. As an example, for the random variable \( T_{A,i} \), the density function is

\[
\psi_A^*(t_i) = \left\{ \psi_A(t_{A,i}, x_i; \theta, \beta) \right\}^c \left[ S_A(t_{A,i}, x_i; \theta, \beta) \right]^{1-c} \left[ 1 - H_A(t_{A,i}; \gamma) \right]^{\gamma} \times \left\{ \psi_A(t_{A,i}, x_i; \theta, \beta) \right\}^c \left[ S_A(t_{A,i}, x_i; \theta, \beta) \right]^{1-c} \left[ 1 - H_A(t_{A,i}; \gamma) \right]^{\gamma} + \int_{[\text{trunc}]} \left\{ \psi_A(v, x_i; \theta, \beta) \right\}^c \left[ S_A(v, x_i; \theta, \beta) \right]^{1-c} \left[ H_A(v; \gamma) \right] dv \right\}^{1-\gamma},
\]
where $S_A(\cdot)$ is the corresponding survivor function, $H_A(\cdot)$ is the heaping function, $y_i$ is the binary variable denoting whether $t_i = h_i$, and $c_i$ is again a binary variable denoting whether the observed spell is a completed spell. An explanation of this expression can be found in Lalla and Pattarin (2001).

The transitions considered in the following only generate completed spells and $c_i$ is always equal to 1. If the heaping effect is not considered, the log-likelihood function is then given by

$$
\log L(\alpha, \beta, \theta) = n \log(\alpha) + n \log(\beta) - \sum_{i=1}^{n} X_i \beta_i + \sum_{i=1}^{n} \log \left( \pi \tilde{\psi}(t_i, x_i; \theta, \beta) + (1 - \pi) \tilde{\psi}(t_i, x_i; \beta, \beta) \right).
$$

where $n$ is the total number of observations, $m$ is the total number of covariates, $\tilde{\psi}(\cdot)$ is the density function of $T_i$ (which could be $T_{A,i}$ or $T_{R,i}$), $\alpha$ is the shape parameter, $\theta$ is the location parameter, and $\beta$ is the usual vector of coefficients for the covariates. The density functions $\tilde{\psi}(\cdot)$ are derived from equation (4) and assume the form

$$
\tilde{\psi}(t_i, x_i; \theta, \alpha, \beta) = \frac{t_i^{\alpha-1}}{\left[t_i^\alpha + \theta^\alpha \exp(-\sum_{j=1}^{m} X_j \beta_j)\right]^2}.
$$

Inclusion of the heaping effect highly complicates the expression of $\tilde{\psi}(\cdot)$, as can be deduced from equation (7).

4. Differential Evolution: some background information

Evolutionary algorithms are stochastic search heuristic algorithms. They can deal with complex optimisation problems. An evolutionary algorithm starts by randomly generating a population of individuals, which encode feasible candidate solutions to the optimisation problem under investigation. Each individual is the mathematical encoding of one candidate solution to the optimisation problem under investigation. A fitness function, which measures the goodness of a candidate solution, is then computed for each individual in the population. The population is then altered, by iteratively applying some mathematical operators. This iterative process is usually defined as the evolution of the population through generations. For each generation some mathematical
operators, namely selection, recombination and mutation, are applied to the individuals of the current population in order to determine the composition of the population of the next generation. Such operators aim to determine candidate solutions with a better fitness value by exploring the search space. The population is evolved through different generations until the algorithm converges to a satisfying solution or the maximum number of iterations is reached. The fittest individual in the final population then encodes the optimal solution to the problem under investigation.

Genetic algorithms (Holland, 1975) are probably the most well-known evolutionary algorithms, even outside the heuristic community. They have been successful in tackling optimisation problems in different fields, such as statistics, engineering, and bioinformatics. Recently, many empirical investigations have shown that Differential Evolution (DE), a relatively unknown approach to numerical optimization (Storn and Price, 1995), has a superior performance compared to other optimisation and evolutionary algorithms. DE is very simple to implement and requires little or no parameter tuning.

DE starts by randomly generating an initial population \( P_{\text{gen}} \) of \( N \) strings \( x_i \) (\( i = 1, \ldots, N \)), which are the individuals in the population. Each individual (chromosome) consists of a string of \( D \) real-value cells (genes), which encodes a candidate solution (genome) to the optimisation problem under investigation. After random initialisation and fitness evaluation, the population is then evolved through generations \( \text{Gen}=1, \ldots, G_{\text{max}} \). In each generation, mutation, recombination and selection operators are applied to the individuals. In every generation, for each individual \( i \), we choose three other individuals \( k, l, \) and \( m \) randomly from the population (with \( j \neq k \neq m \)), calculate the difference of the individual in \( k \) and \( l \), scale it by multiplication with a parameter \( F \) and create an offspring by adding the result to the chromosome of \( m \). However, the entire individual of the offspring is not created in this way, but genes are partly inherited from individual \( i \). The proportion of genes that are also partly inherited is determined by the so-called crossover probability (\( CR \)), which determines how many genes of the difference vector are copied on the average in the offspring. More precisely, the operator iteratively copies consecutive genes (from a random starting point and continuing on with the first gene after the last gene has past) of the difference vector in the offspring until \( U(0,1) \geq CR \). If the offspring \( o \) is better than the parent \( i \), then \( i \) is substituted by \( o \).

The process is repeated for a fixed number of iterations and the optimization result is the best recorded candidate solution and fitness value at the end of the run. Different stopping criteria can be used. Table 4 reports the pseudo-code of the differential evolution (DE) algorithm.
4.1. Differential evolution encoding for the unemployment spells mixture

The DE algorithm could be useful in estimating the parameters of the mixture distribution of the unemployment spells when the spell reconstructions ($T_{R,i}$) and spell answers ($T_{A,i}$) are considered jointly in a mixture model. The aim is to determine the parameters that maximise the likelihood function of the mixture model or minimise the negative of the log-likelihood function (see equation 8).

The analysis was performed considering 2380 individuals, from a sample with no missing values, collected by ISTAT, considering only the transitions from the unemployment to the employment state (see Table 3). For each person in the sample, $m=45$ covariates were considered. As Fig. 4 shows, each DE individual encodes the following parameters: the mixture weight $\pi$, the shape parameter $\alpha$, the location parameter $\beta$, and the 45 coefficients of the covariates $(\beta_1, \ldots, \beta_m)$, where $m=nvar$. The DE explores to determine the parameters that minimise the following fitness function, which corresponds to the log-likelihood of the model described by (8) and (9).

$$fitness = -\log L(\pi, \beta, \alpha, \beta).$$ (10)

DE requires a priori setting of the search space domain and some evolutionary parameters. After some preliminary analyses, we set the population size at 147 (NIND=3*No. of parameters to be estimated), the maximum number of generations $G_{max}$ at 4000, the crossover rate at 0.5 (CR), and the DE-step size at 0.8 (F). Moreover, the search space was constrained so that the mixture weight $\pi$ can vary in [0,1], the shape parameter $\alpha$ in [0,4], the location parameter $\beta$ in [0,8], and the 45 beta parameters $(\beta_1, \ldots, \beta_m)$ in [-2,2]. During the simulations, a parameter was set to the closest domain boundary in case it proved to be equal to a value outside the domain search space. Future investigation will consider a larger search space.
Table 4
Pseudo-code of the DE algorithm

```c
void differential_evolution() {
    initialize();
evaluate();
    While (Gen<Gen_max, Gen++) {
        for (i=1; i<NIND; NIND++) {
            for (j=1; j<D; D++) {
                MutateandRecombine();
            }
            select();
evaluate();
        }
    }
}

Void MutateandRecombine(){
    o.genej = \[\begin{align*}
    & m.genej + F \cdot (k.genej - l.genej) & \text{if } U(0,1)<CR \\
    & i.genej & \text{otherwise}
    \end{align*}\]
}

Void select(){
    If fitness(offspring)>fitness(parent), choose offspring;
    Otherwise, choose parent;
}
```
5. Results of the estimation parameters

The data set was restricted solely to transitions from unemployment to employment, considering only job losers/leavers. The averages of the two durations, $t_{A,i}$ and $t_{R,i}$, were almost the same (see Table 5), but values of $t_{R,i}$ greater than 120 months were eliminated (see §2). In any case, this indicates that the two durations were consistent, on the average, while the mean of their minimums showed a reduction of five months. The means of $t_{A,i}$ and $t_{R,i}$ were somewhat higher than those obtained by Favro-Paris et al. (1996), while the mean of $\min(t_{A,i};t_{R,i})$ was slightly lower than their value. Only $t_{R,i}$ proved to have missing values for the unemployed and they amounted to 206 cases. To avoid the elimination of these statistical units, the missing values of $t_{R,i}$ were replaced with the corresponding available value of $t_{A,i}$, given that their means were almost the same.

The descriptive statistics of the main variables used to model unemployment durations are reported in Table 5, grouped by personal, family, individual job-search, local job-search, previous job, and current job characteristics. The explanatory variables were selected using two main criteria, although they admittedly remain arbitrary. The first used the backward method and maximisation of the partial likelihood estimates for removal, in Cox's semi-parametric proportional hazard model for $t_{A,i}$, for $t_{R,i}$, and for their minimum. The final set was obtained by matching the three outcomes. The second used both personal judgements and the main findings in the literature to select covariates from among those that did not enter the models.

The personal characteristics included gender, age, and education. The proportion of women involved in these types of transitions was only 40.5%, suggesting discrimination (and segregation) in the labour market. In fact, the number of unemployed women was generally greater than that of men, but only a lower proportion of them proved to have found employment and many of them passed into the out-of-the-labour-force state. Age was calculated at the beginning of unemployment and expressed in decades to reduce its value with respect to the dummies. It was introduced in the models as a polynomial of the second order to account for the different behaviours of individuals in different life phases. The level of education did not reveal a significant impact on the transition from unemployment to the employment state and it was not included in the model, for the sake of brevity, even if in other studies, they have sometimes yielded coefficients significantly different from zero (see, inter alia, Roed and Nordberg, 2003; Boheim and Taylor, 2002).
Table 5
Acronym (Acr.), Mean, Standard Deviation (SD), Range, and number of cases (n) –or 1 when the variables is binary (0/1)– for the main variables in the models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Acr.</th>
<th>Unemployed Mean</th>
<th>SD</th>
<th>Range</th>
<th>n or 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spell Answers, $t_{A;i}$ (months)</td>
<td></td>
<td>19.760</td>
<td>22.889</td>
<td>1.9–107</td>
<td>2585</td>
</tr>
<tr>
<td>Spell Reconstructions, $t_{R;i}$ (months)</td>
<td></td>
<td>19.352</td>
<td>22.952</td>
<td>1.5–119</td>
<td>2379</td>
</tr>
<tr>
<td>Min($t_{A;i};t_{R;i}$) – months</td>
<td></td>
<td>15.174</td>
<td>17.770</td>
<td>1.5–107</td>
<td>2585</td>
</tr>
<tr>
<td><strong>Personal characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>W</td>
<td>0.405</td>
<td>0.491</td>
<td>0/1</td>
<td>1046</td>
</tr>
<tr>
<td>Age (in decades)</td>
<td>A</td>
<td>3.276</td>
<td>1.069</td>
<td>1.1–8.0</td>
<td>2585</td>
</tr>
<tr>
<td>High School (Diploma)</td>
<td>HS</td>
<td>0.251</td>
<td>0.433</td>
<td>0/1</td>
<td>648</td>
</tr>
<tr>
<td>College Level (Laurea)</td>
<td>CL</td>
<td>0.056</td>
<td>0.229</td>
<td>0/1</td>
<td>144</td>
</tr>
<tr>
<td><strong>Family characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never Married</td>
<td>NMarr.</td>
<td>0.471</td>
<td>0.499</td>
<td>0/1</td>
<td>1217</td>
</tr>
<tr>
<td>Married</td>
<td>Marr.</td>
<td>0.489</td>
<td>0.500</td>
<td>0/1</td>
<td>1264</td>
</tr>
<tr>
<td>Separated, Divorced, Widowed</td>
<td>SDW</td>
<td>0.040</td>
<td>0.197</td>
<td>0/1</td>
<td>104</td>
</tr>
<tr>
<td>Child (compared to the head of family)</td>
<td>Child</td>
<td>0.418</td>
<td>0.493</td>
<td>0/1</td>
<td>1081</td>
</tr>
<tr>
<td>Number of Family Components</td>
<td>NFC</td>
<td>3.637</td>
<td>1.300</td>
<td>1–12</td>
<td>2585</td>
</tr>
<tr>
<td>Employment Ratio in the Family</td>
<td>ERF</td>
<td>0.191</td>
<td>0.204</td>
<td>0–0.75</td>
<td>2585</td>
</tr>
<tr>
<td><strong>Individual job-search characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-Term Contract</td>
<td>STC</td>
<td>0.041</td>
<td>0.197</td>
<td>0/1</td>
<td>105</td>
</tr>
<tr>
<td>Part-Time Work</td>
<td>PTW</td>
<td>0.102</td>
<td>0.303</td>
<td>0/1</td>
<td>264</td>
</tr>
<tr>
<td>Any Time</td>
<td>AT</td>
<td>0.236</td>
<td>0.425</td>
<td>0/1</td>
<td>610</td>
</tr>
<tr>
<td>Not Willing to Commute</td>
<td>NWC</td>
<td>0.331</td>
<td>0.471</td>
<td>0/1</td>
<td>856</td>
</tr>
<tr>
<td>Willing to Commute</td>
<td>WC</td>
<td>0.459</td>
<td>0.498</td>
<td>0/1</td>
<td>1187</td>
</tr>
<tr>
<td>Any Location</td>
<td>AL</td>
<td>0.210</td>
<td>0.407</td>
<td>0/1</td>
<td>542</td>
</tr>
<tr>
<td>Conditional Availability to Work</td>
<td>CAW</td>
<td>0.621</td>
<td>0.485</td>
<td>0/1</td>
<td>1604</td>
</tr>
<tr>
<td>Number of Job Search Attempts</td>
<td>JSA</td>
<td>1.974</td>
<td>1.365</td>
<td>0–10</td>
<td>2585</td>
</tr>
<tr>
<td>Reservation Wage</td>
<td>RW</td>
<td>1.694</td>
<td>1.589</td>
<td>1–3.5</td>
<td>2585</td>
</tr>
<tr>
<td>Registered with Employment Agency</td>
<td>REA</td>
<td>0.832</td>
<td>0.374</td>
<td>0/1</td>
<td>2150</td>
</tr>
<tr>
<td>Benefits</td>
<td>Benefit</td>
<td>0.164</td>
<td>0.370</td>
<td>0/1</td>
<td>424</td>
</tr>
<tr>
<td><strong>Local job-search characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provincial Unemployment Rate</td>
<td>PUR</td>
<td>15.028</td>
<td>8.724</td>
<td>2.2–34.4</td>
<td>2585</td>
</tr>
<tr>
<td>North-East (4, 5, 6)</td>
<td>NE</td>
<td>0.113</td>
<td>0.317</td>
<td>0/1</td>
<td>293</td>
</tr>
<tr>
<td>Centre: Emilia-Romagna</td>
<td>CE-ER</td>
<td>0.060</td>
<td>0.237</td>
<td>0/1</td>
<td>154</td>
</tr>
<tr>
<td>Centre (9, 10, 11)</td>
<td>CE</td>
<td>0.098</td>
<td>0.297</td>
<td>0/1</td>
<td>253</td>
</tr>
<tr>
<td>Latium</td>
<td>LZ</td>
<td>0.051</td>
<td>0.219</td>
<td>0/1</td>
<td>131</td>
</tr>
<tr>
<td>South-East (13, 14, 16)</td>
<td>SE</td>
<td>0.138</td>
<td>0.345</td>
<td>0/1</td>
<td>357</td>
</tr>
<tr>
<td>1st Quarter</td>
<td>D1Q</td>
<td>0.234</td>
<td>0.424</td>
<td>0/1</td>
<td>606</td>
</tr>
<tr>
<td>2nd Quarter</td>
<td>D2Q</td>
<td>0.289</td>
<td>0.453</td>
<td>0/1</td>
<td>747</td>
</tr>
<tr>
<td>3rd Quarter</td>
<td>D3Q</td>
<td>0.133</td>
<td>0.340</td>
<td>0/1</td>
<td>345</td>
</tr>
<tr>
<td>4th Quarter of previous year</td>
<td>D4Qpy</td>
<td>0.220</td>
<td>0.414</td>
<td>0/1</td>
<td>568</td>
</tr>
</tbody>
</table>

(continued on the next page)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Acr.</th>
<th>Unemployed Mean</th>
<th>SD</th>
<th>Range no. of</th>
<th>n or no. of</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Previous job characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dismissed from the Previous Job</td>
<td>DPJ</td>
<td>0.240</td>
<td>0.427</td>
<td>0/1</td>
<td>621</td>
</tr>
<tr>
<td>Left for Other Reasons</td>
<td>LOR</td>
<td>0.203</td>
<td>0.403</td>
<td>0/1</td>
<td>526</td>
</tr>
<tr>
<td>Agriculture Sector</td>
<td>AS1</td>
<td>0.123</td>
<td>0.329</td>
<td>0/1</td>
<td>318</td>
</tr>
<tr>
<td>Construction Sector</td>
<td>CS1</td>
<td>0.175</td>
<td>0.380</td>
<td>0/1</td>
<td>453</td>
</tr>
<tr>
<td>Trade Sector</td>
<td>TS1</td>
<td>0.220</td>
<td>0.414</td>
<td>0/1</td>
<td>569</td>
</tr>
<tr>
<td>Service Sector</td>
<td>SS1</td>
<td>0.101</td>
<td>0.302</td>
<td>0/1</td>
<td>262</td>
</tr>
<tr>
<td>Public Administration</td>
<td>PA1</td>
<td>0.115</td>
<td>0.319</td>
<td>0/1</td>
<td>297</td>
</tr>
<tr>
<td><strong>Current job characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working in a Different Region</td>
<td>WDR</td>
<td>0.474</td>
<td>0.499</td>
<td>0/1</td>
<td>1226</td>
</tr>
<tr>
<td>Forced Part-Time</td>
<td>FPT</td>
<td>0.147</td>
<td>0.354</td>
<td>0/1</td>
<td>380</td>
</tr>
<tr>
<td>Chosen Part-Time</td>
<td>CPT</td>
<td>0.063</td>
<td>0.242</td>
<td>0/1</td>
<td>162</td>
</tr>
<tr>
<td>Forced Temporary Employment</td>
<td>FTE</td>
<td>0.271</td>
<td>0.444</td>
<td>0/1</td>
<td>700</td>
</tr>
<tr>
<td>Chosen Temporary Employment</td>
<td>CTE</td>
<td>0.172</td>
<td>0.377</td>
<td>0/1</td>
<td>444</td>
</tr>
<tr>
<td>Owner Position</td>
<td>OP</td>
<td>0.042</td>
<td>0.201</td>
<td>0/1</td>
<td>109</td>
</tr>
<tr>
<td>Professional Man</td>
<td>PM</td>
<td>0.109</td>
<td>0.312</td>
<td>0/1</td>
<td>283</td>
</tr>
<tr>
<td>Coadjuvant Position</td>
<td>CP</td>
<td>0.026</td>
<td>0.160</td>
<td>0/1</td>
<td>68</td>
</tr>
<tr>
<td>Agriculture Sector</td>
<td>AS2</td>
<td>0.159</td>
<td>0.366</td>
<td>0/1</td>
<td>412</td>
</tr>
<tr>
<td>Construction Sector</td>
<td>CS2</td>
<td>0.168</td>
<td>0.374</td>
<td>0/1</td>
<td>435</td>
</tr>
<tr>
<td>Trade Sector</td>
<td>TS2</td>
<td>0.261</td>
<td>0.439</td>
<td>0/1</td>
<td>674</td>
</tr>
<tr>
<td>Service Sector</td>
<td>SS2</td>
<td>0.138</td>
<td>0.345</td>
<td>0/1</td>
<td>356</td>
</tr>
<tr>
<td>Public Administration</td>
<td>PA2</td>
<td>0.113</td>
<td>0.317</td>
<td>0/1</td>
<td>293</td>
</tr>
</tbody>
</table>

Family characteristics were limited to marital status, the «position» in the family, the number of family components, and the employment ratio in the family. Marital status was subdivided in three dummies: never married, married, and other (separated, divorced, widowed). The position in the family contains dummies distinguishing between head of family, spouse, child, and other. The employment ratio is the division of the number of employed by the number of family components, ranging from 0 to 1.

The individual job-search characteristics also included three covariates that were not significant. Conditional availability to work is a negative attitude with respect to unconditional availability, the latter being a preliminary condition to be classified as unemployed. The number of job search attempts represents an indicator of the intensity of the job-search, but a high value could denote some intrinsic problems of job seekers (Favro-Paris et al., 1996). The reservation wage is an important reference for job search theory (Lippman and McCall, 1976), but it has not proven to be significant in other studies (Lalla and Pattarin, 2001). However, it is known that answers regarding wages are generally missing and often unreliable, introducing a bias in the estimates.
The local job-search characteristics include the Provincial unemployment rate, which was constant over the period considered. This rate captures local differences without fulfilling them. Therefore, some dummies for the geographic areas of Italy were introduced in the model using the aggregations proposed by Attanasio and Padoa Schioppa (1991), with a small change in the definition of Centre where the region of Emilia-Romagna was introduced as a single area (see Table 5). Other dummies concern the temporal period (i.e., quarter) before finding a job: the reference dummy was the fourth quarter of 1998.

All the available previous job characteristics were included in the model and they proved to have a significant impact on the probability of leaving the unemployment state, while current job characteristics did not always yield coefficients significantly different from zero. In particular, the sectors proved to be significant only in the model based on $t_{R,i}$. However, the latter were included for completeness.

Table 6 reports the estimates of the parameters of the spell mixture models when 15 runs of DE were considered. In each run, the population was iteratively evolved for 4000 generations. Fig. 4 shows that (after 500 generations) the best fitness value in each generation decreased markedly at the beginning (0-500 generations) and then (after 500 generations) the fitness value of the best individual continued to decrease, but in a more stable manner. Four thousand generations can be acceptable even if a better solution to the optimisation problem is found when the number of generations increases. Column 2 shows the «best» values of $\beta$, column 3 the mean values, and column 4 the standard error of the parameters for 15 runs. The other columns show, respectively, the value of $\beta$, its standard error, the relative p-value ($p < $), and the corresponding estimates of the percentage of change in risk with each unit change in the covariate or $\exp(\beta)$, which can be interpreted as the relative risk when the covariate takes only the values of 0 and 1. Cox's model was based only on the spells given by $\min(t_{A,i}, t_{R,i})$ and the parameters $\lambda$ and $\alpha$ were estimated through a nonlinear model of the empirical baseline cumulative hazard. The coefficient of determination for this post-hoc estimation was $R^2 = 0.992$.

The estimated value, $\hat{\alpha}$, of the shape parameter is particularly interesting because some labour market theories predict $\hat{\alpha} < 1$, implying negative state dependence (Salant, 1977), but the empirical results showed all possible values, i.e., either negative (Ham and Rea, 1987; Torelli and Trivellato, 1988; Lynch, 1989), or positive (Hujer and Sneider, 1989), or null (Flinn, 1986; Groot, 1990).

All experiments were run using Windows XP on an IBM desktop PC. With the parameters described above, each DE simulation required 8h 26m. The estimation of the parameters in Cox's regression model was obtained through SPSS (1997).
state dependence, although the segment of the surveyed population was not always the same. The estimated shape, $\hat{\alpha}$, showed a value slightly, but significantly, lower than 1, which could depend on the noticeable heterogeneity not being captured completely by the selected covariates. The estimates of $\hat{\beta}$ (or $\lambda$) refer to the average spell length, as it is the scale parameter. The $\hat{\beta}$ coefficient represents the predicted change for a unit increase in the predictor. Therefore, a positive value implies an increase in the probability (acceleration) of leaving the unemployment state with the $\exp(\beta_i)$ being greater than 1, while a negative value entails a decrease in the probability (deceleration) of leaving the unemployment state with the $\exp(\beta_i)$ being lower than 1.

![Best Fitness Value in 4000 generations](image)

Fig. 4. Fitness value of the DE optimal solution in 4000 generations.

Of the personal and family characteristics, only the "child" covariate proved to have a positive coefficient $\hat{\beta}$ statistically different from zero, implying a reduction of unemployment duration. Children looking for a job could be induced to accept an offer because either the family could be pressed for money or the employed in the family could represent an aid in finding a job, but a negative sign also has an interesting interpretation. Therefore, its sign is not univocally determined a priori.

Among the individual job-search characteristics, only the covariates REA and "Benefits" proved to be significant. The coefficient of REA was negative implying, surprisingly, people registered with employment agencies tend to increase their unemployment spells, although the registration would denote a
deliberate intention to find a job. A positive coefficient for benefits does not have a causal interpretation, as it could represent a spurious correlation with the omission of important firm-specific variables from the model (Roed and Nordberg, 2003). However, it is often argued that recipients should wait longer than those without benefits, but the empirical outcomes sometimes contradict one another (see, inter alia, Nickell, 1979b; Narendranathan et al., 1985). In this case, benefits tend to decrease unemployment duration.

The significant local job characteristics included the Provincial unemployment rate with an expected negative coefficient and the dummies for the North-East with an expected positive coefficient. The other dummies — Latium, the South-East, the second quarter, and the fourth quarter of the previous year— yielded negative coefficients. The second and the fourth quarters denote some seasonal effects on unemployment flows, as they indicate periods in which job offers are high for seasonal workers.

All the available previous job characteristics proved to be significant and the coefficients showed the expected signs. Dismissed workers and individuals belonging to the residual category of “leavers for other reasons” showed longer unemployment lengths than those of voluntary leavers. Workers in industry proved to have spells longer than those of workers in agriculture, construction, trade, services, and also even public administration. For agriculture, construction, and the trade sector it could depend on the presence of a large number of seasonal workers, while the service sector and public administration could denote the specificity of the current labour market in which industry tends to be less important in the dynamics of unemployment.

The selected current job characteristics were all significant, except the dummies for the sectors. Individuals accepting (forced or chosen) temporary employment showed an increase in the probability of leaving the unemployment state. Owner positions, professionals and coadjutant positions yielded negative coefficients with respect to the other positions, such as clerical, low- and high-skilled workers.

Estimates of parameters, obtained through the DE algorithm, concerned the values of the “best” $\hat{\beta}$ coefficients, i.e., the values of parameters generating the lowest value of the likelihood. Each run produces a set of values that is the best in that population. Therefore, the mean of the various “best values” of the same $\hat{\beta}$ coefficient, obtained in the 15 runs, is also reported. Finally, the evaluation of the significance of the parameters required the calculation of the corresponding standard errors. The naïf procedure adopted estimated the standard error of each parameter using the standard deviation of the different best values obtained in the various runs. The results are reported in Table 6.
Table 6
Estimates of parameters of the mixture model for completed spells of unemployment obtained by the DE method — the "best" $\hat{\beta}$, the mean of the $\hat{\beta}$'s, and the relative standard error SE($\hat{\beta}$) — and for Cox's proportional hazards model, using min($t_{A;i}; t_{B;i}$)

<table>
<thead>
<tr>
<th>Variables</th>
<th>DE estimates</th>
<th>Cox's proportional hazards model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Best $\hat{\beta}$</td>
<td>$\hat{\beta}$</td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>0.276</td>
<td>0.267</td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>2.133</td>
<td>2.094</td>
</tr>
<tr>
<td>$\hat{\alpha}$ / $\hat{\lambda}$</td>
<td>6.254</td>
<td>6.472</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>W</td>
<td>0.514</td>
<td>-0.026</td>
</tr>
<tr>
<td>A/10</td>
<td>0.334</td>
<td>0.333</td>
</tr>
<tr>
<td>(A/10)**2</td>
<td>-0.039</td>
<td>-0.028</td>
</tr>
<tr>
<td>Child</td>
<td>0.155</td>
<td>0.317</td>
</tr>
<tr>
<td>NCF</td>
<td>0.089</td>
<td>0.096</td>
</tr>
<tr>
<td>ERF</td>
<td>0.077</td>
<td>0.168</td>
</tr>
<tr>
<td>STC</td>
<td>-0.077</td>
<td>0.416</td>
</tr>
<tr>
<td>PTW</td>
<td>-0.471</td>
<td>-0.285</td>
</tr>
<tr>
<td>AT</td>
<td>-0.322</td>
<td>-0.126</td>
</tr>
<tr>
<td>AL</td>
<td>0.529</td>
<td>0.146</td>
</tr>
<tr>
<td>CAW</td>
<td>0.441</td>
<td>0.107</td>
</tr>
<tr>
<td>JSA</td>
<td>-0.160</td>
<td>-0.056</td>
</tr>
<tr>
<td>RW</td>
<td>0.031</td>
<td>-0.024</td>
</tr>
<tr>
<td>REA</td>
<td>-0.706</td>
<td>-0.604</td>
</tr>
<tr>
<td>Bnft</td>
<td>1.080</td>
<td>0.848</td>
</tr>
<tr>
<td>PUR</td>
<td>-0.036</td>
<td>-0.051</td>
</tr>
<tr>
<td>NE</td>
<td>0.635</td>
<td>0.282</td>
</tr>
<tr>
<td>CE-ER</td>
<td>0.299</td>
<td>-0.034</td>
</tr>
<tr>
<td>CE</td>
<td>-0.321</td>
<td>-0.441</td>
</tr>
<tr>
<td>LZ</td>
<td>-1.340</td>
<td>-0.965</td>
</tr>
<tr>
<td>SE</td>
<td>-0.158</td>
<td>-0.083</td>
</tr>
<tr>
<td>D1Q</td>
<td>-1.245</td>
<td>-0.592</td>
</tr>
<tr>
<td>D2Q</td>
<td>-1.223</td>
<td>-0.835</td>
</tr>
<tr>
<td>D3Q</td>
<td>-0.005</td>
<td>0.153</td>
</tr>
<tr>
<td>D4Qpy</td>
<td>-1.304</td>
<td>-1.210</td>
</tr>
<tr>
<td>DPJ</td>
<td>-0.329</td>
<td>-0.315</td>
</tr>
<tr>
<td>LOR</td>
<td>-0.110</td>
<td>-0.367</td>
</tr>
<tr>
<td>A1</td>
<td>1.166</td>
<td>1.357</td>
</tr>
<tr>
<td>CS1</td>
<td>0.375</td>
<td>0.648</td>
</tr>
<tr>
<td>TS1</td>
<td>0.522</td>
<td>0.722</td>
</tr>
<tr>
<td>SS1</td>
<td>0.226</td>
<td>0.665</td>
</tr>
<tr>
<td>FA1</td>
<td>1.022</td>
<td>1.222</td>
</tr>
</tbody>
</table>

(continued on the next page)
Table 6 (continued)

<table>
<thead>
<tr>
<th>Variables</th>
<th>DE estimates</th>
<th>Cox's regression model estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>SE($\hat{\beta}$)</td>
</tr>
<tr>
<td>WDR</td>
<td>$-0.708$</td>
<td>$0.223$</td>
</tr>
<tr>
<td>FPT</td>
<td>$-0.193$</td>
<td>$0.344$</td>
</tr>
<tr>
<td>CPT</td>
<td>$-0.563$</td>
<td>$0.565$</td>
</tr>
<tr>
<td>FTE</td>
<td>$0.270$</td>
<td>$0.194$</td>
</tr>
<tr>
<td>CTE</td>
<td>$0.049$</td>
<td>$0.333$</td>
</tr>
<tr>
<td>OP</td>
<td>$-1.592$</td>
<td>$0.708$</td>
</tr>
<tr>
<td>PM</td>
<td>$-0.533$</td>
<td>$0.403$</td>
</tr>
<tr>
<td>CP</td>
<td>$-0.774$</td>
<td>$0.663$</td>
</tr>
<tr>
<td>AS2</td>
<td>$-0.319$</td>
<td>$0.403$</td>
</tr>
<tr>
<td>CS2</td>
<td>$-0.497$</td>
<td>$0.425$</td>
</tr>
<tr>
<td>TS2</td>
<td>$-0.427$</td>
<td>$0.297$</td>
</tr>
<tr>
<td>SS2</td>
<td>$-0.514$</td>
<td>$0.341$</td>
</tr>
<tr>
<td>PA2</td>
<td>$-0.085$</td>
<td>$0.473$</td>
</tr>
<tr>
<td></td>
<td>$-\log(L)$</td>
<td>$8779.863$</td>
</tr>
</tbody>
</table>

$^a$ Border-Line.  $^*$ Significant at the 5% level.  $^{**}$ Significant at the 1% level.

The DE algorithm yielded the estimates of all parameters in the model. The estimate of $\pi$ indicated that $t_{R,i}$ was heavier than $t_{A,i}$, while the estimate of $\gamma$ proved to be negligible. It should be noted that the standard errors of the estimates were higher than those obtained by the partial maximum likelihood applied in estimating the parameters of the Cox’s proportional hazards model (Cox, 1972). Therefore, the adopted procedure should be refined or the BHHH method (Greene, 1993) should be used, i.e., the outer product of the gradient which requires only the calculation of the first derivatives. Moreover, it is not useful to compare the different significance of parameters, as the latter depend on the value of their corresponding standard errors and the models are not exactly the same. However, only two significant covariates presented different signs: the short-term contract (STC) parameter, which was significant at the 5% level, and the parameter of chosen temporary employment (CTE) with borderline significance (6%).

6. Conclusions

This paper has examined unemployment durations in Italy detected in the period ranging from 1997:4 to 1999:1. Two measures were determined: spell answers and spell reconstructions. As is known, results obtained from unemployment spells analyses alone are generally ambiguous. In fact, conclusions concerning policy indications are not grounded upon a valid and
stable basis (Heckman and MaCurdy, 1992). However, the analysis of unemployment duration remains important for an approximate understanding of the pattern of relationships between covariates.

On the average, the spell reconstructions proved to be similar, to the spell answers, when the former were bounded within the same range as the latter. Therefore, the missing data for one type of measure could be (and were) replaced with the data of the other type.

The use of the mixture distribution to model unemployment spell lengths could constitute an alternative «data-based» strategy against the choice of the minimum between the two types of spells. The mixture parameter could be interpreted as the weight of one with respect to the other, but it tends to increase the average length of the spells.

The DE method proved to be a valuable alternative to classical methods, although it is computationally expensive and convergence can be slow. However, it appears to be very useful in dealing with the complex log-likelihood function because the optimisation problem is then NP-hard and classical methods cannot provide a solution. In fact, the latter require the determination of the first (and sometimes also the second) derivatives, a low number of covariates, and a rather simple type of function. On the contrary, although an unsolved problem still remains concerning the calculation of the standard errors and the procedure used above requires further refinements, the limitations of the other methods are not present in the DE method.

Acknowledgements

The work reported here was partially supported by a grant from the Italian Ministry for Education, University and Research (MIUR) for the project, «Industrial Districts as Complex Systems», approved in 2002 (prot. 2002133972) and directed by David Avra Lane. Partial financial support was also provided by the Italian MIUR for the project «Dynamics and inertia in the Italian labour market: databases, measurement and modelling issues», coordinated by Ugo Trivellato. Individual anonymous data, elicited from the Italian Labour Force Survey, were kindly provided by Istat (the Italian national statistical agency), under a research agreement with the Department of Statistics of the University of Padua.
References


188. Elena Bonfiglioli, Paolo Bosi, Stefano Tosi [1996] “L'equilibrio del contributo straordinario per l'Europa” pp. 20
190. Gianna Boero, Costanza Torricelli [1997] “The Expectations Hypothesis of the Term Structure of Interest Rates: Evidence for Germany” pp. 15
193. Mauro Dell'Amico [1997] “An Linear Time Algorithm for Scheduling Outforests with Communication Delays or Two Processor” pp. 18
196. Mauro Dell'Amico, Francesco Maioli e Marco Tolum [1997] “New bounds for optimum traffic assignment in satellite communication” pp. 21
198. Barbara Pistorese e Marcello D'Amato [1997] “Persistence of relative unemployment rates across Italian regions” pp. 25
204. Francesca Bergamin e Stefano Bondini [1997] “Una verifica empirica di un nuovo metodo di selezione ottima di portafoglio” pp. 22
223. Stefano Bordoni [1997] “Supporte Informatici per la Ricerca delle soluzioni di Problemi Decisionali” pp. 30
Measure of Comovement for Economic Indicators: Theory and Evidence... [1998]

Marcello D'Amato e Barbara Pistoresi [1998] “Delegation of Monetary Policy to a Central Banker with Private Information” pp. 15


Carlo Alberto Magni [1998] “A systemic rule for investment decisions: generalizations of the traditional DCF criteria and new conceptions” pp. 30

Marcello D'Amato e Barbara Pistoresi [1998] “Interest Rate Spreads Between Italy and Germany: 1995-1997” pp. 16

Paola Bertalini e Alberto Bertacchini [1998] “Il distretto di lavorazione earn sums in province of Modena” pp. 29

Costanza Toncelli e Gianluca Di Lorenzo [1998] “Una nota sui fondamenti matematico-finanziari della teoria delle aspettative della struttura della scadenza” pp. 15


Cesare Alberto Magni [1998] “Note sparse sul dilemma del pregiudizio (e non solo)” pp. 15


476 Anna Maria Sala [2004] “Sistema urbano e dinamiche insediative Una verifica empirica” pp

477 Paolo Bosi [2005] “Paradigmi economici e riforma del welfare nelle politiche europee” pp. 16

478 Claudio Marra [2005] “Percorsi, aspettative e valutazioni nell’esperienza lavorativa degli immigrati stranieri in Emilia Romagna: i casi di Modena e Reggio Emilia” pp 39

479 Paolo Bosi, M.Cecilia Guerra, Paolo Silvestri [2005] “Il finanziamento dei servizi per la non autosufficienza nel quadro della riforma del Titolo V” pp. 49


482 Gianluca Di Lorenzo e Giuseppe Marotta [2005] “Una politica monetaria meno efficace con l’UME? Evidenza dal passthrough nei tassi d’interesse attivi” pp. 28


485 Margherita Russo e Federica Rossi [2005] “Stimolare l’innovazione con strumenti innovativi: reti di partenariato e sviluppo locale nei programmi comunitari” pp. 32

486 Luigi Brighi e Reinhard John [2005] ”A Hypothesis Guaranteeing the Weak Weak Axiom” pp. 11

487 Massimo Baldini e Luca Beltrametti [2005] “Modelli di finanziamento di un fondo pubblico per la non autosufficienza” pp. 26