Unemployment at Entry and the Occupational Injury Risk of Young Workers*

Roberto Leombruni (University of Torino and LABOR)

Tiziano Razzolini (University of Siena and LABOR)

Francesco Serti (University of Alicante)

September 14 , 2014

Abstract

Using a unique dataset from Italy, we analyze the impact of the initial and contemporaneous unemployment rates on the workplace injury hazards of young workers entering the labor market in 1994-2003. Our findings confirm that workplace injuries are procyclical and also indicate that a 1 percentage point increase in the unemployment rate at entry reduces the time to the first injury by a factor of 0.80. Assuming workplace injuries are associated with low quality jobs, our results indicate that unfavourable initial labor market conditions negatively affect human capital accumulation and the career prospects of young workers.

Keywords: Initial labor market conditions, Procyclicality of injuries

JEL classification codes: J28, J24

*We thank Alberto Dalmazzo, Anzelika Zaiceva and participants at the EALE Conference 2012 and at the seminars of the University of Alicante and the University of Siena for helpful comments.

1 Introduction

Epidemiological studies have extensively demonstrated that inexperience at work is one of the primary cause of workplace accidents and that younger workers are more likely to incur non fatal injuries (Breslin and Smith, 2006; Salminen, 2004). However, along with other components of the job compensation package such as wages (Oreopoulos et al., 2008; Kwon et al., 2010; Kahn, 2010; Brunner and Kuhn, 2010), workplace safety of young workers may be also affected by the initial conditions faced at the time of entry.

This paper mainly focuses on the effects of unemployment rate at entry and current unemployment on job-related injury hazard rates for Italian labor market entrants. Thanks to a unique employer-employee dataset, survival analysis is employed to fully exploit available information on the exact entry date and the timing of each workplace injury. Work histories from the Italian administrative data ("Work Histories Italian Panel", WHIP) have been combined with individual work-related injuries from the Italian Workers' Compensation Authority (INAIL) for the period 1994-2003. As the literature suggests, we address the potential endogeneity of labor market entry over the business cycle, both from the perspective of the timing of entry and sorting into different local labor markets. Our results show the significant effects of past and current unemployment levels on the injury risk-experience profiles of young workers.

Previous studies (Kossoris, 1938; Boone and van Ours, 2006; Boone et al., 2011) found that job-related accident rates are pro-cyclical. This evidence seems to indicate that workplace safety decreases during economic expansions. This is consistent with Ruhm's (2000) conclusion, who attributed the positive relationship between macroeconomic conditions and mortality to, inter alia, a decrease in job safety due to the high effort levels that workers exert during economic expansions. Boone and van Ours (2006, 2011) offer an alternative explanation based on the reporting behaviour of workers: reporting a moderate accident increases the probability of being laid off; therefore, workers under(over)-report moderate injuries when the probability of being displaced is high(low), such as during recessions(expansions).

The effects of business cycle conditions on the entire set of job attributes may vary depending on the institutional settings in the labor market. As shown by the literature on job displacements, the shock of a job loss is more likely to affect unemployment spells than wages in European countries characterized by strict employment legislations (Hijzen et al., 2010; Kuhn, 2002). Apart from delaying labor market entry, bad labor market conditions cannot force young entrants to remain unemployed indeterminately; other job characteristics are likely to absorb negative shocks. Indeed, the study by Leombruni et al. (2013) has shown that Italian laid-off workers do not experience relevant wage losses, but face longer unemployment rates and most importantly exhibit a considerable increase in workplace risk at reemployment. Thus, wage compression and employment protection legislation may amplify the effect of negative shocks on other job amenities, such as workplaces safety.[1]

A persistent effect of initial macroeconomic conditions on labor market outcomes can be the consequence of various mechanisms, including the evolution of the quality of the initial job across the economic cycle and its effects on-the-job human capital accumulation; employers' imperfect information about workers' productivity; and long-term contracts. According to Gibbons and Waldman (2006), if new entrants facing poor macroeconomic conditions enter the labor market in lower-quality jobs or in tasks which offer relatively fewer opportunities for skill accumulation and career prospects (or are characterized by a lower transferability of the accumulated skills to higher level occupations), initial job or task assignment may have longer term consequences by affecting actual worker's productivity.[2] In addition, prospective employers could perceive the initial low-rank job as a signal of the worker's ability, without taking into account the macroeconomic conditions at the time of labor market entry (Oyer, 2006). Finally, a persistent effect of macroeconomic conditions at entry is also consistent with the existence of implicit/insurance contracts that insure workers against wage decreases (e.g., Harris and Holstrom, 1982;

[1] Job search literature focusing on labor markets with a less compressed wage distribution has emphasized the importance of non-wage job characteristics. See among others Blau (1991) and Dey and Flinn (2005).

[2] Empirical evidence indicates that during recessions workers are more likley to enter in lower-level jobs. Solon et al. (1997) and Devereux (2000) study the influence of the business cycle on workers' current job assignments, and show that during slowdowns workers are assigned to lower skilled jobs. Baker et al. (1994) find that during a negative macroeconomic period a higher proportion of newly hired workers are placed in lower job levels.

Beaudry and DiNardo, 1991). For example, workers entering the labor market during recessions may tend to accept long-term contracts characterized by low wages and job mobility may be costly (Kwon et al., 2010).

Our empirical analysis complements the extant literature in a number of ways. Although several studies have analyzed the relationship between the (initial) type of occupation and general health outcomes (Sindelar et al., 2007; Fletcher et al., 2011), no study to date has examined the effects of initial labor market conditions on workplace safety. First, the effect of the economic cycle on job safety is an important phenomenon that deserves attention because of the consequences in terms of lost work days and health assistance costs. Second, workplace safety is an important dimension that, together with wages, contributes to job quality. To the extent that injury hazard rates indicate job disamenities, physical effort and stress conditions, our results shed light on the link between macroeconomic entry conditions and subsequent job quality. By verifying whether initial unemployment levels accelerate or delay these hazard rates, and thus whether workplace safety improves with time, we also test the impact of starting conditions on time spent by new workers in low-quality jobs. The analysis of job-safety is particularly relevant for European countries such as Italy in which wages are regulated by stringent institutional rules and, therefore, the non-pecuniary characteristics of the jobs could represent an important dimension to account for in the overall variability of job quality. Finally, cohort differences in the distribution of hazardous tasks might have a long lasting effect and influence the interpretation of the empirical findings of the extant literature regarding procyclicality. This is because the elasticity of injury rates to the economic cycle could also be a function of the composition of tasks inherited from the past and, therefore, of the initial labor market conditions. Hence, our analysis also provides evidence on the sensitivity of the contemporaneous unemployment-injuries nexus to different starting conditions.

The remainder of the paper is organized as follows. The next section provides some institutional details and decribes the data. The econometric framework is presented in Section 3. Section 4 reports the empirical results, Section 5 describe the robustness check and and Section 6 concludes our discussion.

2 Data Description

We merged the WHIP dataset and the administrative records from the INAIL for 1994-2003. The resulting dataset provides a random sample of workers employed in the private non-agricultural sector of the Italian economy. The WHIP dataset contains data on the start and conclusion dates, as well as on the duration (number of weeks) of each employment relationship for the period 1985-2003. It also provides information on workers' characteristics (age, sex, birthplace, place of work, type of occupation, maternity leave and sick leave); standard labor market outcomes (number of weeks worked in a year and annual earnings); and characteristics of the firms in which individuals are employed (the number of employees, firms' opening and closing dates and sector). The INAIL dataset contains the date of workplace injuries (i.e., accidents that have occurred during a work task) and the duration of injury-related leaves at the employer-employee level for the period 1994-2003. It includes all injuries leading to a leave of more than three days.[3] Physicians reported and certified each diagnosis and prognosis for workers involved in these accidents.

Our sample includes men who had their first labor market experience between 1994 and 2003 and who were less than 24 years old at the time of entry.[4] Althought no information on schooling is available, the restriction on age in practice excludes individuals with higher education/skills (i.e., with at least a university degree)[5] and therefore reduces potential unobserved heterogeneity problems related to this important dimension. Moreover, job safety should be less relevant for labor market entrants with higher education, which tend to perform non-manual tasks. Therefore, due both to data limitations and

[3] Less serious injuries are not reported.

[4] We observe the labor market history of individuals from 1985 to 2003. We define "first time labor market entrants" those workers who are observed for the first time in the sample in 1994 or later on. By considering entrants who are less than 24 y.o. at the time of entry we exclude the pre-1971 birth cohorts. The resulting sample is representative of 70% of first time labor market entrants in Italy during 1994-2003.

[5] According to the AlmaLaurea surveys (http://www.almalaurea.it/en/), in 2003, only 0.7 % of students who completed their undergraduate studies were 23 y.o. or younger, 28 being the average age of graduation . During the previous years included in our sample the average age of graduation was even greater.

for conceptual reasons, we concentrate on low to medium skilled entrants. The restriction on gender is aimed to reduce the unobserved heterogeneity that reflects the complexity of female labor supply behaviour over the life cycle.

Following the above-cited literature, we use unemployment rates to proxy for the economic cycle. In particular, we use data on regional Italian unemployment rates over the period 1985-2003 from the Italian National Institute of Statistics (ISTAT). Given that our identification strategy uses the unemployment rate at the age of the expected labor market entry in the region of birth (e.g., to correct for the endogenous timing and location of the actual labor market entry), we select cohorts of individuals for which we can recover the regional unemployment rate of the region of birth at the end of compulsory schooling (i.e., 14 years old).[6] Therefore, our final sample consists of Italian-born, low to medium skilled men who started their first employment between 1994 and 2003, and who belong to the 1973-1989 birth cohorts.[7]

Figure 1 depicts the unemployment rate for the male population over the 1992-2004 period. The slowdown of the Italian economy after 1993 resulted in an increasing trend in uneployment until 1998. The recovery occurred thereafter.

[6]This restriction excludes the 1971 and 1972 birth cohorts and also foreign-born individuals, whose country of origin is unknown. The inclusion of these individuals does not change the main results in the baseline specifications.

[7]As the INAIL dataset does not cover the public sector, we also exclude labor market entrants and employment spells in these industries (the ATECO 1991/ISIC rev 1.1 codes: L, M, N, O). This hardly affects the representativeness of the data. Indeed, only 4.9% of the selected labor market entrants begin their career in the public sector and only 4.7% of individuals in the final sample have some job spells in the public sector.

Figure 1: Unemployment rate.

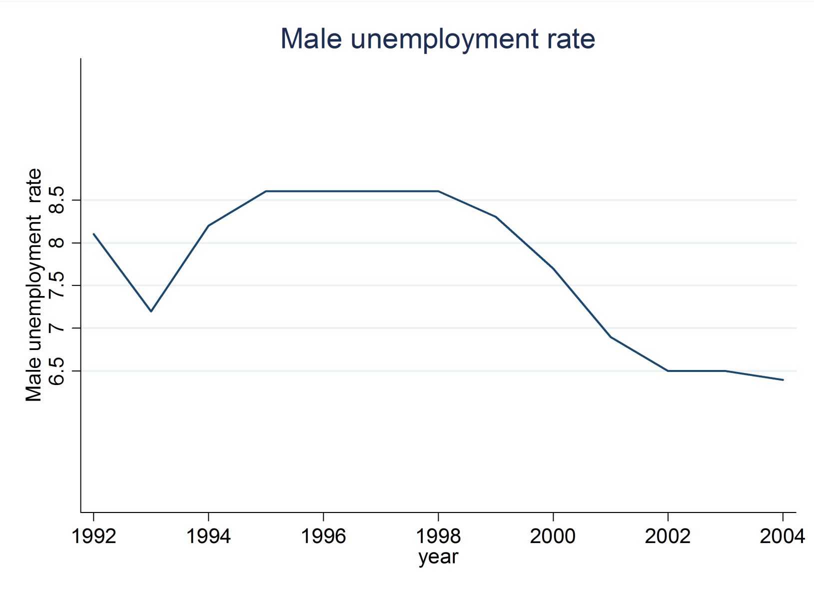
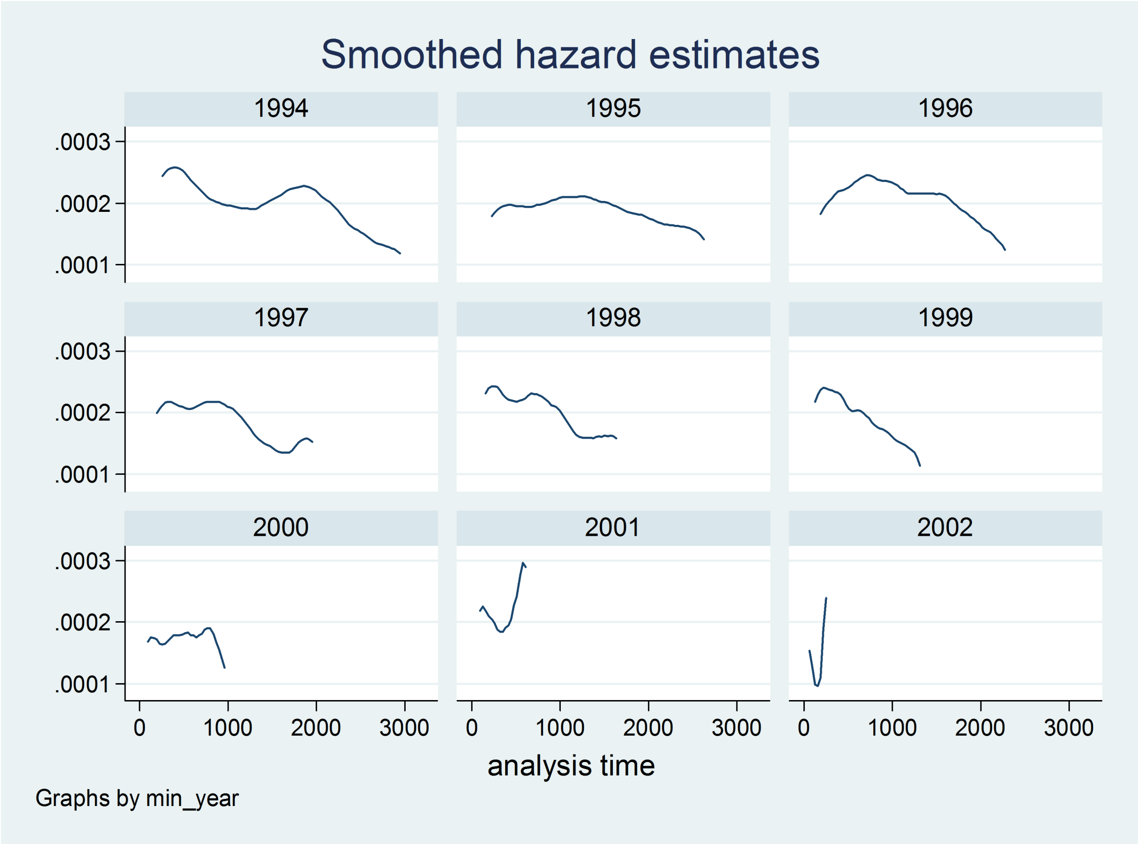


Figure 2 shows the smoothed injury hazard estimates for our sample by year of entry. The shape of the hazard seems to follow a lognormal distribution. The hazard rate declines as workers accumulate the experience to deal with risks. Interestingly, the hazard rate is low during the first days of a job, probably reflecting the fact that new entrants are trained and introduced to the new jobs in a sheltered environment, and then increases as the workers start to undertake their tasks independently. As it is evident from Figure 2, the hazard rates are sensitive to the business cycle. For example, the hazard rate of entrants in 1994 increased again after 1998 (around 1,500 days) as the unemployment rate started to decrease. Similarly, for entrants in 1995, 1996 and 1997 a positive relationship appears evident between the economy's upturn (which takes place approximately after three, two and one years respectively) and the hazard rates. Also notable, the estimated

Figure 2: Hazard rate by year.



hazard rates in 1999 and 2000, that is the first two years of the recovery of the Italian economy, exhibit a rapid decline after the first working days. The estimated hazard rates for entrants in 2001 and 2002, after a decline during the first working days, increase rapidly, probably because of the good performance of the economy. In general, the hazard rates during the 1994-1998 period of rising unemployment are at higher levels and are flatter than the hazard rates in the post 1998 recovery. This descriptive analysis suggests that hazard rates are not only sensitive to the business cycle but are also persistently higher for those who enter during relatively worse labor market conditions.

Panel A of Table 1 indicates the number of individuals in the sample, the number of failures (i.e., injured workers) and the number of never injured (censored) workers. Panel B of Table 1 shows the average and maximum survival time for these two types of workers.

Table 2 includes some descriptive statistics for our sample of workers by year of entry. Although a negative trend in the percentage of new entrants' manufacturing jobs is accompanied by a stable increase of the proportion of entrants in the service sector (e,g., commerce, hotels and restaurants and communication) and in other services (e.g., financial intermediation and business services), the share of workers starting their carrers as apprentices and blue collar workers is quite stable around the 90%. However, the categorization of blue versus white collars is likely to conceal a significant amount of heterogeneity in terms of task assignment.

The proportion of entrants born in the North of Italy exhibits a negative trend, whereas the percentage of workers born in the South and **in** the Islands increases over time. However, consistent with the historical economic duality between the richer North and the less developed Southern part of the country, more jobs for new entrants are created in the North. It is also interesting that the proportion of entrants in the Northern regiorns is higher during the years of increasing unemployment and decreases during the post-1998 recovery period when the proportion of jobs created in the Southern regions increases. Moreover, the difference between the proportion of entrants in the Southern(Northern) labor markets and the proportion of entrants born in the South(North) is always negative(positive). This evidence points to the relevance of mobility from the disadvantaged Southern regions toward the richer Northern part of the country.

Finally, the difference between the unemployment rate at age 14 in the region of birth and the current unemployment rate is smaller and negative during the period of the recession and becomes larger and positive during the economic upturn. These results, together with the relatively larger number and percentage of entrants in the period of recovery (shown in the last rows of Table 2), suggest that some individuals might delay the time of entry and wait for more favourable labor market conditions.

3 Econometric framework

In the present paper, we are interested in the estimation of the effects of initial and current regional unemployment (denoted by ur_o and ur_t) on the injury hazard rate. Toward this aim, and following the descriptive statistics, we use the lognormal regression model in the accelerated failure time (AFT) metric.[8] In this parametrization we have:

[8] We have also estimated several Cox proportional hazard models. The results from Cox regressions indicate qualitatively the same effect as the lognormal regression.We do not show these results since the proportional hazard assumption is always rejected by tests based on Schoenfeld residuals. In particular, the analysis of the variable-by-variable tests reveals that the Schoenfeld residual for the initial unemployment rate varies with time.

$$\varepsilon_i = \exp(-x_j\beta)t_j$$

where t_j is failure time and ε_i is distributed as a lognormal with mean μ and variance σ. In this specification a negative coefficient β accelerates failure time; in other words, injuries occur earlier.

In addition, we estimate a specification in which we model the variance σ as a function of the initial unemployment level and current unemployment. This requires estimating two additional coefficients, the ancillary parameters, which define the effect of the initial and current unemployment rates on the shape of the hazard distribution. As Zorn (2000) suggested, this flexible parametrization permits us to model time dependency if duration dependence is mainly due to state dependence.

As a robustness check, we also estimate the lognormal regression with frailty to control for potential unobserved heterogeneity (i.e., we add individual frailty or frailties that are shared by individuals in groups defined by year and region of first entry).

Moreover, we employ an additional specification in which we address the potential endogeneity of the timing and the region of entry, and thus of initial unemployment rate. Due to the non-linearity of duration models we cannot implement instrumental variable regressions. Using maximum likelihood, we jointly estimate the duration equation and the unemployment rate at entry. We assume that the initial level of unemployement is a function of unemployment at age 14 in the region of birth, region of birth dummies and a normally distributed error term η. Exploiting the properties of the lognormal-normal distributions, we model the hazard rates and the survival function conditional on the error term η , and we estimate the model by maximum likelihood. More details are provided in the Appendix.

To complete the picture and to provide new evidence on the Italian labor market, we analyze the effect of macroeconomic conditions on average weekly earnings (i.e., annual earnings divided by the number of weeks worked annually). We estimate a standard Mincer regression augmented with contemporaneous regional unemployment rate and the regional unemployment rate at the time of entry. We also interact the regional unemployment rate at the time of entry with potential experience (the number of years since entry) to allow the effect of initial macroeconomic conditions to vary over time. In all specifications we control for contemporaneous year effects, potential experience and squared potential experience.

As the timing and the location of initial labor market entry could be endogenous to macroeconomic conditions, we instrument for the regional unemployment rate at the time of entry with the unemployment rate in the region of birth at the end of compulsory schooling (i.e., 14 years old; denoted by ur_{14}). In both the first and second stage equations we control for contemporaneous year effects, birth cohort dummies and region of birth dummies. Finally, we instrument for the quadratic in potential experience and the interaction between potential experience and initial regional unemployment rate with a quadratic in age and the interaction between the square of age and the region of birth unemployment rate at the end of compulsory schooling.[9]

4 Results

Table 3 displays the estimated coefficients from the lognormal regressions. The first column reports the baseline specification using initial and current unemployment rates only. In column 2, we add dummies for the region of birth. In columns 3 to 12 we gradually introduce additional controls for geographical location and job characteristics that are not predetermined. We believe including dummies for the region of entry, although potentially endogeneous, is necessary to capture important regional differences in labor markets and in working conditions.

As can be seen from Table 3, the estimated coefficients of initial and current unemployment rates are statistically siginificant at least at the 5% level in all specifications. The coefficient on initial unemployment is relatively small in magnitude for the first two specifications. After we control for the region of entry, this coefficient increases and remains around -0.20 in all subsequent specifications. In quantitative terms, this means

[9] Oreopoulos et al. (2006, 2010), Kahn (2010) and Brunner and Kuhn (2010) adopt a very similar IV strategy.

that a one point increase in the initial unemployment rate decreases survival time by a factor of exp(-0.20)=0.80, that is, injuries occur earlier. The estimated coefficient of current unemployment is always in the range of 0.09 to 0.12, thereby indicating an increase in the survival time as the economy slows down. The inclusion of dummies controlling for current (and initial) sector and occupation does not alter the magnitude of the estimated coefficients. Therefore our results are not driven by changes in the distribution of entrants across industrial sectors and occupations. Nevertheless, the way tasks are assigned within occupations/sectors may vary according to the business cycle.

Table 4 reports the estimated coefficients from lognormal regression with frailties shared by individuals in groups defined by the year and region of entry.[10] The shared frailty term is assumed to have a gamma distribution. From the loglikelihood test we find that the shared frailty term is significant at the 1 % level. The coefficients of the initial and current unemployment, however, are unaffected by introducing frailties and remain at the same levels as in Table 3.

The results shown in Table 5 indicate that the introduction of ancillary parameters (to shape the hazard distribution and model time dependency) does not modify the values and the statistical significance of the initial and current unemployment coefficients. Interestingly, the ancillary parameters for ur_o and ur_t are negative and positive, respec-

[10]We have estimated an individual frailty model. We do not report the results since the coefficients of interest do not vary and the estimated frailty term is not significant . Estimates are available upon request.

tively. These results suggest that initial unemployment reduces σ and thus shifts the hazard distribution up and to the right. Intuitively, this means that for entrants in poor labor market conditions the instantaneous rate of injury is less concentrated in the first days of work and remains at higher levels as time continues. A higher level of current unemployment increases the concentration of injuries at early stages, but reduces the value of the hazard as t increases.

In Table 6, we report the coefficients from the joint estimation of the lognormal regression and the initial unemployment rate. The first panel shows the results for the lognormal regression. The estimates for the ur_0 equations are shown in the second panel. The estimates of the coefficient associated to ur_t are fairly stable in the range of 0.9 to 0.124 and close to the ones obtained using the same specification in Table 3. The effect of ur_0 increases when we control for its potential endogeneity. The effect of unemployment in the region of birth at the age of 14 on ur_0 is negative, thus suggesting that individuals facing particularly negative labor conditions at the age of 14 delay their entry and wait for more favourable economic conditions. The correlation coefficient ρ is positive but significant only in the specifications in columns 1 and 2. These results thus suggest that entry could be endogeneous and that during downturns only a sample of positively selected workers enter the labor marked in relatively higher quality jobs. This positive selection could potentially reduce the effect on ur_o on survival time.

Table 7 summarizes the results for the log of average weekly earnings. The first three columns show the estimates for different OLS specifications. In all regressions we control for contemporaneous year effects, potential experience and squared potential experience. In the first column we add only predetermined controls (region and year of birth dummies), whereas in the second specification we add dummies to indicate the region and the year of the first labor market experience. In the third column both sets of additional controls are included. The results are qualitatively similar in all three regressions. Contemporaneous unemployment is negatively correlated with weekly earnings. The magnitude of the above relationship (a point decrease in the contemporaneous unemployment rate is associated with a wage premium of about 0.5 to 0.7 %) suggests that the responsiveness of wages to local labor market conditions is relatively low. Unemployment at the time of entry has a positive effect on wages during the first years of labor market experience, thus pointing to a positive selection of entrants in periods of high unemployment. However, this effect is relatively low in magnitude and gradually vanishes as experience accumulates.

Once we instrument for potentially endogenous variables (column 4), contemporaneous unemployment and unemployment at the time of entry estimated coefficients turn out to be non statistically different from zero. Although the estimated coefficient on the interaction between initial unemployment and experience doubles and is statistically significant, the overall effect of unemployement at entry is not statistically different from zero for all possible values of experience. These results have important implications. First, wages seem to not respond to local unemployment rates, as one should expect given the high centralization of Italian wage setting institutions and the existence of downward wage rigidity (which can be particoularly relevant for this sample of young low to medium skilled labor market entrants that we consider).[11] Second, the larger losses due to the initial macroeconomic conditions (with respect to OLS results) corroborates the previous hypothesis (also drawn from the results in Table 6) on a possible positive selection of new labor market entrants during periods of relatively high unemployment. In other words, when macroeconomic conditions are relatively negative, only the more productive individuals succeed in entering the labor market.[12]

5 Robustness checks

In our survival analysis, the dataset and log likelihood function are set to account for interval truncations (Cleves et al., 2010), that is periods in which some workers are not observed because they are not employed in the sectors under analysis.[13] However,

[11] In Italy there is no legislated national minimum wage. Instead, minimum wage rates typically are set via industry-specific national collective bargaining agreements, which are then applicable to all workers in that industry. According to Dolado et al. (1996), the minimum wage as a fraction of average earnings in Italy is 0.71, which is the highest value across western European countries and USA during the ninentees.

[12] The estimated losses due to bad initial macroeconomic conditions are greater in the IV specifications (with respect to OLS) also in Kahn (2010) and Brunner and Kuhn (2010). OLS and IV estimates are similar in Oreopoulos et al. (2006, 2008).

[13] See Table A.2 for details.

non-employed workers may face different scenarios during these intervals. To check the sensitivity of our results to these interval truncations, in a first robustness check the exposure to risk is assumed to be zero during periods of non employment. The time elapsed in non-employment status is thus ignored and all employment spells are considered as contiguous.[14] In a second robustness check, individuals are assumed to be at risk and to experience zero injuries during periods of non-employment.[15] In both cases, the main results are qualitatively similar and significantly different from zero.

Because some individuals experience more than one injury, as shown in Table 1, we have conducted a multiple events analysis (Cleves, 2000). We have assumed different baseline specifications for the period until the first injury, the period between the first and the second injury and period after the third injury.[16] We ran this conditional risk model estimation with and without resetting the clock to zero after each injury (Cleves, 2000).[17] The results on initial employment and current unemployment are similar and remain statistically significant. However, we decide not to report these estimates because

[14] Let us consider, as an example, the spells of individual 1808 in Table A.1. This individual end an employment spell in year 1998 at day 14106. If we ignore the non-employment spell, the next spell will start at day 14106 instead of 14167, and will end at day 14183, instead of 14244. All start and end date of successive spells are rescaled accordingly (i.e. 77 days are subtracted to all these dates).

[15] Let us consider again individual 1808 in Table A.1. In this case we add a spell whose starting day coincides with the ending day of the previous spell and ends in the starting day of the next spell. Since during the non employment spells we do not observe the region of work we have imputed the region of work and the related current level of the unemployment rate of the region of work in the spell preceding the interval truncation.

[16] In the lognormal model this boils down to estimating a specific constant and a specific ancillary parameter for each strata.

[17] In the first case, the conditional-risk set model measures time to each event from the time of the previous event.

after an injury event the individual may experience a large reduction in working capacity and may rationally decide or be rationally allocated to a completely different task. The type of the job, its characteristics and therefore the hazard rate for the subsequent injury will thus be an endogenous result of the first injury.

6 Conclusions

The present paper investigates the effect of macroeconomic conditions on the injury hazard rates of male workers who entered the Italian labor market in 1994-2003. Our findings confirm a strong procyclicality of the injury hazard rates and, in addition, point to a positive and significant effect of unemployment rate at entry on the subsequent workplace-injury risk.

Although the Italian labor market is characterized by a strict employment protection legislation for "insiders", the result on the effect of the current unemployment level could be explained at least in part by the procyclicality of the workers' reporting decisions connected to the fear of being laid off during unfavorable local labor market conditions (as suggested by Boone et al., 2011 and Boone and van Ours, 2006). Consequently, it cannot be exclusively imputed to a worsening of job-safety during periods of intensive economic activity.

However, we do believe that higher hazard rates for entrants during recessions mainly indicate more time spent by these individuals in low quality matches. On the one hand, employers may vary the composition of the tasks offered along the business cyle, that is, they may take into account that during a recession it is relatively easy to fill in more risky/lower quality vacancies and offer this kind of job-tasks (Gibbons and Waldman, 2006). On the other hand, it is also reasonable to expect that the willingness of labor market entrants to accept relatively low-quality and more risky job-tasks, should be comparatively greater during slack labor market conditions. According to this view, it is likely that labor market entrants during unfavourable macroeconomic conditions end up in low-quality/more risky jobs that often only imperfectly match their ex-ante expectations and accumulated skills (probably causing low job satisfaction).[18] These explanations are consistent with the possibility that entering the labor market during poor macroeconomic conditions leads to disparities in the accumulation of the kinds of human capital associated with these low quality tasks. These workers invest more in this specific human capital and consequently they have a higher probability to persist in this kind of job or tasks. This persistence could be also due to stigma effects (i.e., with imperfect information future employers may take the worker's initial task as a signal of ability) and to the particular importance of early human capital investment, (i.e., early investments are particularly important bacause the individual can reap benefits over a

[18]It is also likely that an opposite mechanism applies to labor market entrants during favourable macroeconomic conditions.

longer span of time). Therefore, the effect of initial labor market conditions on the injury rates could be connected to persistence in more risky tasks and/or to the possibility that entrants during relatively bad labor market conditions report more injuries because they are less satisfied with the quality of their jobs.

These results complement the literature that investigates the effect of the initial labor market conditions on wages which has focused on countries other than Italy (Oreopoulos et al., 2008; Kwon et al., 2010; Kahn, 2010; Brunner and Kuhn, 2010). Our findings on wages, which do not seem to respond to local unemployment rates, are consistent with the high centralization of Italian wage setting institutions and with the existence of downward wage rigidity (which can be particularly relevant for this sample of young low to medium skilled labor market entrants we consider). For example, Bertola and Rogerson (1997) noted that the apparently surprisingly high rates of worker and job turnover in Italy in spite of the strict employment protection legislation are due to wage compression, that is, firms not being able to adjust the salary adjust the size of their workforce. The results presented in the present paper suggest additional adjustment mechanisms that are related to the initial unemployment rate and to job safety. With the business cycle, employers may vary the composition and the quality of the jobs offered at a constant wage, that is, during a recession, offering low quality/more risky jobs that workers are less willing to accept during other periods.

Our results are relevant for policymakers, because they identify the cohorts of workers at higher risk of workplace injury. The design of training programs and policies against injuries could be calibrated over the business cycle and optimally targeted to the correct population of workers in order to increase the returns and minimize the costs of such interventions.

References

Beaudry, P., DiNardo, J., August 1991. The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data. Journal of Political Economy 99 (4), 665–88.

Bertola, G., Rogerson, R., June 1997. Institutions and labor reallocation. European Economic Review 41 (6), 1147–1171.

Blau, D. M., April 1991. Search for nonwage job characteristics: A test of the reservation wage hypothesis. Journal of Labor Economics 9 (2), 186–205.

Boone, J., van Ours, J. C., November 2006. Are recessions good for workplace safety? Journal of Health Economics 25 (6), 1069–1093.

Boone, J., van Ours, J. C., Wuellrich, J.-P., Zweimuller, J., July 2011. Recessions are bad for workplace safety. Journal of Health Economics 30 (4), 764–773.

Breslin, F. C., Smith, P., 2006. Trial by fire: a multivariate examination of the relation between job tenure and work injuries. Occupational and Environmental Medicine 63 (1), 27–32.

Brunner, B., Kuhn, A., Dec. 2010. The impact of labor market entry conditions on initial job assignment, human capital accumulation, and wages. IZA Discussion Papers 5360, Institute for the Study of Labor (IZA).

Cleves, M., 2000. Analysis of multiple failure-time data with stata. Stata Technical Bulletin 9 (49).

Cleves, M., Gould, W. W., Gutierrez, R. G., Marchenko, Y., January 2010. An Introduction to Survival Analysis Using Stata. No. saus3 in Stata Press books. StataCorp LP.

Devereux, P. J., January 2000. Task assignment over the business cycle. Journal of Labor Economics 18 (1), 98–124.

Dey, M. S., Flinn, C. J., 03 2005. An equilibrium model of health insurance provision and wage determination. Econometrica 73 (2), 571–627.

Dolado, J.and Kramarz, F., Machin, S., Manning, A., Margolis, D., Teulings, C., 1996. The economic impact of minimum wages in europe. Economic Policy 23, 319–372.

Fletcher, J. M., Sindelar, J. L., Yamaguchi, S., May 2011. Cumulative effects of job characteristics on health. Health Economics 20 (5), 553–570.

Gibbons, R., Waldman, M., January 2006. Enriching a theory of wage and promotion dynamics inside firms. Journal of Labor Economics 24 (1), 59–108.

Harris, M., Holstrom, B., July 1982. A theory of wage dynamics. Review of Economic Studies 49 (3), 315–33.

Hijzen, A., Upward, R., Wright, P. W., 2010. The income losses of displaced workers. Journal of Human Resources 45 (1).

Kahn, L. B., April 2010. The long-term labor market consequences of graduating from college in a bad economy. Labour Economics 17 (2), 303–316.

Kossoris, M. D., 1938. Industrail injuries and the business cycle. Monthly Labour Review 46, 579–594.

Kuhn, P. J. (Ed.), Septiembr 2002. Losing Work, Moving On: International Perspectives on Worker Displacement. No. lwmo in Books from Upjohn Press. W.E. Upjohn Institute for Employment Research.

Kwon, I., Milgrom, E. M., Hwang, S., 2010. Cohort effects in promotions and wages: Evidence from sweden and the united states. Journal of Human Resources 45 (3).

Leombruni, R., Razzolini, T., Serti, F., 2013. The pecuniary and non-pecuniary costs of job displacement. the risky job of being back to work. European Economic Review 61 (0), 205 – 216.

Oreopoulos, P., Wachter, T. v., Heisz, A., Jul. 2008. The short- and long-term career effects of graduating in a recession: Hysteresis and heterogeneity in the market for college graduates. IZA Discussion Papers 3578, Institute for the Study of Labor (IZA).

Oyer, P., Summer 2006. Initial labor market conditions and long-term outcomes for economists. Journal of Economic Perspectives 20 (3), 143–160.

Ruhm, C. J., May 2000. Are recessions good for your health? The Quarterly Journal of Economics 115 (2), 617–650.

Salminen, S., 2004. Have young workers more injuries than older ones? an international literature review. Journal of safety research 35 (5), 513–521.

Sindelar, J. L., Fletcher, J., Falba, T., Keenan, P., Gallo, W. T., Dec. 2007. Impact of first occupation on health at older ages. NBER Working Papers 13715, National Bureau of Economic Research, Inc.

Solon, G., Whatley, W., Stevens, A. H., April 1997. Wage changes and intrafirm job mobility over the business cycle: Two case studies. Industrial and Labor Relations Review 50 (3), 402–415.

Zorn, C. J. W., 2000. Modeling duration dependence 8 (4), 367–380.

Table 1: Descriptive statistics on survival time for injured and never injured workers.

	(A) Statitics on the injured workers				
N. individuals	In the sample	with at least one injury	with a unique injury	with two injuries	with three or more injuries
	20575	4,966	3,840	815	311
	(B) Statistics on survival time (first injury or censoring)				
	Injured Workers		Never injured (censored)		
Average n. of days to the first injury	940.88 (801.05)		1200.47 (1013.81)		
Max number of days	3628		3637		

Table 2: Descriptive statistics on workers by year of entry

Year of entry	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Av. Age	18.904	19.208	19.215	19.359	19.344	19.492	19.501	19.392	19.404	19.343
% apprentice and blue collars	0.912	0.910	0.897	0.898	0.915	0.910	0.910	0.890	0.909	0.895
% entrants in manufactoring sect.	0.478	0.487	0.456	0.476	0.416	0.379	0.329	0.325	0.303	0.284
% entrants in construction sect.	0.184	0.180	0.188	0.163	0.193	0.175	0.206	0.195	0.200	0.205
% entrants in Service sector	0.285	0.263	0.289	0.285	0.310	0.315	0.291	0.290	0.343	0.355
% entrants in other services	0.049	0.067	0.064	0.076	0.077	0.128	0.171	0.188	0.152	0.155
Diff. between ur_{14} at and ur_0	-1.281	-1.049	-0.863	-0.142	-0.237	0.446	1.464	2.735	3.217	3.448
% entrants born in the North	0.532	0.536	0.510	0.479	0.472	0.435	0.411	0.406	0.365	0.383
% entrants born in the Center	0.163	0.154	0.147	0.168	0.159	0.185	0.171	0.162	0.183	0.174
% entrants born in the South and Islands	0.305	0.310	0.343	0.353	0.370	0.380	0.418	0.432	0.452	0.443
% entrants in North	0.567	0.588	0.568	0.541	0.534	0.499	0.471	0.476	0.428	0.449
% entrants in Center	0.172	0.166	0.171	0.186	0.181	0.210	0.198	0.192	0.203	0.196
% entrants in South and Islands	0.261	0.246	0.262	0.274	0.285	0.291	0.331	0.332	0.370	0.355
Number of entrants	1883	2228	2015	2007	2059	2190	2325	2162	1918	1788
% of entrants over the total period	0.092	0.108	0.098	0.098	0.100	0.106	0.113	0.105	0.093	0.087

ur_{14} is the unemployment rate in the region of birth whan the individual was 14 y.o.

Table 3: Lognormal regression – accelerated failure-time form

	1)	2)	3)	4)	5)	6)
Ur_0	-.045***	-.033**	-.193***	-.218***	-.210***	-.210***
	(.014)	(.014)	(.026)	(.037)	(.037)	(.037)
Ur_t	.087***	.092***	.090***	.120***	.127***	.128***
	(.014)	(.014)	(.014)	(.033)	(.033)	(.033)
Region of birth dummies	No	Yes	Yes	Yes	Yes	Yes
Region_0 dummies	No	No	Yes	Yes	Yes	Yes
Region_t dummies	No	No	No	Yes	Yes	Yes
Sector_0 dummies	No	No	No	No	Yes	Yes
Sector_t dummies	No	No	No	No	No	Yes
Occupation_0 dummies	No	No	No	No	No	No
Occupation_t dummies	No	No	No	No	No	No
year_t dummies	No	No	No	No	No	No
Age at entry	No	No	No	No	No	No
N. individuals	20575	20575	20575	20575	20575	20575
N. records	85071	85071	85071	85071	85071	85071
Log likelihood	-11315.75	-11269.697	-11213.792	-11202.782	-11131.782	-11063.079

Standard errors clustered at individual levels (in parentheses). *** significant at 1%,
** significant at 5%. Continues on the next page

Table 3: Continued

	7)	8)	9)	10)	11)	12)
Ur_0	-.212***	-.206***	-.208***	-.198***	-.203***	-.196***
	(.036)	(.036)	(.037)	(.037)	(.038)	(.038)
Ur_t	.116***	.117***	.122***	.123***	.111**	.109**
	(.032)	(.033)	(.033)	(.033)	(.046)	(.045)
Region of birth dummies	Yes	Yes	Yes	Yes	Yes	Yes
$Region_0$ dummies	Yes	Yes	Yes	Yes	Yes	Yes
$Region_t$ dummies	Yes	Yes	Yes	Yes	Yes	Yes
$Sector_0$ dummies	No	No	Yes	Yes	Yes	Yes
$Sector_t$ dummies	No	No	No	Yes	Yes	Yes
$Occupation_0$ dummies	Yes	Yes	Yes	Yes	Yes	Yes
$Occupation_t$ dummies	No	Yes	No	Yes	Yes	Yes
$year_t$ dummies	No	No	No	No	Yes	Yes
Age at entry	No	No	No	No	No	Yes
N. Individuals	20575	20575	20575	20575	20575	20575
N. Records	85071	85071	85071	85071	85071	85071
Log likelihood	-11093.086	-11001.359	-11047.801	-10901.472	-10891.519	-10853.757

Standard errors clustered at individual levels (in parentheses). *** significant at 1%,

** significant at 5%

Table 4: Lognormal regression with gamma distributed frailty at the level groups defined by year and region at entry.

	1)	2)	3)	4)
Ur_0	-.056***	-.039***	-.187***	-.235***
	(.015)	(.014)	(.027)	(.039)
Ur_t	.093***	.10***	.096***	.153***
	(.013)	(.014)	(.014)	(.034)
θ	.076***	.052***	.016***	.019***
	(.018)	(.015)	(.008)	(.009)
Region of birth dummies	No	Yes	Yes	Yes
Region_0 dummies	No	No	Yes	Yes
Region_t dummies	No	No	No	Yes
N. individuals	20575	20575	20575	20575
N. records	85071	85071	85071	85071
Log likelihood	-11274.903	-11248.00	-11208.789	-11196.467

Standard errors in parentheses. *** significant at 1%,** significant at 5%.

The significance level of θ refers to the Likelihood-ratio test of theta=0

Table 5: Lognormal regression with ancillary parameters in the baseline hazard function

	1)	2)	3)	4)
Ur_0	-.038***	-.024**	-.183***	-.200***
	(.008)	(.009)	(.024)	(.036)
Ur_t	.095***	.095***	.098***	.119***
	(.011)	(.011)	(.013)	(.034)
Ancillary Ur_0	-.031***	-.031***	-.025***	-.025***
	(.005)	(.006)	(.006)	(.007)
Ancillary Ur_t	.036***	.036***	.030***	.030***
	(.005)	(.006)	(.006)	(.007)
Region of birth dummies	No	Yes	Yes	Yes
$Region_0$ dummies	No	No	Yes	Yes
$Region_t$ dummies	No	No	No	Yes
N. individuals	20575	20575	20575	20575
N. records	85071	85071	85071	85071
Loglikelihood	-11300.086	-11255.369	-11201.853	-11191-953

Standard errors clustered at individual levels (in parentheses).

*** significant at 1%, ** significant at 5%.

Table 6: Results from the joint estimation of lognormal regression and unemployment rate at entry.

	1)	2)	3)	4)
Equation 1: Lognormal regression				
Ur_0	-.063***	-.185***	-.243***	-.283***
	(.014)	(.073)	(.072)	(.081)
Ur_t	.094***	.093***	.091***	.124***
	(.014)	(.014)	(.014)	(.033)
Region of birth dummies	No	Yes	Yes	Yes
Region_0 dummies	No	No	Yes	Yes
Region_tdummies	No	No	No	Yes
Equation 2: $\mathrm{Ur}_0 = \alpha \mathrm{Ur}_{14} + Z\gamma + \eta$				
Ur_{14}	-.231***	-.230***	-.230***	-.230***
	(.028)	(.028)	(.028)	(.028)
ρ	.055***	.277**	.099	.123
	(.014)	(.118)	(.133)	(.133)
Region of birth dummies	Yes	Yes	Yes	Yes
Individuals	20575	20575	20575	20575
Observations	85071	85071	85071	85071
Log pseudolikelihood	-274854.69	-274813.52	-274759.86	-274748.7

Standard errors clustered at individual levels (in parentheses). *** significant at 1%, ** significant at 5%. ρ is the correlation between the two normally distributed error terms ε and η. See technical appendix for details.

ur_{14} is the unemployment rate in the region of birth whan the individual was 14 y.o.

Table 7: OLS and IV regression of log(wage) on current unemployment rate and unemployment rate at entry.

	1) OLS	2) OLS	3) OLS	4) IV*
Ur	-.005***	-.007***	-.007***	-.027
	(.001)	(.001)	(.001)	(.030)
Ur_0	.009***	.005***	.009***	.036
	(.002)	(.001)	(.003)	(.042)
$Ur_0 *$Experience	-.002***	-.002***	-.002***	-.004***
	(.000)	(.000)	(.000)	(.001)
Experience	.100***	.056***	.053***	.200**
	(.003)	(.003)	(.003)	(.078)
Experience squared	-.004***	-.004***	-.004***	-.018***
	(.000)	(.000)	(.000)	(.006)
$Year_t$ dummies	Yes	Yes	Yes	Yes
$Region_0$ dummies	Yes	No	Yes	No
$Year_0$ dummies	Yes	No	Yes	No
region of birth dummies	No	Yes	Yes	Yes
year of birth dummies	No	Yes	Yes	Yes
Individuals	20575	20575	20575	20575
Observations	87618	87618	87618	87618

S.E. clustered at the level of groups defined by year and region at entry

Age, age^2, ur_{14} and ageur_{14} are the excluded intruments.

Table 8: Summary results for first stage regressions:

Variable	F(4,189)	P-value
Ur_0	17.68	0.0000
$Ur_0 *$Experience	22.31	0.0000
Experience	67.30	0.0000
Experience squared	156.91	0.0000

F-test (cluster-robust) of the excluded instruments

in the corresponding first-stage regressions.

7 Data Appendix

7.1 Structure of the dataset

Table A.1 displays an extract of the dataset. The first column identifies the individual. Because we include time-varying covariates, that is current unemployment rate, we have split each employment spell to obtain year-specific records. The second and third columns indicate the end and the beginning of each time span. The numbers in these two columns are dates recorded as integers representing the number of days from January 1, 1960. Thus 14137 represents September 15, 1998. The fourth column displays the year that contains the records . The columns labeled t0 and t1 define the time span in days for each record. Each record starts at t0 and end at t1. These variables define the time span used in the loglikelihood specification.

For example, individual 881 has been employed continuously since September 15, 1998 until Decembre 31, 2003. This employment spell has been divided into six year-specific records. The number in bold indicates when the beginning and the end of the year-specific record coincide with the beginning and the end of the year. Individual 881 has never been injured, as the fifth column indicates. Thus, this individual is censored at the end of 2003.

As another example, individual 1808 has been continuously employed from 13163 until 14106. He is not observed from 14106 to 14167 and from 14593 to 14868. Note that because the individual is not observed for a short period in 1998, we have two records in 1998. This individual experiences an injury on 14873 (September 20, 2000) and if we do not implement multiple failures he exits from the analysis.

Table A.1: Structure of the dataset

individuals	Working spell		year	injury	t0	t1	initial	current
	spell start	spell stop					unemployement	unemployment
[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]
881	14137	**14244**	1998	0	0	107	5.8	5.8
881	**14244**	**14609**	1999	0	107	472	5.8	4.7
881	**14609**	**14975**	2000	0	472	838	5.8	4.4
881	**14975**	**15340**	2001	0	838	1203	5.8	3.7
881	**15340**	**15705**	2002	0	1203	1568	5.8	3.8
881	**15705**	**16070**	2003	0	1568	1933	5.8	3.6
[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]
1808	13163	**13514**	1996	0	0	351	6.2	6.2
1808	**13514**	**13879**	1997	0	351	716	6.2	5.9
1808	**13879**	14106	1998	0	716	943	6.2	5.8
1808	14167	**14244**	1998	0	1004	1081	6.2	5
1808	**14244**	14593	1999	0	1081	1430	6.2	4.7
1808	14868	14873	2000	1	1705	1710	6.2	3.7
[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]

7.2 Log likelihood specification

7.2.1 Episode splitting and interval truncation

In our preferred specification we use a lognormal hazard function. As shown in Table A.1 we have split the episodes in year-specific records to incorporate time-varying covariates. The contribution of each record to the loglikelihood function for individual i is as follows:

$$LogL_i = c_i \log[\theta(T_i)] + \log[S(T_i)]$$

where $c_i = 1$ if the spell ends with an injury and 0 if the individual i is censored. $\theta(T_i)$ is the hazard rate and is equal to $f(T_i)/S(T_i)$. Given the structure of our dataset we consider the individuals to be at risk of injury only during episodes of employment. Since some individuals experience periods of unemployment, we have interval truncation. More precisely, we observe each worker at the time of first entry; thus, we know the spell start date and the time at which the individual is first at risk. However, if a spell of unemployment occurs, we observe subsequent year specific records with delayed entry. For example, if individual i is employed for t_1 days in his first job but experiences unemployment between t_1 and t_2, the hazard function and survivor function computed at the first year-specific record at reemployment should be conditional on the survival function computed at time t_2.

Table A.2 summarizes the contributions to the loglikelihood of each year-specific record for individuals, censored and not censored, with an without interval truncations.

Table A.2: Loglikelihood contributions

Individual	Record	Year	Censoring Indicator	Survival Time	Entry Time	Time Var. Covariates	Contribution to the log likelihood
Multiple data records after episode splitting in year specific records and no interval truncation							
1	1	1994	0	t_1	0	ur_{1994}	$log(S(t_1))$
1	2	1995	0	t_2	t_1	ur_{1995}	$log(S(t_2)/S(t_1))$
1	3	[..]	[..]	[..]	[..]	[..]	[..]
2	1	1994	0	t_1	0	ur_{1994}	$log(S(t_1))$
2	2	1995	1	t_2	t_1	ur_{1995}	$log(f(t_2)/S(t_1))$
Multiple data records after episode splitting in year specific records and interval truncation							
3	1	1994	0	t_1	0	ur_{1994}	$log(S(t_1))$
3	2	1995	0	t_3	t_2	ur_{1995}	$log(S(t_3)/S(t_2))$
3	3	[..]	[..]	[..]	[..]	[..]	[..]
4	1	1994	0	t_1	0	ur_{1994}	$log(S(t_1))$
4	2	1995	1	t_3	t_2	ur_{1995}	$log(f(t_3)/S(t_2))$

7.2.2 Selection analysis

To deal with the likely endogeneity of initial unemployment level we take advantage of the lognormal nature of the hazard function. We consider the following equation:

$$ur_{i,0} = \alpha ur_{i,14} + Z_i\gamma + \eta_i \tag{A.1}$$

where the initial unemployment level $ur_{i,0}$ is a function of the unemployment rate at 14, some predetermined variables Z_i and a normal error term η_i. We assume that the error term η_i and error term ε_i, defining the hazard distribution, follow a bivariate normal distribution. We thus exploit the properties of le lognormal-normal distribution to define the loglikelihood function. The likelihood contribution of individual i becomes:

$$\log L_i = \sum_{j=1}^{n_i} \left(c_{ij} \log[\theta_{\varepsilon|\eta}(T_{ij})] + \log[S_{\varepsilon|\eta}(T_{ij})] \right) \frac{1}{\sigma_\eta} \phi_\eta \left(ur_{0i} - \alpha ur_{i,14} - Z_i\gamma \right)$$

where n_i represent the number of records observed for individual i . ϕ_η is density function of the error term η in the equation *A.1.* The hazard function $\theta(T_{ij})$ conditional on η is defined as:

$$\theta_{\varepsilon|\eta}(T_{ij}) = \frac{\frac{1}{T_{ij} \cdot \sigma_\varepsilon} \phi_{\varepsilon|\eta} \left(\frac{\ln T_{ij} - X_{ij}\beta - \rho \frac{\sigma_\varepsilon}{\sigma_\eta}(ur_{0i} - \alpha ur_{i,14} - Z_i\gamma)}{\sigma_\varepsilon \sqrt{1-\rho^2}} \right)}{\left(1 - \Phi_{\varepsilon|\eta} \left(\frac{\ln T_{ij} - X_{ij}\beta - \rho \frac{\sigma_\varepsilon}{\sigma_\eta}(ur_{0i} - \alpha ur_{i,14} - Z_i\gamma)}{\sigma_\varepsilon \sqrt{1-\rho^2}} \right) \right)}$$

where the denominator $(1 - \Phi_{\varepsilon|\eta}(.))$ defines the survivor function $S_{\varepsilon|\eta}(T_{ij})$. Both the hazard and survivor function, in case of interval truncation, are divided by the survivor function computed at the time of reemployment as explained in section *6.1.* Analytical gradients are used to compute clustered standard errors and are available upon request.