Past unemployment influences on the incidence of Belgian employees unemployment

ERIC CAHUZAC

Abstract

In this paper we are mainly interested in the question of whether spells of unemployment lead to subsequent unemployment spells. Unemployment probabilities are evaluated on a panel of Belgian white-collar workers, controlling the heterogeneity by an individual specific effect. Therefore, a fixed effect conditional logit model is implemented to partly avoid spurious state dependence. First, the paper examines the first-order persistence in unemployment over waves of four or five years long. Thereafter, we examine whether individual behaviour facing unemployment, follows a Markov process. Results suggest that, given the white-collar present status, his past history in unemployment is not informative about future unemployment. However, the first order persistence in unemployment is partly due to data collection procedure and unmeasured personal characteristics.

1. Introduction

As economic theory often suggests, the relationship between past and future spells of unemployment, and their duration, is tenuous. Several reasons can explain this relationship. Based on human capital theory, if we assume that productivity is decreasing as time spent unemployed, a worker in this situation, reduces his chances of finding a new job. One may expect as well that, for homogeneous workers, specific human capital increases with tenure. Clearly, if employers are not constrained, they shed last-in workers and those having many spells of unemployment. In addition, because unemployment reduces workers

1 The author is grateful to the National Fund for Scientific Research (Belgium, FRFC convention n° 2.4529.90) for financial support and to Bruno Van der Linden for helpful discussions. Comments received from the participants of the Summer School GDR CNRS 995 (Paris 1994) are also gratefully acknowledged.

2 INRA, Economie et Sociologie Rurales (ESR).
requirements, the larger the number of previous spells of unemployment, the more precarious the job an individual accepts. Moreover, precarious jobs may not be rewarding and could therefore create a bad signal for future jobs. So, it is undoubtedly true that reoccurrence of unemployment are mutually due to repeated spells of unemployment and the individual’s or job’s characteristics. In this paper we are particularly concerned in the ways in which an unemployment spell during a year affects future unemployment spells. This kind of dependence, under consideration is labelled in the literature as a structural state dependence or a true state dependence (Heckman 1978).

It must be first recognised that past and future spells of unemployment can also be related simply because some individual or environmental variables are missing. For example, the reservation wage, the level of education, employment conditions may be correlated in time. Leaving out these variables may create serial correlation. In this case, the correlation is due to unobserved individual heterogeneity or to the auto correlation of some exogenous missing variables. Second, past and current unemployment might also be associated, because of the way the information on unemployment is collected. For example the question “have you ever been unemployed during the previous year?”, will result in a correlation between past and future experiences of unemployment even if there are no causal links between them. We would observe some persistence simply because a same unemployment spell may span two years. This spurious state dependence can be avoided if chronological data are available and different spells of unemployment are distinguished.

Following Heckman (1978), Chamberlain’s works was epoch-making, due to his introduction of a new method to investigate the structural state dependence problem with limited dependent variables in the framework of panel data. In his seminal paper (Chamberlain 1985) he presents a survey of the new methods used in the estimation of models for longitudinal data. More particularly, he discusses the heterogeneity treatment problem with reference of panel models and introduce conditional likelihood functions. This paper is a sequel to an earlier one, (Chamberlain 1980). Following Chamberlain’s works some empirical studies appear. Corcoran (1982), use first Chamberlain’s autoregressive logistic model to investigate the extent to which a woman’s work history influences her current work behaviour. She uses a subsample of the National Longitudinal Survey (NLS) of women between 14 and 24 years old who where interviewed an-
nually from 1968 to 1973. Shortly afterwards, Corcoran and Hill (1985) use the same approach to know whether unemployment at one period leads to later unemployment for adult men (35-64 years old) household heads and labour force participants, taken from the Panel Study of Income Dynamics (PSID) over the 1968-1977 period. Narendranathan and Elias (1993), investigate a similar unemployment behaviour with the same methodology using a birth cohort panel survey – the National Child Development Study (NCDS) – a monthly record longitudinal study of all persons born in Great Britain in one week in March 1958. Finally, in a recent paper, Magnac (1995) highlights the heterogeneity problem studying the transition from school to work. He uses a subsample of the 1990-1992 French Labour Force panel Survey of young between 18 and 29 years old.

As discussed in detail in the third section, we use in this paper longitudinal administrative data on a representative set of Belgian white collar workers to investigate unemployment behaviour. In a first approach we analyse unemployment persistence problem through a first order autoregressive model. In the sequel we ask whether or not the whole unemployment history helps us to predict future unemployment spells. The two approaches take unobserved personal characteristics into account. Few will dispute the claim that individuals precariousness on the labour market and their come-back to unemployment is connected to numerous factors as age, education, firm activity, location of the job, or dimension of the firm. Data at hand only permit us to neutralise the age and sex effects. Other variables are included in the heterogeneity term in view to capture all time-invariant characteristics effects which could influence men’s propensities for unemployment. With respect towards above-mentioned empirical studies and the references therein, we consider a wider period of observation (nine years). Dividing the observation period into several short spells, the well known time-invariant hypothesis of unemployment propensity is then less constraining.

The remainder of the paper is organised as follows. In the next Section, we present the model. In Section 3 we briefly present the data at hand and the choices made for a pertinent study of the persistence of unemployment. Section 4 deals with the econometric methods of estimation. In Section 5 we summarise our findings. For the purposes of comparison the model is implemented with two consecutive periods (1977-1980 and 1981-1984). Some concluding remarks are drawn in the last Section.
2. The model

To introduce the methodology used in this paper it is somewhat useful to consider first the autoregressive and serial correlation models in a usual dynamic regression framework with reference to continuous data. Hence, a state dependence model tests the significance of the coefficient $\gamma$ in

(i) \[ y_t = \alpha + \gamma y_{t-1} + \epsilon_t \]

where depict the heterogeneity term and the $\epsilon_t$ are i.i.d. N(0, 1). While a serial correlated model is specified as

(ii) \[ y_t = \alpha + u_t \]

with $u_t = \rho u_{t-1} + \epsilon_t$ and where the $\epsilon_t$ are i.i.d. N(0, 1). Testing serial correlation come down testing the significance of $\rho$. In Model (i), we have

\[ P(y_t = 1|\alpha, y_{t-1}, y_{t-2}, \ldots) = P(y_t = 1|\alpha, y_{t-1}) \]

while in Model (ii), the value of $y_t$ depends on the entire history and not only on the preceding state (like in Model (i)). Thus the lagged Model (i) implies a first order Markov chain, whereas the residual first order Markov process does not imply a first order Markov chain for $y_t$ (Chamberlain 1984). The distinction between state dependence and serial correlation cannot be satisfactory made on the autoregression order in Model (i). As in continuous models, a key distinction between these two models is whether there is a dynamic response to a change in the exogenous variables. We cannot proceed much further (to dissociate pure heterogeneity from state dependence or from serial correlation) unless we impose more structure in Model (i). Without this structure, we will show in this paper how to take into account data sampling specifications and individual heterogeneity to reduce spurious state dependence.

With limited dependent variable, the binary variable $y_{it}$ records whether individual $i$ experiences unemployment during year $t$ ($y_{it} = 1$) or not ($y_{it} = 0$). In a general point of view the incidence of unemployment might be regarded as a continuous latent variable $y^*_{it}$, like

\[ y^*_{it} = \alpha_i + \beta' x_{it} + \gamma' z_i + \lambda y_{it-1} + u_{it} \text{ where } i = 1, \ldots, N \quad \text{and } t = 1, \ldots, T \]
where $\alpha_i$ depict a constant individual heterogeneity term, $x_{it}$ denote a vector of individual and environmental exogenous characteristics, and $z_i$ are time-invariant exogenous observable personal characteristics, such as educational indicators or social level. The incidence of unemployment is modelled by means of a first-order autoregressive process where $\lambda$ captures unemployment persistence of order one. In addition, we estimate the probability that $y_{it}$ equals 1 or 0, assuming

$$y_{it} = \begin{cases} 1 \text{ if } y^*_t \geq 0 \\ 0 \text{ otherwise} \end{cases}$$

(2)

Now, we are looking for consistent estimates of the parameters $\beta$, $\gamma$ and $\lambda$. If the $u_i$ are independently and identically distributed and follow a logistic distribution, the contribution to the likelihood function for an unemployed individual at any time $t$, conditionally on the observable, is

$$P(y_{it} = 1) = P(y^*_t \geq 0) = \frac{\exp(\alpha_i + \beta' x_{it} + \gamma' z_i + \lambda y_{it-1})}{1 + \exp(\alpha_i + \beta' x_{it} + \gamma' z_i + \lambda y_{it-1})}$$

(3)

We have to estimate $N+3$ parameters, $N$ specific effects ($\alpha_i$, $i = 1, \ldots, N$), and 3 structural parameters ($\beta$, $\gamma$, and $\lambda$). According to Neyman and Scott (1948), when the number of parameters to estimate increases with the number of observations, the Maximum Likelihood Estimator (MLE) for fixed $T$ loses his properties. From this it follows that the structural parameters estimators are no more consistent and we come up against an incidental parameter problem.

The key to obtain consistent estimates of structural parameters in attendance of incidental parameters was first presented by Andersen (1970). If sufficient statistics$^3$ $T(y)$ exist for the $\alpha_i$ 's, and if they are independent of the structural parameters, the likelihood conditional to these statistics does not depend upon the incidental parameters $\alpha_i$. Then, an alternative means to estimate the joint likelihood with an increasing number of parameters, is to avoid incidental parameters by estimating the conditional likelihood. This procedure gives consistent Conditional Maximum Likelihood Estimates (CMLE)$^4$. This approach was used by Chamberlain (1980, 1985) to implement an autoregressive logit model for the study of unemployment persistence. The main advantage of CMLE lies

---

$^3$ $T(y) : \PsiIR^n$, is called sufficient statistic for $\theta$ if the conditional distribution $f(y \mid T(y))$ is independent from $\theta$.

$^4$ A comprehensive treatment of the information matrix related to these models may be found in Cahuzac (1995).
in the fact that it is very easily carried out. But the model, as specified in Equation (3), cannot be estimated without some restrictions. It should be stressed that consistent estimates of $\lambda$ can only be obtained if $\beta = 0$ (Chamberlain, 1980). As far as our concern, this is not an unrealistic assumption because the sample could be shared according to observable and time invariant individual characteristics.

An alternative approach to avoid such a tedious incidental parameter problem is to consider each $\alpha_i$ as a random variable with a specific distribution whose parameters as mean and variance must be estimated. In a parametric approach we must specify a distribution for $\alpha_i$ conditional on the explanatory variables. Consistent estimators of structural parameters can be drawn by way of Marginal Likelihood Estimation (MLE) where it may be conceded that errors terms are normally distributed. Nothing unfortunately ensure that the $\alpha_i$’s nor the $u_{it}$’s are normally distributed. This technique is particularly interesting in that it enables us to take into account observable individual characteristics and environmental time-dependent characteristics. Relevant models are presented by Heckman and Willis (1976), Avery, Hansen and Hotz (1983) and Chamberlain (1980, 1985).

Although CMLE imposes some restrictions under the structural parameters, it does not require any hypothesis under the distribution of $\alpha$. In addition, in our case, the introduction of was motivated by missing variables that are correlated with individual or environmental characteristics. From all of this, it follow that as a first step, we will estimate a fixed effects logit model.

3. The data

The panel data at hand are collected for administrative purpose, specifically in order to compute retirement rights. We observe for a 9 years period all the non-retired Belgian people covered by the general pension scheme (three quarters of the active population). It should be stressed that the way the data are collected only allows us to know if the observed worker has been unemployed a given year. In fact, events (unemployment) can occur at any point in time, but our data only record the particular interval of time (year) in which each event occurs (interval sampling). For technical reasons we concentrate our analysis on a specific cohort,
the French-speaking generation born in 1951. By so-doing the analysed population is about 42,000 individuals who are 26 years old the first year of observation.

Similarly, it seems necessary to focus on an homogeneous population, the white-collars, to decrease the amount of unemployed for economic reasons. Indeed, a distinctive feature of the labour market in Belgium is that blue-collars workers can experience unemployment for economic, climatic or technical reasons. This notion is quite different from the one of temporary layoff since the contract of partially unemployed people is not interrupted. In Belgium the employer can suspend the contract partially or completely in case of economic difficulties (e.g. a drop in demand), but we cannot isolate this phenomenon in our data bank. Ultimately, this kind of unemployment is so different from the usual definition, that we can not consider partial unemployment in our analysis. This reduce our sample close to 60% with 7,375 men and 9,918 women.

When we speak about “true” and “spurious” state dependence, we have in mind that spurious state dependence is due to a lack of information but also to a sampling problem. Our interval sampling data are obviously not appropriated to measure the average duration of a completed unemployment spell. However, a similar data base (with more distinction), from a census of Belgian unemployed between 25 and 50 years old, gives us the average duration, in year, of a completed spell of unemployment starting year t (see Table 1).

**TABLE 1**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>0.81</td>
<td>0.83</td>
<td>0.9</td>
<td>0.93</td>
<td>1.35</td>
<td>1.6</td>
<td>1.57</td>
<td>1.54</td>
<td>1.34</td>
</tr>
<tr>
<td>Women</td>
<td>2.38</td>
<td>2.27</td>
<td>2.02</td>
<td>2.5</td>
<td>2.61</td>
<td>2.59</td>
<td>2.49</td>
<td>2.51</td>
<td>2.43</td>
</tr>
</tbody>
</table>

---

5 For a complete description of the data, see Cahuzac et al (1994).
6 Here, the distinction white-collars blue-collars is rough, it just allows to distinguish between manual workers from intellectual workers.
7 The employer is allowed to shut down its operation during at the most 4 weeks or to reduce the working time during at the most 3 months. During that time the worker is allowed to receive unemployment compensation. After 4 weeks (resp. 3 months) a normal week of work is necessary before a new period of partial unemployment can take place.
8 However, during the period we cover, Belgian national statistics reveal that between 5 and 7 percent of the blue-collar workers were partially unemployed.
9 As far as the period of observation, the age and place of residence are concerned. Duration are complete, that is there is no left nor right censoring.
Table 1 is not as satisfactory as one would expect. Indeed, the results of this Table do not tell apart the status of occupation, the level of education or the activity sector. Nevertheless, Table 1 provides us with a useful information on unemployment duration. Undoubtedly, spells of unemployment are generally greater than 2 years for women and the propensity is increasing for both sex after the 1980 crisis. So paradoxical thought it may seem, even if women’s past and current unemployment behaviour are not related, we could expect the opposite simply because an unemployment spell may span two consecutive years. This dependence is clearly not structural, but only due to a sampling problem. Because this phenomenon is less important for men, we concentrate the analysis on the population of white-collar men. Therefore, we can expect that the unemployment persistence study carried out here, will be meaningful.\footnote{The spurious state dependence problem is not largely resolved by this way but it is reduced comparing to the one for the women.}

4. Estimation method

In this Section we answer two questions. First, whether or not unemployment in a period t-1 is relevant to explain unemployment in period t. Second, whether or not the whole unemployment history helps to predict future chances of unemployment. Because of the limited number of personal characteristics in the data bank, we can easily assume that $\beta$ equals 0 in Equation (1) and estimate a first order autoregressive logit model as presented in Chamberlain (1985). The above-mentioned conditioning with respect to a set of sufficient statistics for the incidental parameters $\alpha_i$ also eliminates the time invariant observable characteristics. To take into account these later variables, the sample must be stratified across the main $z_i$ variables. Under this hypothesis, Equation (3) became

$$P(y_{it} = 1) = \frac{\exp(\alpha_i + \lambda y_{it-1})}{1 + \exp(\alpha_i + \lambda y_{it-1})}$$ (4)

It has been said earlier that the joint likelihood maximisation with respect to the parameters $\alpha_i$ and $\lambda$ gives inconsistent estimates due to the incidental parameter problem. Consistent estimates of can be obtained by estimating the conditional likelihood (Andersen 1970, Chamberlain 1980, 1985). This bring us to suppose that the number of years a man is unemployed over the period $\sum_{t=1}^{T} y_{it}$
and his unemployment status for the final year of observation \(y_{iT}\) are sufficient statistics for the incidental parameters \(\alpha_i\). In addition, initial conditions are dealt with by conditioning a man’s unemployment status in the first full year \(y_{i1}\). Given these assumptions, we can write the conditional probability that an individual is unemployed a given year as

\[
P(y_{i1}, \ldots, y_{iT} \mid y_{i1}, \sum_{t=1}^{T} y_{it}, y_{iT}) = \frac{\exp\left(\lambda \sum_{t=2}^{T} y_{it}y_{i,t-1}\right)}{\sum_{d \in B_i} \exp\left(\lambda \sum_{t=2}^{T} d_t d_{t-1}\right)}
\]

where \(B_i = \{d = (d_1, \ldots, d_T) \mid d_t = 0 \text{ or } 1, \sum_{t=1}^{T} d_t = \sum_{t=1}^{T} y_{it}, d_T = y_{iT}\}\) represents the set of all the possible realisations of the vector \((y_{i1}, \ldots, y_{iT})\) for a given \(i\). Hence, the conditional log-likelihood can be written as

\[
L(\lambda \mid y, \tau) = \sum_{i=1}^{n} \ln \left\{ \frac{\exp(\lambda \sum_{t=2}^{T} y_{it}y_{i,t-1})}{\sum_{d \in B_i} \exp(\lambda \sum_{t=2}^{T} d_t d_{t-1})} \right\}
\]

with \(\tau = \left\{ y_{i1}, \sum_{t=1}^{T} y_{it}, y_{iT}\right\}\). Since the likelihood depends on \(y_{i1}, \sum_{t=1}^{T} y_{it}\) and \(y_{iT}\), an observation period greater than 4 years \((T \geq 4)\) is necessary to estimate \(\lambda\). With \(T = 9\) (our case) there exist \(2^9\) possible probabilities. But some of them do not contribute to the conditional likelihood because they are just equal to a constant. Accordingly, some probabilities are independent of \(\lambda\) and will not be estimated. Consequently, only a subset of our sample will be necessary for the estimation of \(\lambda\).

Note that this model is not without drawback. Particularly, \(\alpha\) and \(\lambda\) are time invariant. Related to \(\alpha\), this means that the non observable individual characteristics that enter in the conditional probability of getting a new job are constant over the nine years. For \(\lambda\), the time invariant hypothesis means that the effect of

---

11 A demonstration is given in Appendix A.
12 See Appendix B for details.
13 Technical details are provided in Appendix B.
an unemployment spell on the probability of the following spell is the same during the observation period for all individuals. Few will dispute the claim that this hypothesis seems unrealistic. Therefore in the next Section we decide to split the sample into two sub samples periods corresponding to the 1977 to 1980 and 1981 to 1984. This approach is not completely satisfactory but the real question at issue is that the model does not integrate the $x_{it}$ variables. From the data at hand we are able to soften time invariant assumption, but we are not able to get rid completely of the serial correlation problem.

It remains to see now whether, given the current unemployment status, the whole unemployment history helps to predict future chances of unemployment or not. Equivalently, does the behaviour of an individual facing unemployment The non Markovian hypothesis (the state dependence hypothesis, according to Chamberlain, 1985) will be tested below. Rejection of this hypothesis involves independent and exponentially distributed duration of the spells of unemployment. When data record only whether there is a spell of unemployment during the year or not (interval sampling case), the corresponding binary sequence does not have the first-order Markov property. For example if $y_{it-1} = 1$ we only know that the individual was unemployed sometime in period $t-1$. Observing $y_{it-2}$ is relevant because it could affect the probability that he was unemployed early in the year $t-1$ rather than late in that year. A test of the Markov property can however be implemented because if $y_{it-1} = 0$ then the individual was never unemployed the preceding year, therefore $y_{it-2}, y_{it-3}$, would be irrelevant if the unemployment process is a Poisson process. The conditional probability of unemployment at time $t$ is given by:

$$P(y_{it} = 1|y_{it-1}, y_{it-2}, \ldots) = P(y_{it} = 1|y_{it-1} = y_{it-2} = \ldots = y_{it-J} = 1, y_{it-J-1} = 0)$$

$$= P(y_{it} = 1 | J)$$

where $J$ is the number of consecutive preceding years where the individual was unemployed. Hence, if the unemployment process is a Poisson process and if there is no state dependence the probability to be unemployed at time $t$ only depends on the number of consecutive years an individual was unemployed preceding $t$. A Logistic parameterisation allows us to write this probability as:

$$P(y_{it} = 1|y_{it-1}, y_{it-2}, \ldots) = \frac{\exp(A_i)}{1 + \exp(A_i)}$$

where $A_i = \alpha_i + \sum_{k=1}^{\infty} \psi_{ik} \Pi_{j=1}^{k} y_{it-j}$
Here, \( \alpha_i \) and \( \psi_{ik} \) are the individual set of parameters for the Poisson distribution of time spent in unemployment. Writing \( A_i \) it should be stressed that past unemployment occurs until the first spell of employment is reached. This strong hypothesis simplify the model unlike a model where unemployment history occurs through any function of the past. However, a realistic assumption is that weights \( \psi_{ik} \) fluctuate according to unemployment duration. That is, testing state dependence amounts to give to each individual his own first order Markov chain \( (A_i) \) and test the magnitude and significance of \( \lambda_2 \) in

\[
P(y_{it} = 1 | y_{it-1}, y_{it-2}, \ldots) = \frac{\exp(A_i + \lambda_2 y_{it-2})}{1 + \exp(A_i + \lambda_2 y_{it-2})}
\]

As in the simple case (Equation (4)) we are now facing up to an incidental parameter problem and we must find sufficient statistics for \( \alpha_i \) and \( \psi_{ik} \) to obtain consistent estimates of \( \lambda_2 \). Using the same argument as in Appendix A it is easy to show that sufficient statistics for the parameters \( \alpha_i \) and \( \psi_{ik} \) are \( S_{i01}, S_{i011}, S_{i0111}, \ldots \), where \( S_{i0111} \) for example is the number of times that \( y_{it} = 1 \) is preceded by the sequence \( (0, 1, 1) \). In addition, we must condition on \( y_{i1} \) and \( y_{iT-1} \) and on \( n_{i1} \) (resp. \( n_{iT} \)) the number of consecutive ones at the beginning (resp. end) of the period. Finally, \( n_{i1} \) (resp. \( n_{iT} \)) gives us information on \( y_{i2} \) (resp. \( y_{iT} \)). We can write now the conditional likelihood as described in Appendix B

\[
L = (\lambda | y, \tau) = \sum_{i=1}^{N} \ln \left( \frac{\exp(\lambda_2 \sum_{t=N_{i1} + 2}^{T} y_{it} y_{it-2})}{\sum_{d \in B_i} \exp(\lambda_2 \sum_{t=N_{d1} + 2}^{T} d_i d_{i-2})} \right)
\]

where \( \tau = \{S_{i01}, S_{i011}, \ldots \} \) and \( B_i = \{d=(d_1, \ldots, d_T); d_i = 0 \text{ or } 1, n_{d1} = n_{i1}, n_{dT} = n_{iT}, d_2 = y_{i2}, d_{T-1} = y_{iT-1}, S_{d01} = S_{i011}, \ldots \} \)

Given \( n_{i1}, n_{iT}, y_{i2} \) and \( y_{iT-1} \), we need at least 6 waves of observations \( (T \geq 6) \) to obtain probabilities depending on \( \lambda_2 \). Here we have to estimate the magnitude and significance of \( \hat{\lambda}_2 \). If it is not significantly different from zero, we say that given the individual present unemployment status, prior unemployment history is not informative about its future unemployment. In other case we can say that there is duration dependence.
5. Empirical findings

As mentioned earlier, we study the unemployment persistence on a sample of white-collar males (7,375 individuals). First of all let us consider this persistence on two consecutive periods of four years 1977 to 1980 and 1981 to 1984. The Chamberlain’s technique presented in the preceding Section is now used to estimate the coefficient of Equation (4). We need to concern ourselves here with all possible sequences of the binary variable $y_{it}$, depending on $\lambda$, over the four years. In case of $T = 4$ we have $2^4$ probabilities but few of them depend on $\lambda$. They are summarised in the following box.

<table>
<thead>
<tr>
<th>$y_1, \ldots, y_4$</th>
<th>$\sum y_{it}$</th>
<th>likelihood contributions</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_1 = 0$, $y_4 = 1$, $\sum y_i = 2$</td>
<td>0011</td>
<td>$\exp(\lambda) / 1 + \exp(\lambda)$</td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>0101</td>
<td>$1 / 1 + \exp(\lambda)$</td>
<td>(2)</td>
</tr>
<tr>
<td>$y_1 = 1$, $y_4 = 0$, $\sum y_i = 2$</td>
<td>1100</td>
<td>$\exp(\lambda) / 1 + \exp(\lambda)$</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>1010</td>
<td>$1 / 1 + \exp(\lambda)$</td>
<td>(4)</td>
</tr>
</tbody>
</table>

Only two sets of probabilities (four conditional probabilities) allow to estimate $\lambda$ over the sixteen possible ones. Therefore, we are working on a subsample of white-collars i.e. 196 individuals for the (1977-1980) period and 141 individuals for the (1981-1984) period. If $n_i$ ($i = 1, \ldots, 4$) individuals contribute to the likelihood of Type (i) (in the above box) then $T_i / n_i$ gives us an unbiased estimator of $\theta_i = \exp(\lambda)$ with a standard error of $n / n_1 n_2$. In the same way, $T_2 = n_3 / n_4$ is also an unbiased estimator of $\theta_2 = \exp(\lambda)$ with a standard error of $n / n_3 n_4$. Hence the weighted sum of the $T_i$ is still an unbiased estimator of $\exp(\lambda)$ defined as

$$T^* = \frac{(n_1 + n_2) \cdot T_1 + (n_3 + n_4) \cdot T_2}{n_1 + n_2 + n_3 + n_4}$$

For the 1977-1980 period, the Chamberlain’s autoregressive logistic model gives us an estimation of equal to 2.35 (see Table 2) with a standard error of
0.02, while for the (1981-1984) period, the same model gives a $\hat{\lambda}$ equal to 2.54 with a standard error of 0.03.

**TABLE 2**

<table>
<thead>
<tr>
<th>Years</th>
<th>$\hat{\lambda}$</th>
<th>standard error</th>
<th>$e^{\hat{\lambda}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977-1980</td>
<td>2.35</td>
<td>0.02</td>
<td>10.5</td>
</tr>
<tr>
<td>1981-1984</td>
<td>2.54</td>
<td>0.03</td>
<td>12.7</td>
</tr>
</tbody>
</table>

In both cases the estimates are very significant. So, we can confidently conclude that for a white-collar, an unemployment spell a given year is a good predictor of his unemployment the following year. More precisely, it can be said that before 1980 an unemployed white-collar has 10.5 more chance to be unemployed the following year (see the last column in Table 2) if he was unemployed the preceding year than if he was not. After 1980 this ratio becomes 12.7. This result shows the growth unemployment risk after the 1980 crisis for an individual who knew unemployment the preceding year. This information was already present in Table 1. However, high probabilities ($e^{\hat{\lambda}}$) suggests that serial correlation takes a large part in this first order dependence.\(^{14}\)

For this reason $\lambda$ is estimated again over a period of two consecutive years. If the individual knows an unemployment spell somewhere during the two years, we consider that $y_{it} = 1$. In that way, the number of observations over the nine years is four ($T = 4$).\(^{15}\) But our subsample increases and we have now 278 white-collar workers contributing to the likelihood. Estimation of first order dependence in this case yields $\hat{\lambda}$ equals 2.1 with a standard error of 0.015. As expected a smaller coefficient is obtained since now, an unemployed white-collar has 8.2 ($e^{2.1}$) more chance to be unemployed if he was unemployed the preceding year than if he was not.

We also use Chamberlain's methodology to estimate the dependence over a five waves period ($T = 5$). The corresponding probabilities are presented in Appendix C. As for $T = 4$, if $n_i$ is the number of individuals who contribute to the likelihood $L_i$, we obtain 2 estimators of $\exp(\lambda)$. They are $T_1 = n_1/n_2$ and $T_2 = n_3/n_4$.\(^{14}\) See Chamberlain 1985 p14-15 for details.\(^{15}\) Last year out.
n₄. Over the 1977-1981 period the ratio gives a $\hat{\lambda}$ equal to 2.01 with a standard error of 0.017 (see Table 3), while over the 1981-1985 period $\lambda$ is equal to 2.58 with a standard error of 0.018.

**TABLE 3**

<table>
<thead>
<tr>
<th>Years</th>
<th>$\hat{\lambda}$</th>
<th>standard error</th>
<th>$e^{\lambda}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977-1981</td>
<td>2.01</td>
<td>0.017</td>
<td>7.5</td>
</tr>
<tr>
<td>1981-1985</td>
<td>2.58</td>
<td>0.018</td>
<td>13.2</td>
</tr>
</tbody>
</table>

These estimations are significantly different from zero and tell us again that past unemployment predicts future unemployment. However, in this case there is a fundamental difference between chances of getting a job prior to 1981 and after. As regards the 1977-1981 period, this probability is about a factor of two smaller than for the 1981-1985 period. More precisely, prior to 1981 an unemployed white-collar has 7.5 ($e^{2.01}$) more chance to be unemployed if he was unemployed the preceding year than if he was not. This same ratio is about 13 after 1981.

Comparing our results with those of Corcoran and Hill (1985), we come to similar conclusions\(^\text{16}\). That is, a man's unemployment in one year is a good predictor of his unemployment in the next year, even after adjustments for stationary heterogeneity. However, the way our data are collected seems to be the main reason of the large size of the coefficients, and so of the large association between current and just preceding unemployment. Finally, we can notice a high sensitivity of the results between $T = 4$ and $T = 5$, as already revealed by Corcoran and Hill.

We now deal with the second question. Given a man unemployment current status, is the unemployment history a good predictor or not for white-collar future chance of unemployment? It is really much the same as asking whether or not unemployment behaviour differs from a Markov behaviour. If we consider, like in Equation (6), that $A_i$ represents the unemployment weight, we can estimate $\lambda_2$. We need six (or more) waves to estimate $\lambda_2$, and we notice that only two conditional probabilities are able to estimate $\lambda_2$, they are

\(^{16}\) Nevertheless, for a five-year time span from 1972 to 1976 Corcoran and Hill find an estimate of adjacent year persistence in unemployment behaviour equal to 1.29.
P_1 = P(101000|101000 or 100100) = \frac{\exp(\lambda_2)}{1 + \exp(\lambda_2)}

P_2 = P(101000|101000 or 100100) = \frac{1}{1 + \exp(\lambda_2)}

An estimator of \exp(\lambda_2) will be obtained by the quotient n_1/n_2, where n_i is the number of individuals with pattern of unemployment P_i. The associated standard error of this estimation is (n_1+n_2)/n_1n_2. Results are presented in Table 4 for different periods of observation.

TABLE 4
Estimation of \( \hat{\lambda}_2 \)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\lambda}_2 )</td>
<td>0.34</td>
<td>0.91</td>
<td>0.34</td>
<td>0.69</td>
</tr>
<tr>
<td>Se</td>
<td>0.34</td>
<td>0.35</td>
<td>0.34</td>
<td>0.75</td>
</tr>
<tr>
<td>St</td>
<td>1</td>
<td>2.6</td>
<td>1</td>
<td>0.92</td>
</tr>
</tbody>
</table>

(where Se is the standard error of estimates and St the corresponding Student statistics)

We can notice that in most cases the estimate of \( \lambda_2 \) is still large but is not very significant (Se = 0.34) or non significant (Se = 0.75). A one percent significant estimation is only obtained for the 1978-1983 period, yet the estimations are based on small number of individuals. Therefore we can confidently conclude that, given the white-collar present status in unemployment, his past history in unemployment is not informative about future unemployment spells.

6. Summary and conclusions

In this paper the incidence of past unemployment spells on future employment conditions has been analysed. This exploratory analysis was conducted with a view to compare our results with preceding empirical studies. Most of the time these have been realised on American or English survey panel data. We wanted to see here if our data, not collected for economic but for administrative
purpose, are also able to model individual behaviour on the labour market in European countries. A logit autoregressive model was estimated on a white-collar men population, after adjustments for stationary heterogeneity. When compared to preceding empirical studies, we have at our disposal a larger sample both in number of individuals and in number of periods. This allows us to choose a statistically more relevant population. In addition the available number of waves enables us to soften the hypothesis that the effect of an unemployment spell on the probability of the following spell is the same during the observation period for all individuals.

Estimations on four or five years show that even after adjustment for stationary heterogeneity, dependence between preceding and current spell of unemployment is quite strong. The inescapable conclusion which emerges from our estimates is that for a white collar, unemployed the preceding year, the odds that the same man is unemployed during current year is between 7 to 13 times higher than if he was not unemployed last year. In addition this paper has shown that entire past history on unemployment is not an helpful predictor for future unemployment risks.  

These results are obtained by assuming several restrictions: (i) individual-specific characteristics that influence unemployment behaviour are supposed constant over the period; (ii) exogenous variables are not present in the model; (iii) the probability of transition is assumed to be time invariant. These hypothesis are due to the way we have modelled the unemployment dependence and may be relaxed in a more sophisticated model. The estimation of the Chamberlain first order logit autoregressive model is hence a first stage in a more complete study of unemployment dependence. By the way, a time trend can be added to capture effects of time variant variables. As mentioned in Corcoran (1982), allowing for a time trend reduces estimates of adjacent year persistence $\hat{\lambda}$ by about one-third. But this model first reduces to the usual conditional logit with no $\alpha_t$ and then no ensure that $P(y_{it} = 1)$ will necessarily lie in the range $(0,1)$. Another way is to substitute the fixed effect logit analysis by a random effect one. This last technique allows to take into account individuals characteristics and time variant variables but introduce other problems in the method of estimation and in the choice of the distribution for the random effect.

---

17 These results are similar to those observed by Corcoran and Hill (1985) over 1251 prime-age males between 35 and 65 years old in 1977 and studied over the 1967-1976 period.
Appendix A: Determining sufficient statistics

Let $\alpha_i$ ($i = 1, ..., N$) a set of incidental parameters. If $y_0$ is a sufficient statistic for $\alpha_i$ conditionally to other parameters, then the $y$'s conditional distribution to this sufficient statistic is independent to the $\alpha_i$:

$$P(y|y_0, \lambda) = \frac{P(y|\alpha_i, \lambda)}{P(y_0|\alpha_i, \lambda)}$$

(1A)

In other words, the $y_0$ so that $P(y|y_0, \lambda)$ is independent of the $\alpha_i$ is a sufficient statistic for the $\alpha_i$ conditionally to other parameters. From now, subscript $i$ is dropped from $y_{it}$ for notational convenience.

Let $y = (y_1, ..., y_T)$, for a given individual. Then with $\alpha_i$ an incidental parameter and $\lambda$ the structural parameter, the likelihood is

$$L(\alpha_i, \lambda|y) = P(y_1, ..., y_T|\alpha_i, \lambda) = \prod_t P(y_t|\alpha_i, \lambda)$$

$$= P(y_T|y_{T-1}), P(y_{T-1}|y_{T-2}), P(y_{T-2}|y_{T-3})...P(y_2|y_1). P(y_1)$$

$$= P(y_1). \prod_{t=2}^T P(y_t|y_{t-1})$$

As $y_i$ is a binary variable:

$$L(\alpha_i, \lambda|y) = P(y_1). \prod_{t=2}^T \left[ P(y_t = 1|y_{t-1})^{y_t}. P(y_t = 0|y_{t-1})^{1-y_t} \right]$$

$$= P(y_1) \prod_{t=2}^T \left[ \frac{\exp(\alpha_i + \lambda y_{t-1})^{y_t}}{1 + \exp(\alpha_i + \lambda y_{t-1})} \right]^{y_t} \cdot \left[ \frac{1}{1 + \exp(\alpha_i + \lambda y_{t-1})} \right]^{1-y_t}$$

$$= P(y_1) \prod_{t=2}^T \left[ \frac{\exp(\alpha_i + \lambda y_{t-1})^{y_t}}{1 + \exp(\alpha_i + \lambda y_{t-1})} \right]$$

$$= P(y_1) \frac{\exp \left[ \alpha_i \left( \sum_{t=2}^T y_t \right) + \left( \sum_{t=2}^T y_{t-1} \right) \right]}{\prod_{t=2}^T [1 + \exp(\alpha_i + \lambda y_{t-1})]}$$

The denominator, can be divided in number of times an observation, $y_t$ is preceded by $y_{t-1} = 0$ (m0) or $y_{t-1} = 1$ (m1) so that we write
(2A) \[ L(\alpha_i, \lambda|y) = P(y_1) \cdot \frac{\exp\left[\alpha_i \left(\sum_{t=2}^{T} y_t \right) + \lambda \left(\sum_{t=2}^{T} y_{t-1} y_{t-1}\right)\right]}{(1 + \exp(\alpha_i))^m_0 \cdot (1 + \exp(\alpha_i + \lambda))^{m_l}} \]

with \( m_l = \text{the number of times } y_t \text{ is preceded by } (y_{t-1} = 1); m_l = \sum_{t=2}^{T} y_{t-1} - y_T \)

and \( m_0 = \text{the number of times } y_t \text{ is preceded by } (y_{t-1} = 0); m_0 = (T-1) - m_l \)

In the same way, for a given \( y_0 \) we write

\[ L(\alpha_i, \lambda|y_0) = P(y_1) \cdot \frac{\exp\left[\alpha_i \left(\sum_{t=2}^{T} y_{0t} \right) + \lambda \left(\sum_{t=2}^{T} y_{0t} y_{0t-1}\right)\right]}{(1 + \exp(\alpha_i))^{m_0^*} \cdot (1 + \exp(\alpha_i + \lambda))^{m_l^*}} \]

with \( m_l^* = \sum_{t=1}^{T} y_{0t} y_{0t} \) and \( m_0^* = (T-1) - m_l^* \)

(1A) can now be written as:

\[ L(\lambda|y, y_0) = \frac{\exp\left[\alpha_i \left(\sum_{t=2}^{T} y_t \right) + \lambda \left(\sum_{t=2}^{T} y_{t-1} y_{t-1}\right)\right]}{(1 + \exp(\alpha_i))^m_0 \cdot (1 + \exp(\alpha_i + \lambda))^{m_l}} \times \frac{\exp\left[\alpha_i \left(\sum_{t=2}^{T} y_{0t} \right) + \lambda \left(\sum_{t=2}^{T} y_{0t} y_{0t-1}\right)\right]}{(1 + \exp(\alpha_i))^m_0^* \cdot (1 + \exp(\alpha_i + \lambda))^{m_l^*}} \]

(3A) = \[ \exp\left[\alpha_i \left(\sum_{t=2}^{T} y_t - \sum_{t=2}^{T} y_{0t}\right) + \lambda \left(\sum_{t=2}^{T} y_{t-1} - \sum_{t=2}^{T} y_{0t} y_{0t-1}\right)\right] \left[1 + \exp(\alpha_i)\right]^{m'} \cdot (1 + \exp(\alpha_i + \lambda))^{m''} \]

with \( m' = m_0^* - m_0 = m_l - m_l^* = \sum_{t=1}^{T} (y_t y_{0t}) - (y_T y_{0T}) \) and \( m'' = m_l^* - m_l \).

If in Equation (3A) we put \( \sum_{t=2}^{T} y_{0t} = \sum_{t=2}^{T} y_t \), and \( y_{0t} = y_t \), the incidental parameters problem disappear because \( L(\lambda|y, y_0) \) does not depend on \( \alpha_i \).

Thus given \( y_1 \) as initial condition it can be said that \( \sum_{t=1}^{T} y_t \) and \( y_T \) is the set of minimal sufficient statistics for the \( \alpha_i \), conditionally to \( \lambda \).
Appendix B: The conditional likelihood

In Appendix A we saw that the joint likelihood is defined as (2A):

\[
L(\alpha_i, \lambda | y) = P(y_1) \cdot \frac{\exp \left[ \alpha_i \left( \sum_{t=2}^{T} y_t \right) + \lambda \left( \sum_{t=2}^{T} y_t y_{t-1} \right) \right]}{(1 + \exp(\alpha_i))^{m_0} \cdot (1 + \exp(\alpha_i + \lambda))^{m_1}}
\]

with \( m_1 = \text{the number of times } y_t \text{ is preceded by } (y_{t-1} = 1) \):

\[
m_1 = \sum_{t=2}^{T} (y_{t-1}) = \sum_{t=2}^{T} y_t - y_T
\]

and \( m_0 = \text{the number of times } y_t \text{ is preceded by } (y_{t-1} = 0) \): \( m_0 = (T-1) - m_1 \)

Conditional to the set of sufficient statistics the likelihood function is the quotient of the joint likelihood and the marginal likelihood:

\[
L(\lambda | y, \tau) = \frac{\exp \left[ \alpha_i \left( \sum_{t=2}^{T} y_t \right) + \lambda \left( \sum_{t=2}^{T} y_t y_{t-1} \right) \right]}{(1 + \exp(\alpha_i))^{m_0} \cdot (1 + \exp(\alpha_i + \lambda))^{m_1}} \cdot \frac{(1 + \exp(\alpha_i))^{m_0} \cdot (1 + \exp(\alpha_i + \lambda))^{m_1}}{\sum_{d \in Bi} \exp \left[ \alpha_i \left( \sum_{t=2}^{T} y_t \right) + \lambda \left( \sum_{t=2}^{T} y_t y_{t-1} \right) \right]}
\]

\[
= \frac{\exp \left[ \alpha_i \left( \sum_{t=2}^{T} y_t \right) + \lambda \left( \sum_{t=2}^{T} y_t y_{t-1} \right) \right]}{\sum_{d \in Bi} \exp \left[ \alpha_i \left( \sum_{t=2}^{T} y_t \right) + \lambda \left( \sum_{t=2}^{T} y_t y_{t-1} \right) \right]}
\]

As \( \tau = \left\{ y_1, \sum_{t=1}^{T} y_t, y_T \right\} \) is sufficient for \( \alpha_i \), we write:

\[
L(\lambda | y, \tau) = \frac{\exp \left[ \lambda \left( \sum_{t=2}^{T} y_t y_{t-1} \right) \right]}{\sum_{d \in Bi} \exp \left[ \lambda \left( \sum_{t=2}^{T} d_t d_{t-1} \right) \right]}
\]

with \( B_i = \{ d = (d_1, \ldots, d_T) | d_t = 0 \text{ or } 1, d_1 = y_1, \sum_{t \in i} d_t = \sum_{t \in i} y_t, d_T = y_T \} \) the set of all the possible combinations of the \( (y_1, \ldots, y_T) \) vector for a given \( i \).
Taking the natural log of the product over all the individuals we obtain above-mentioned relation (5):

\[
L(\lambda | y, \tau) = \sum_{i=1}^{N} \ln \left\{ \frac{\exp \left( \lambda \sum_{t=2}^{T} y_{it} y_{it-1} \right)}{\sum_{deB} \exp \left( \lambda \sum_{t=2}^{T} d_{it} d_{i(t-1)} \right)} \right\}
\]
Appendix C: Conditional likelihood estimation

To achieve estimation of $\lambda$ in (5) we need all the possible combinations of $y_{it}$ over an estimation period greater than 4 years. The problem is that the number of possible trajectories increases with the observation period in the proportion $2^T$ where $T$ is the number of year an individual is observed. Then for $T = 5$ we have to estimate 32 conditional probabilities. But some of them are independent of $\lambda$. For $T = 5$ probabilities that depend on can be summarised in the following box.

<table>
<thead>
<tr>
<th>$y_1$</th>
<th>$y_5$</th>
<th>$\Sigma_t y_t$</th>
<th>$\Sigma_t y_{t-1}$</th>
<th>likelihood contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>01100</td>
<td>$\exp(\lambda) / 1 + 2\exp(\lambda)$ (1)</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>01010</td>
<td>$1 / 1 + 2\exp(\lambda)$ (2)</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>00110</td>
<td>$\exp(\lambda) / 1 + 2\exp(\lambda)$ (3)</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>10010</td>
<td>$1 / 2 + \exp(\lambda)$ (4)</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>10100</td>
<td>$1 / 2 + \exp(\lambda)$ (5)</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>11000</td>
<td>$\exp(\lambda) / 2 + \exp(\lambda)$ (6)</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>01001</td>
<td>$1 / 2 + \exp(\lambda)$ (7)</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>00101</td>
<td>$1 / 2 + \exp(\lambda)$ (8)</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>00011</td>
<td>$\exp(\lambda) / 2 + \exp(\lambda)$ (9)</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>3</td>
<td>11100</td>
<td>$\exp(\lambda) / 2 + \exp(\lambda)$ (10)</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>3</td>
<td>11010</td>
<td>$1 / 2 + \exp(\lambda)$ (11)</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>3</td>
<td>10110</td>
<td>$1 / 2 + \exp(\lambda)$ (12)</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>3</td>
<td>00111</td>
<td>$\exp(\lambda) / 2 + \exp(\lambda)$ (13)</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>3</td>
<td>01011</td>
<td>$1 / 2 + \exp(\lambda)$ (14)</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>3</td>
<td>01101</td>
<td>$1 / 2 + \exp(\lambda)$ (15)</td>
</tr>
</tbody>
</table>
It can be clearly seen that individuals contribute to the likelihood function by four different manners:

\[ L_1 = \frac{\exp(\lambda)}{1 + 2 \exp(\lambda)} \quad L_2 = \frac{1}{1 + 2 \exp(\lambda)} \quad L_3 = \frac{\exp(\lambda)}{2 + \exp(\lambda)} \quad L_4 = \frac{1}{2 + \exp(\lambda)} \]
References


Cahuzac, E., M. Mouchart, and B. Van der Linden (1994), "Examining the econometric relevance of discretizing panel data: an application to the Belgian labour market, Labour, Vol 8, Number 3, Autumn.


Heckman, J.J. (1978), "Simple statistical models for discrete panel data developed and applied to test the hypothesis of true state dependence against the hypothesis of spurious state dependence.", Annales de l'INSEE, No 30-31.


