INFLUENCES OF PAST HISTORY ON THE INCIDENCE OF YOUTH UNEMPLOYMENT: EMPIRICAL FINDINGS FOR THE U.K.**

by

Wiji Narendranathan+ and Peter Elias*

January, 1990.

Revised: November, 1990.

No. 369

+ Department of Economics, University of Warwick, Coventry, CV4 7AL.

* Institute for Employment Research, University of Warwick, Coventry, CV4 7AL.

** This is a revised version of the paper 'A discrete model of state dependence in youth unemployment: Empirical findings for the U.K.' which was circulated before.

This paper is circulated for discussion purposes only and its contents should be considered preliminary.

ABSTRACT

The issue of whether or not there is evidence of a causal relationship between the experience of unemployment and the future economic activity of an individual is, as yet, unresolved by labour economists. Theoretical reasoning suggests that one may expect to find such a relationship, either through a reduction in human capital or through employer 'labelling' or 'screening' processes.

This paper considers the issues which lend complexity to the problem and seeks to address these issues using a variety of techniques. In particular, problems arising from interval and point sampling of longitudinal information, the effects of observed and unobserved heterogeneity and from the lack of control for serially correlated exogenous factors are investigated. The study focuses upon the work histories of a group of young males who left school in 1974 at the age of 16 years in the U.K. We find that, having controlled for observed and unobserved heterogeneity, the odds of becoming unemployed are 2.3 times higher for youths who were unemployed last year than for youths who were not unemployed; but, given the current status, the past unemployment history of the individual is not informative about his future chances of being unemployed.

KEY WORDS: STATE DEPENDENCE, YOUTH UNEMPLOYMENT, LONGITUDINAL ANALYSIS, OMITTED HETEROGENEITY.

JEL CLASSIFICATION: 210, 810.

1. INTRODUCTION

The issue of whether or not there is evidence of a causal relationship between the experience of unemployment and the future economic activity of an individual is, as yet, unresolved by labour economists. Theoretical reasoning suggests that one may expect to find such a relationship, either through a reduction in human capital or through employer 'labelling' or 'screening' processes. Earlier empirical studies of state dependence in youth unemployment have based their analyses on continuous time models and focussed mainly upon the issue of duration dependence in spells of unemployment[1]. Based upon U.S. data of young men who just graduated from high school, Heckman and Borjas (1980) found no evidence that previous occurrences of unemployment or their duration affect future labour market behaviour once sample selection and omitted heterogeneity biases were controlled for in their analysis. Ellwood (1982) who also based his study upon US data, found no evidence of a negative relationship between the duration of a spell of unemployment and the probability of exit from unemployment. However, using a larger and more detailed data set, Lynch (1989) concluded that .... 'as the spell of non-employment increases, the probability of becoming re-employed declines sharply' (p.45). In her similar study of the unemployment durations among a sample of 68 youths from Inner London boroughs, Lynch (1985) reaches the same conclusion.

The importance of this issue, in terms of its implications for policies which address youth labour markets, should not be underestimated. Young people demonstrate high rates of labour turnover, relative to other age groups, for reasons related to their 'sampling' of different types of work and different employers, their low acquisition of employer or occupation-specific human capital and their relative lack of constraints associated with family formation. But high turnover carries with it a risk of unemployment. If the experience of unemployment itself increases the risk of future unemployment, the resulting concentration of unemployment can have serious repercussions as young people enter early adulthood. Thus, although the question of whether or not there is duration dependence in the exit probability from a relatively short spell of unemployment is of intrinsic interest, the issue of whether or not a long term pattern of recurrent unemployment begins to develop in the early years of a person's working life is also very important.

This paper considers the issues which lend complexity to this problem and seeks to address these issues using a variety of techniques. In particular, problems arising from interval and point sampling of longitudinal information, the effects of observed and unobserved heterogeneity and from the lack of control for serially correlated exogenous factors are investigated. By focussing upon the seven-year work histories of a group of young males, all of whom left school in 1974 at the age of 16 years, this study also enables us to address the exogeneity/endogeneity issue raised by the initial status of these individuals.

The plan of the paper is as follows: the next section gives a brief introduction to the data set analysed. Section 3 sets out the model, discusses the different estimation techniques used in the analysis and addresses some important econometric issues raised by this type of analysis. A detailed description of the sample used and the main findings are presented in section 4 and the paper summarises and concludes in section 5.

2. DATA

To address these issues, we require detailed longitudinal information on the time order of spells of unemployment, their incidence and duration. Such information is available from two sources; administrative records (examples are National Insurance, Unemployment Benefit, etc.) or from panel surveys. The information studied in this paper is of the latter variety, given that the records of 'unemployment' contained within the former are influenced by changes in rules and regulations governing eligibility.

Panel data can be grouped into two main categories; those which follow a crosssection of the population and those which track the evolution of a specific cohort, defined in terms of some unique event in time. In the absence of sufficient high quality UK panel data of the 'cross-sectional' variety, we utilise a major birth cohort study for our purpose - the National Child Development Study (NCDS), a longitudinal study of all persons born in Great Britain in one week in March 1958. Originally containing information on over 18,000 persons, attrition had reduced the sample to just over 12,500 persons by 1981[2].

Unlike longitudinal information from administrative sources, records of unemployment from longitudinal surveys are subject to *time interval censoring*. Because of the extremely high cost of conducting such repeated surveys, respondents are asked to recall the incidence and timing of their economic activity status for the period between surveys. In the case of NCDS, this consisted of a seven-year period between the third sweep in 1973/74 and the fourth sweep in 1981. This study utilises information from the retrospective work histories collected at the time of fourth sweep, when study members were aged 23 years.

Time interval censoring of work history information occurs because it is unreasonable, costly and inaccurate to attempt to record a detailed day-by-day or week-by-week recollection of events spanning a long time period. For NCDS, the time interval utilised was month-by-month, asking respondents to characterise each month since leaving school as 'in employment', 'unemployed' or 'out of the labour force'[3]. Inevitably, very short spells of unemployment are censored in the process. The observed 'spells' of unemployment are in fact, a concatenated series of months for which respondents considered themselves to be 'unemployed' (not in employment and wanting work) in each month.

Focussing specifically upon male minimum-age school leavers, this study follows the employment/unemployment experience of 4,067 males all of whom left school at approximately the same time (April to June, 1974) until late in 1981. This group constitutes 65 per cent of all males covered in the 1981 sweep of the birth cohort. Of these, 48 per cent recorded at least one spell and, on average, 2 spells of unemployment in their work histories.

The distribution of the durations of all completed spells of unemployment recorded by this group is shown in Figures 1 and 2. The average duration of a completed spell of unemployment is just under 5 months. It is appropriate, therefore, to utilise an interval sampling approach for analysis, thereby recognising the discrete nature of the underlying data. Further details of data construction and methods of analysis are given in section 4 of this paper.

3. THE MODEL AND ESTIMATION

3.1 The Model and Econometric Issues

Any analysis which attempts to answer questions regarding whether past experience of unemployment increases the likelihood of future unemployment experiences raises two issues. First, the distinction between genuine state dependence and spurious state dependence (Heckman, (1981a)). Second, the treatment of initial observation (Heckman (1981b), Pickles (1987)).

Genuine state dependence, the so-called 'scarring' effect, occurs if past unemployment experience actually changes the likelihood of experiencing unemployment in the future. This can arise, for example, if either an individual's past experiences result in some loss of accumulated human capital, or if the past experiences of unemployment are used as an indicator of 'unreliability' by future employers in their hiring decisions.

Spurious state dependence, on the other hand, can arise for at least three different reasons. First, individuals can differ in their propensity to experience unemployment. This, termed heterogeneity, both observable and unobservable, if not appropriately controlled for, will result in a correlation between past experiences of unemployment and future experiences even if there are no causal links between them. This has been commonly dealt with by including as controls a variety of measured individual characteristics. With longitudinal information, it becomes possible to attempt to control for unobservable characteristics. The potential effect of unobservable characteristics is well documented[4]. Second, there is a need to control for exogenous factors such as marital status, number of children, local labour market conditions faced by the individual etc. If these exogenous factors are serially correlated and are omitted from the model, this can induce spurious state dependency in the unemployment incidence. Third, the observed correlation between the past and present unemployment experiences might be solely due to the sampling framework employed in the collection and analysis of the data. Interval sampling of such data can lead us to observe a causal link between past and present unemployment even if there were no causal links since a single unemployment spell could overlap between two consecutive periods. The model presented below allows us to account explicitly for these problems.

We have chosen to model the incidence of unemployment as a second order auto-regressive latent continuous random variable y_{it}^*. Data limitations preclude us from allowing for higher order processes. We thus have,

$$y_{it}^* = x_{it}'\beta + z_i'\delta + \gamma_1 y_{it-1} + \gamma_2 y_{it-2} + \alpha_i + v_{it} \quad , \qquad i = 1,..,N \text{ and } t = 1,...,T$$

where

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^* > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

and

$$\begin{aligned} y_{it} &= 1 \quad \text{if individual } i \text{ is unemployed in period } t^5 \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

It is also assumed that the following variables influence the conditional probability of the individual i experiencing a spell of unemployment in period t. The x_{it} is a vector of exogenous observable personal and environmental characteristics such as marital status, number of children, local labour market conditions etc., which are assumed to vary over time. The z_i are time invariant exogenous observable personal characteristics such as family social background of the individual, educational indicators of performance etc. and the α_i is an unobservable time invariant individual specific variable. We also assume that the error term v_{it} is an independently and identically distributed random variable which is also distributed independently of all the explanatory variables.

The above assumptions, coupled with the assumption that v_{it} is distributed as a Logistic, gives us the familiar logit model:

$$\text{Prob}(\, y_{it} = 1 \mid y_{it-1}, y_{it-2}, x_{it}, z_i, \alpha_i\,) = \frac{\exp(\, x_{it}'\beta + z_i'\delta + \gamma_1 y_{it-1} + \gamma_2\, y_{it-2} + \alpha_i\,)}{1 + \exp(\, x_{it}'\beta + z_i'\delta + \gamma_1 y_{it-1} + \gamma_2\, y_{it-2} + \alpha_i\,)} \qquad (2)$$

As we shall see, the assumption of a logit model, unlike a probit model, enables the estimation of the model using conditional maximum likelihood methods which help us to eliminate the unobservable α_i's.

The other issue which we have to address is the initial conditions problem. This refers to the way in which pre-sample characteristics of the data are dealt with in a dynamic model where the past experience influences the current outcome. The two most common assumptions (though both are subject to criticisms) that are made about initial conditions are (i) the initial conditions or the relevant pre-sample history of the process are truly exogenous, or (ii) the process generating the data is in equilibrium. Given that we have carried out our analyses on the unemployment experience of young men from the time they left school, the assumptions that are needed to circumvent the problems of initial conditions in our model are that (i) the decision to leave school itself is not dependent upon the employment prospects and (ii) any attrition observed is exogenous[6].

3.2 Estimation

Given a set of observations on N individuals over T time periods, one could use maximum likelihood estimation techniques to estimate the parameters of the above model. First, however one has to decide upon the treatment of the the α_i term, the unobservable variable. Treating these α_i's as parameters (i.e. as fixed effects) and estimating them in a direct maximum likelihood estimation of the above model leads to the well known problem of incidental parameters (see Cox and Hinkley (1974)). Since the consistency property of the parameters of interest are based on large N and not large T, and the number of α_i's increase with N, we have to either eliminate the α_i's from the estimation procedure or treat the α_i's as random variables with a specific distribution and estimate the parameters of this distribution along with the other parameters of the model. The first of the above is achieved by the use of the maximisation of the conditional likelihood (CMLE) and the second with the use of the maximisation of the marginal likelihood (MMLE) where the marginalisation is carried out with respect to the unobservable variable α. Both these procedures have their advantages and disadvantages. The main advantage of CMLE is that, unlike the MMLE, it does not require the assumption of independence of the α_i and the other included variables and the main disadvantage is that the model as presented in equation (1) can not be estimated without further restrictions on the parameters of the model. But, if this independence assumption is not violated, then, the MMLE is more efficient compared to the CMLE[7] since the former uses more information.

The main disadvantage with the MMLE is that it requires an assumption regarding the parametric form of the distribution of the α_i. Since the CMLE is very easily carried out compared to the MMLE, the results from both estimations are presented and discussed in this paper. We also check for the sensitivity of our results to the assumed form of the density for the α.

Conditional likelihood maximisation

A technique, originally due to Andersen (1970) and then used by Chamberlain (1980; 1985), allows us to obtain consistent estimates of some of the parameters of interest in logit models of this type by maximising a conditional likelihood function where the conditioning is carried out with respect to a set of sufficient statistics. The only short-coming of this approach is that the model as specified in equation (1) cannot be estimated as it stands without imposing further restrictions. If $\gamma_1 = \gamma_2 = 0$, then consistent estimation of β coefficients are possible as shown in Chamberlain (1980). On the other hand, one could obtain a consistent estimate of γ_2 if $\beta = 0$. The conditioning on the set of sufficient statistics, while eliminating the α_i's, also eliminates the δ coefficients in the first case and, δ and γ_1 in the second case. For an application of this technique in the univariate case, see Corcoran and Hill (1980, 1985) and in the bivariate case, see Narendranathan, Nickell and Metcalf (1985).

Assuming that $\beta = 0$, i.e. that there are no exogenous variables that vary over time affecting the probability of the individual experiencing unemployment, we have, for individual i,

$$\text{Prob}(\ y_{it} = 1 \mid y_{it-1}, y_{it-2}, z_i, \alpha_i\ ) = \frac{\exp(\ z_i'\delta + \gamma_1 y_{it-1} + \gamma_2\, y_{it-2} + \alpha_i\ )}{1 + \exp(\ z_i'\delta + \gamma_1 y_{it-1} + \gamma_2\, y_{it-2} + \alpha_i\ )} \tag{3}$$

Thus,

$$\text{Prob}(y_{i1}, y_{i2}, \ldots, y_{iT}) = p(y_{iT} \mid y_{iT-1}, y_{iT-2})\, p(y_{iT-1} \mid y_{iT-2}, y_{iT-3}) \ldots p(y_{i2}, y_{i1})$$

$$= A/B \text{ say,}$$

where,

$$A = p(y_{i2}, y_{i1}) \exp\{\alpha_i \sum_{3}^{T} y_{it} + \gamma_1 \sum_{3}^{T} y_{it}\, y_{it-1} + \gamma_2 \sum_{3}^{T} y_{it}\, y_{it-2}\} \tag{4}$$

and $$B = [1+\exp(\alpha_i+\gamma_1+\gamma_2)]^{\sum_3^T y_{i\,t-1} y_{i\,t-2}} [1+\exp(\alpha_i+\gamma_1)]^{\sum_3^T y_{i\,t-1}(1-y_{i\,t-2})}$$

$$[1+\exp(\alpha_i+\gamma_2)]^{\sum_3^T y_{i\,t-2}(1-y_{i\,t-1})} [1+\exp(\alpha_i)]^{P_1}$$

$$P_1 = T - 2 - \sum_3^T y_{i\,t-1} y_{i\,t-2} - \sum_3^T y_{i\,t-2}(1-y_{i\,t-1}) - \sum_3^T y_{i\,t-1}(1-y_{i\,t-2})$$

$$= T - 2 + \sum_3^T y_{i\,t-1} y_{i\,t-2} - 2\sum_1^T y_{i\,t} + y_{i1} + y_{i\,T-1} + 2\,y_{iT}$$

($z_i'\delta$ term has been absorbed into the α_i s). Looking at the terms containing the α_i s, we see that, the following constitutes a set of sufficient statistics for the α_i s:

$$y_{i1},\ y_{i2},\ y_{iT},\ y_{i\,T-1},\ \sum_1^T y_{it},\ \text{and} \sum_2^T y_{it}\, y_{i\,t-1}\ .$$

Since, γ_1 occurs in A with the term $(\Sigma y_{it}\, y_{i\,t-1})$, when we condition the analysis on the set of sufficient statistics, γ_1 will also get eliminated[8]. That is, γ_1 is not identifiable in this model. Given the above set of conditions, one needs T to be at least 6.

For example, consider a model with 6 time periods and focus on the contribution of those individuals to the conditional likelihood whose sequences satisfy the following (suppressing the i index):

$$y_1 = 1,\ \ y_2 = 0,\ \ y_5 = 0,\ y_6 = 0,\ \sum_1^6 y_t = 2 \ \text{ and } \sum_2^6 y_t\, y_{t-1} = 0 \qquad (5)$$

There are two possible six period sequences satisfying these restrictions, namely {1 0 1 0 0 0} and {1 0 0 1 0 0}. The probability of the first sequence conditional on (5) then simply reduces to

$$\text{Prob}(\,1\,0\,1\,0\,0\,0 \mid 1\,0\,1\,0\,0\,0 \text{ or } 1\,0\,0\,1\,0\,0\,)$$

$$= \frac{\text{Prob}(\,1\,0\,1\,0\,0\,0\,)}{\text{Prob}(\,1\,0\,1\,0\,0\,0\,) + \text{Prob}\,(\,1\,0\,0\,1\,0\,0\,)} = \frac{\exp(\gamma_2)}{(\,1 + \exp(\gamma_2)\,)}. \qquad (6)$$

The total conditional likelihood contribution is then built up of a product of terms such as those given in (6), with each individual contributing one such term. Many individuals, such as those who have only one spell, will only contribute a constant to the likelihood because none of the parameters enter. When T=6, there are only 8 patterns of sequences which have a contribution to the conditional likelihood which is not equal to just a constant. They are the following:

the pattern of sequence	the contribution to the likelihood function
(1) 1 0 1 0 0 0	$\exp(\gamma_2)\ /\ (1+\exp(\gamma_2))$
(2) 1 0 0 1 0 0	$1\ /\ (1+\exp(\gamma_2))$
(3) 0 1 0 1 1 1	$\exp(\gamma_2)\ /\ (1+\exp(\gamma_2))$
(4) 0 1 1 0 1 1	$1\ /\ (1+\exp(\gamma_2))$
(5) 0 0 0 1 0 1	$\exp(\gamma_2)\ /\ (1+\exp(\gamma_2))$
(6) 0 0 1 0 0 1	$1\ /\ (1+\exp(\gamma_2))$
(7) 1 1 1 0 1 0	$\exp(\gamma_2)\ /\ (1+\exp(\gamma_2))$
(8) 1 1 0 1 1 0	$1\ /\ (1+\exp(\gamma_2))$

where 0 denotes the state of employment and 1 the state of unemployment[9].

The conditional maximum likelihood estimate of the parameter γ_2, which is easily derivable, is given by $\hat{\gamma}_2 = \ln(m_1/m_2) = \ln($no. of individuals with patterns 1,3,5 and 7 / no. of individuals with patterns 2, 4, 6 and 8). The associated standard error = $(m_1+m_2)/(m_1m_2)$. A test of the hypothesis that $\gamma_2 = 0$ is a test of the Markov process for the incidence of unemployment in this model with stationary heterogeneity.

If we find that the null hypothesis of $\gamma_2 = 0$ cannot be rejected, we could then impose this restriction on the model. This restriction then allows one to estimate γ_1 using similar technique. Since T is 6 in our sample, the number of sequences of different patterns which contribute towards the likelihood function is now 48 as shown in the appendix.

Marginal likelihood maximisation

As we saw earlier, the above conditional likelihood maximisation does not allow us to model the effects of personal and environmental characteristics which vary over time on the probability of experiencing an incidence of unemployment. We therefore, consider a different approach to the estimation which will enable us to estimate the full model as presented in equation (1). This is the maximisation of the marginal likelihood function where the marginalisation is carried out with respect to the incidental or 'nuisance' parameters, the α_i s.

Conditioning on the first two observations for each individual, the marginal likelihood contribution for the ith individual is

$$L_i(\beta,\delta,\gamma_1,\gamma_2,\theta \mid y_{i1},y_{i2}) = \int \left[\prod_{t=3}^{T} \frac{[\exp(x_{it}'\beta + z_i'\delta + \gamma_1 y_{i,t-1} + \gamma_2 y_{i,t-2})]^{y_{it}}}{1 + \exp(x_{it}'\beta + z_i'\delta + \gamma_1 y_{i,t-1} + \gamma_2 y_{i,t-2})} \right] f(\alpha)\, d\alpha \qquad (7)$$

where f(.) is the probably density function of the the unobservable random error α, usually referred to as the 'mixing distribution', with parameter vector θ.

For the above marginalisation to be valid one needs to assume that the omitted variable is independent of all the included explanatory variables. In our model, where we have allowed the past two period experiences of unemployment to affect present unemployment, the omitted variables may be correlated with the initial conditions. We assume, therefore, that there is one set of initial conditions common to all individuals in the sample and choose a relatively homogeneous sample of young males who left school at approximately the same time, to overcome this initial conditions problem in our model.

The next stage of the analysis requires the specification of $f(\alpha)$. There are two alternative techniques that have been followed in the literature in the operationalisation of the above model. One is a fully parametric approach and the other a non-parametric approach. In the parametric approach, one can either assume an analytically tractable distribution for f(.) that will give a closed form expression for the marginal likelihood (see Heckman and Willis (1977) who use a Beta-Logit pair and Lancaster (1979) who uses a Gamma-Weibull pair) or some other parametric form and then carry out the integration using numerical methods. In the non parametric approach the integral in equation (7) is replaced by a finite sum, i.e. f(.) is replaced by an empirically determined number of mass points is used (see for example, Heckman and Singer (1984), Davies and Crouchley (1984) for an application of this technique). We follow both approaches in order to check for the sensitivity of the results.

<u>Parametric Approach</u>: In the absence of any guidance from economic theory as to the form of the density f(.), we have assumed a Normal distribution for the α to estimate[10] the parameters of interest.

An obvious weakness of a parametric model of this kind is that the tail behaviour does not allow enough flexibility to model those individuals who never change their status (see Barry et al (1989)). The model estimated in this paper takes care of this by estimating empirically determined masses at the two extremes, i.e. plus and minus infinity of the Normal mixing distribution f(.). This gives us the following likelihood for individual i,

$$L_i^* = \frac{\psi_0}{1+\psi_0+\psi_1}\left[\prod_{t=3}^{6}(1-y_{it})\right] + \frac{\psi_1}{1+\psi_0+\psi_1}\left[\prod_{t=3}^{6} y_{it}\right] + \frac{L_i}{1+\psi_0+\psi_1} \tag{8}$$

where, L_i is given by equation (8) and the ψ_0 and ψ_1 are the unknown end point parameters. Thus, the proportion of individuals who are stayers in state 0, i.e. the proportion of individuals who have not experienced a spell of unemployment at the same time each year is given by,

$$p_0 = \frac{\psi_0}{1 + \psi_0 + \psi_1} \tag{9}$$

and the estimated proportion of stayers in state 1, i.e. those who are always unemployed is given by

$$p_1 = \frac{\psi_1}{1 + \psi_0 + \psi_1} \quad . \tag{10}$$

Non-parametric Approach: In this approach, which is completely non-parametric, the integral in equation (7) is replaced by a finite sum. That is, f(.) is replaced by an empirically determined number of mass points,

$$f(\alpha_i) = p_k \quad \text{for} \quad \alpha_i = \xi_k \quad k=1,.....m \quad \text{and} \quad \sum_{k=1}^{m} p_k = 1 \tag{11}$$

$$= 0 \quad \text{otherwise,}$$

where m is the appropriate number of mass points. Although the resulting model is algebraically identical to a finite mixture model it is conceptually very different from it. In the finite mixture models, it is assumed that these mass points reflect finite divisions within the population (see for example, Bartholomew (1959) and Nickell (1979)). But, in this model, the underlying assumption is that, a finite number of mass points are adequate to control the unexplained variation represented by the α_i. It should also be noted that this is not a simplification of the parametric approach as the identification of the number of mass points, their locations and the associated probabilities present formidable computational problems.

4. SAMPLE AND MODEL ESTIMATION RESULTS

The labour market history data used in the analysis records each individuals' main economic status in every month since leaving school in 1974 till 1980. This month by month continuous history was converted to an *annual* framework of point sampling for the discrete model, in an attempt to avoid the problems associated with spurious state dependence.

Spurious state dependence can arise from an interval sampling of work history information if the average duration of the events under investigation (in this case, periods of employment and unemployment) is similar to the intervals used to characterise the sample. For example, if one sampled every third month of a longitudinal record of employment and unemployment, and the average duration of a spell of unemployment was nine months, one would detect a consistent relationship between unemployment in periods t-2, t-1 and t. The choice of twelve month intervals in our model, where the average unemployment duration is around 5 months, avoids this particular cause of spurious state dependence.

Ideally, sampling should be undertaken from successive intervals which are further apart than the average duration of the event under investigation, in this instance unemployment, from longitudinal information in which the intervals, themselves are short enough to capture a record of events which may be of short duration. Clearly, however, there is a trade-off between the expense of collecting reliable and accurate longitudinal information in which extremely short intervals (say 1 week) have been characterised as employment or unemployment and the need to record short spells of unemployment which may be significant in terms of their impact upon an individual's future employment.

In codifying their work histories, respondents were requested to categorise each month since leaving school as 'paid employment' (including time on training courses if part of a paid job), 'full-time education', 'unemployment' or 'out of the labour force'. The latter category was a catch-all category for use by respondents who could not categorise any particular month into one of the three preceding states. Very few minimum age male school-leavers return to full-time education. Also, use of the category 'out of the labour force' by males is rare. For this reason, the work histories of minimum age male school- leavers were re-classified as 'in unemployment' or 'not in unemployment' for each month after leaving school in 1974. 'Unemployment' consists of all periods of time when the respondent was not in employment and was wanting work, even if a period lasted less than a month. The work history data were then reconstructed as a series of 6 bi-monthly observations of economic activity as follows:

(i) Twelve monthly observations of economic activity were paired into six bi-monthly periods; January-February, March-April, May-June, July-August, September-October, November-December.

(ii) Each bi-monthly period was characterised as a period spent in unemployment if, and only if, for both months within that period the study member was unemployed. This procedure captures significant spells of unemployment in the longitudinal work histories. Transitional spells resulting from one or two weeks without work between jobs are ignored.

Thus, in our model, if we find that we can not reject the hypothesis that $\gamma_2 = 0$ then this would imply that, given an individual's current unemployment status, his prior unemployment history is not informative about his future. This implies a first order Markov process for the labour market status variable. That is, ceteris paribus, all information required to predict the individual's labour market status at the same time next year is contained in the current labour market status. On the other hand, if we find that we can not reject the hypothesis that $\gamma_1 = \gamma_2 = 0$, then this would imply that, the individual's previous year's unemployment status does not allow us to say anything about his current unemployment status[11].

To check for unmodelled seasonal effects in the model, we have carried out all our analyses by varying the months under consideration[12]. The descriptive statistics of some variables of interest for the sample members are given in Tables 1 and 2.

The NCDS contains detailed information relating to the social and educational development of the birth cohort, their medical history, and a brief record of each individual's work history from the age of 16 years. On the basis of earlier research, examining the links between social and educational development and the observed work history (Elias and Blanchflower (1989)), certain variables have been selected as key factors which are related to early labour market experience. Foremost among these factors is information upon childhood mathematical abilities and reading development. This information consists of score results from tests of maths and reading comprehension taken by cohort members at the age of 11 years[13]. Test scores were then transformed to a binary variable, with a value of unity if a score on either test was below the average for all cohort members.

Additionally, information on parental social class was developed from details of the occupation of the respondent's father (or mother in the absence of paternal information) at the time of the birth of the cohort member in 1958.

It can be seen from Table 1 that male minimum age school leavers tended to score slightly below average on the reading and maths tests at age 11 years and were predominantly the sons of manual workers. The years 1974 to 1979 saw an increase in the average unemployment rate as recorded in official sources for the mid year, from an average of 3.7 per cent across the travel-to-work- areas in which respondents lived in 1974, to 7.0 per cent by 1981. Although 11 percent of the school leavers in the sample were unemployed in the sampled 'double-month' of the year after leaving school, this figure dropped down to around 5 to 7 percent over the next six years during the same double months. Around 22 percent of the sample members were married by year 6 after leaving school. Eleven percent had fathered at least one child by the same date.

For those who experienced the patterned sequences of unemployment, a much higher proportion scored below average on the maths and reading tests, had a father/guardian with a manual occupation at the time of their birth and were resident in a high unemployment travel-to-work area in 1981.

Earlier research (Elias and Blanchflower, 1989) has indicated that children with a high birth order (fifth born or higher) were much more likely to have scored below average in the reading and maths tests, to have left school at the minimum school leaving age and to have experienced a discontinuous work history. This is again evident in Table 2.

4.1 Conditional Maximum Likelihood results

The results of the estimation of equation (2) is presented are column 1 of Table 3. If there is no duration dependence in the unemployment process (see footnote 11) then we would expect the effect of the second order term to be insignificant since the model would reduce to a first order Markov chain. As we can see from Table 3 results, although the coefficient estimate of γ_2 varies with the reference pair of months, the null hypothesis of no effect cannot be rejected.

We next impose this restriction on the model and estimate the effect of the first order term. These based on about 218 individuals on average and produce estimates of γ_1 which varies from 1.14 to 1.30 with the larges effect estimated for the March/April months. The null hypothesis of $\gamma_1 = 0$ is very easily rejected in all cases. Thus, for example, given that the individual was unemployed in July/August months last year, the odds that he will be unemployed the same time this year, is about 3.3 times higher ($e^{1.20}$) than if he was not unemployed last year. This figure is not sensitive to the choice of the months sampled.

The above models do not allow us to control for influences of exogenous factors which might be correlated over time. Thus, in the next model analysed, we relax the assumption of $\beta = 0$ and estimate the model presented in equation (7) to see whether the time varying variables such as local unemployment rate and some personal characteristics do have any effect on the probability of becoming unemployed, ceteris paribus.

4.2 Marginal Maximum Likelihood results

The results of the marginal likelihood maximisation for the July/August bi-monthly periods are presented in Tables 4 and 5. Table 4 contains the results of the estimation where a Normal heterogeneity was assumed and Table 5 where this assumption was replaced by a non-parametric mass-points[14]. Model 3 results refer to the MMLE which correspond to the CMLE results of Model 1[15]. Unlike Models 6 and 7, Models 4 and 5 do not allow for any lagged unemployment effects. Models 4 and 6 are presented to show the effects of not allowing for omitted heterogeneity. Model 8 allows for 2 mass points and Model 9 for 3 mass points. We could not achieve any increase in the maximised value of the log likelihood function with the addition of extra mass points to Model 8.

We have chosen to present and concentrate our discussion mainly on the July/August model results for the following reasons. Our sample refers to school leavers who left school when they were 16 years old. Since the school year ends in July in the U.K., the July/August period marks the start of the employment history for these individuals. Thus, bearing in mind the discussion on the initial conditions problems (section 3), we cannot infer whether some of the differences we find from different bi-monthly models are due to seasonal effects or to some bias induced by the non-exogeneity of the initial conditions. We expect this problem to be minimal in the July/August analysis.

The main results which emerge from the estimation of these models is, as expected, that on the one hand, allowing for omitted heterogeneity in these models increases the effects of the explanatory variables, and on the other hand, including lagged unemployment effects decrease the effects of the explanatory variables. That is, the effects of omitted heterogeneity and omitted lagged unemployment effects go in opposite directions.

Comparing Model 3 results with that of 7, we find that the likelihood ratio test of the joint null hypothesis of zero effects from the time varying variables is very easily rejected at conventional levels of significance with a test statistic value of 109. Similar to Model 1 results, we still find the 2nd order effect of past unemployment to be still insignificant.

The important point to note about the results in Tables 4 and 5 is that the results from Model 7 are very similar to results from Model 9. That is, replacing the fully parametric assumption that the omitted heterogeneity variable is distributed as Normal and allowing for differing end point behaviour produces similar results to the fully non-parametric model for the omitted heterogeneity variable. We also find that the non-parametrically estimated heterogeneity distribution is skewed to the left with the largest mass on the left side which is indicative of the fact that there is a large group who never experience unemployment.

Since both models 7 and 9 have produced very similar results we shall concentrate our discussion on Model 9 results. First important point to note is that, as we found with the CMLE, γ_2 is insignificantly different from zero. That is, given the current status, the past unemployment history is not informative about his future status. But, we do find a strong first order effect. Given that the individual was unemployed in the July/August months last year, the odds that he will be unemployed at the same time this year is about 2.3 times higher than if he was not unemployed last year. This compares to a figure of 3.3 found from the CMLE.

The local unemployment rate refers to the area where the individual lived when he left school in 1974. By 1981 about 14 per cent of the individuals had moved to a different location. We do not have information on when they actually moved nor information on how many times they had moved in the 6 intervening years. Though the results were not sensitive to whether we used the unemployment rate relevant to the area of residence in 1974 or 1981, because of the possible endogeneity associated with any variable other than the 1974 location variable, we only report results of estimation which used this variable. We do find that the local unemployment rate has a small but significant effect on the incidence of unemployment ceteris paribus. A one percentage point increase in the local unemployment rate increases the odds of a minimum age school-leaver being unemployed in that year by about 1.2.

Whether the individual had an above average maths and reading scores also matter for the unemployment experience of these individuals. For example, the odds that an individual with a below average maths score will be unemployed this year is 1.8 times higher compared to an individual who has an above average maths score. With regard to reading scores, this figure is slightly lower at 1.6. As expected, sons of non-manual men have a lower probability of becoming unemployed ceteris paribus. The most affected individuals are the sons of unskilled manual workers.

We also find that, being married decreases the chances of becoming unemployed. Given current status, the odds that a married man will be unemployed next year compared to an unmarried man is about 1.5 times lower. The probability of becoming unemployed in any year is estimated to increase with the appearance of children; the odds that a minimum age school leaver who has fathered two children will be unemployed in any year after the birth of the second child is over six times greater than for a similar man with no children. This could relate to the flat rate nature of benefit system relative to low earnings of youths[16].

5. SUMMARY AND CONCLUSION

We have presented and estimated a model for explaining the incidence of unemployment where we have explicitly taken account of omitted heterogeneity and initial condition problems and also allowed for state dependence in the unemployment process.

We find that modelling state dependency while allowing for omitted heterogeneity in the incidence of youth unemployment in the U.K. is very important. Once having allowed for the effects of observed and unobserved heterogeneity, we only find a significant first order effect. More specifically, given the current status, the past unemployment history of the individual is not very informative about his future unemployment experiences. The odds of becoming unemployed are 2.3 times higher for youths who were unemployed last year than for youths who were not unemployed. The time varying variables such as the local unemployment rate, marital status, children etc. and the time invariant variables such as father's social class, whether one scored a mark which was above or below the average score in maths and reading achievement at 11 years of age etc., and whether one was unemployed last year, were all significant determinants of the probability of being unemployed this year. Of particular surprise is the finding that, having controlled for social background, educational test scores, previous employment history and unobserved heterogeneity, the fathering of children has such a pronounced effect upon the odds of being unemployed. This suggests that the high benefit replacement ratio for low-income males with adult dependants and children has a significant work disincentive effect.

FOOTNOTES:

1 In all of these studies, the unemployment duration is modelled by specifying the conditional probability of leaving unemployment (the hazard function) without distinguishing the exit state. However, some of those who exit the unemployment do not do so into employment. Thus, one should strictly interpret the evidences as being about the exit probability from unemployment rather than as being about the re-employment probability.

2 For further information about the NCDS, see National Children's Bureau (1984). Further details about the development of these longitudinal work histories are given in Elias and Blanchflower (1989).

3 For full details, see Elias and Blanchflower (1989).

4 See for example, Lancaster and Nickell (1980), Heckman and Singer (1984), Lancaster (1985).

5 Since the choice of the time period t is of crucial importance for the avoidance of the econometric problems discussed earlier we shall give precise definition in section 4.

6 Heckman (1981b) and Pickles (1987) discuss various ways in which one could deal with the initial condition problems under different set of assumptions and their effects on the maximum likelihood estimates.

7 See Crouchley (1987) for a comparison in terms of simulations.

8 Obviously, this does not preclude us from specifying the model in more general terms by giving each individual a different γ_1 parameter.

9 For purpose of clarity of exposition, we shall refer to all states of non-unemployment as employment.

10 We are grateful to the authors of SABRE (Barry et al (1989)) for providing us with the software for the estimation of this model.

11 If the underlying process of time spent in the unemployment and employment states followed an alternating renewal process, then the binary sequence generated by point sampling would be a first-order Markov chain (see Chamberlain (1985) p.15). Thus, a test of $\gamma_2 = 0$ in our model, is a test for duration dependence.

12 In the interests of brevity and for reasons regarding the initial conditions stated in the results section, we have only presented the results for the July/August months.

13 In some instances where test scores were missing at age 11 years, imputed scores were used from similar tests taken at age 16 years. For full details of the imputational procedures, see Elias and Blanchflower (1989:24).

14 We are grateful to the authors of MIXTURE (Ezzet and Davies, 1987) for providing us with software for the estimation of this model.

15 They are not directly comparable since Model 1 is consistent with a more general model which allows each individual to have his own γ_1 parameter.

16 In 1977, the benefit replacement ratio for a single male aged between 18 and 21 years was 34% of the average earnings of male manual employees aged 18-21 years. With an adult dependent plus one child, the ratio rises to 63%, with the second child it rises again to 70% of average earnings. For males in non-manual employment the ratios are approximately 10% higher.

REFERENCES:

Andersen, E. B. (1970) - 'Asymptotic properties of conditional maximum likelihood estimators', Journal of the Royal Statistical Society, Series B, 32, 283-301.

Barry, J., Francis, B., and Davies, R. B. (1989) - SABRE (Software for the Analysis of Binary Recurrent Events), Centre for Applied Statistics, Lancaster University.

Bartholomew, D. J. (1959) - 'Note on the measurement and prediction of labour turnover', Journal of the Royal Statistical Society, Series A, 122, 232-239.

Chamberlain, G. (1980) - 'Analysis of covariance with qualitative data', Review of Economic Studies, XLVII, 225-238.

Chamberlain, G. (1985) - 'Heterogeneity, omitted variable bias, and duration dependence', Chapter 1 in Longitudinal Analysis of Labour Market Data, Ed. by Heckman, J. J. and Singer, B. L., Cambridge University Press.

Corcoran, M. and Hill, M. S. (1980) - 'Persistence in unemployment among adult men', Chapter 2 in Five Thousand American Families - Patterns of Economic Progress, Vol. 8, Ed. by Freeman, R. B., Wise, D. A. and Morgan, J. N., Ann Arbor.

Corcoran, M. and Hill, M. S. (1985) - 'Reoccurrence of unemployment among adult men', The Journal of Human Resources, XX, 2, 165 - 183.

Cox, D. R., and Hinkley, D. V. (1974) - Theoretical Statistics, Chapman and Hall, London.

Crouchley, R. (1987) - 'A comparison of conditional and marginal likelihood methods for estimating the logistic model in collections of binary sequences' in Longitudinal Data Analysis, ed. by R. Crouchley, pp. 150-76, Avebury, England.

Davies, R. B. and Crouchley, R. (1984) - 'Calibrating longitudinal models of residential mobility and migration', Regional Science and Urban Economics, 14, 231-247.

Elias, P. and Blanchflower, D. (1989) - 'The occupations, earnings and work histories of young adults - who gets the good jobs?', Research Paper No. 68, Department of Employment, UK.

Ellwood, D. (1982) - 'Teenage unemployment: Permanent scars or temporary blemishes', in The Youth Labour Market Problem: Its nature, causes and consequences, ed. R. Freeman and D. Wise, University of Chicago Press.

Ezzet, F. L. and Davies, R. B. (1987) - A Manual for MIXTURE, Centre for Applied Statistics, University of Lancaster.

Heckman, J. J. (1981a) - 'Statistical models for discrete panel data' in Structural Analysis of Discrete Data with Econometric Applications, ed. by C. F. Manski and D. McFadden, pp. 114-78, Cambridge: MIT Press.

Heckman, J. J. (1981b) - 'The incidental parameters problem and the problem of initial conditions in estimating a discrete time - discrete data stochastic process' in Structural Analysis of Discrete Data with Econometric Applications, ed. by C. F. Manski and D. McFadden, pp. 179-95, Cambridge: MIT Press.

Heckman, J. J. and Borjas, G. J. (1980) - 'Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence', Economica, 47, 247-283.

Heckman, J. J., and Singer, B. L. (1984) - 'Econometric duration analysis', Journal of Econometrics, 24, 63-132.

Heckman, J. J., and Singer, B. L. (1984) - 'A method for minimising the impact of distributional assumptions in econometric models for duration data, Econometrica, 52, 271-230.

Heckman, J. J., and Willis, R. (1977) - 'A beta-logistic model for the analysis of sequential labour force participation by married women', Journal of Political Economy, 85, 27-58.

Lancaster, T. (1985) - 'Generalised residuals and heterogeneous duration models: with application to the Weibull model', Journal of Econometrics, 28, 155-169.

Lancaster, T., and Nickell, S. J. (1980) - 'The analysis of re-employment probabilities for the unemployed (with discussions), Journal of Royal Statistical Society, A, 143, part 2, 141-65.

Lynch, L. M. (1985) - 'State dependency in youth unemployment: A lost generation?', Journal of Econometrics, 28, 71-84.

Lynch, L. M. (1989) - 'The youth labour market in the eighties: Determinants of re-employment probabilities for young men and women', The Review of Economics and Statistics, 27-45.

Narendranathan, W., Nickell, S. J. and Metcalf, D. (1985) - 'An investigation into the incidence and dynamic structure of sickness and unemployment in Britain, 1965-75', The Journal of the Royal Statistical Society, Series A, 148, part 3, 254-267.

National Children's Bureau (1984) - National Child Development Study (1958 Cohort): Fourth Follow-up. Final Report to Sponsors. London: National Children's Bureau.

Nickell, S. J. (1979) - 'Estimating the probability of leaving unemployment', Econometrica, 47, 1249-66.

Pickles, A. R. (1987) - 'The problems of initial conditions in longitudinal analysis' in Longitudinal Data Analysis, ed. by R. Crouchley, pp. 129-49, Pub. Avebury, England.

APPENDIX

The following gives a list of the patterns of sequences which make a contribution to the conditional likelihood function in the presence of restrictions $\beta = \gamma_2 = 0$. The subscript 1 in γ_1 has been dropped for notational convenience. The set of sufficient statistics for this analysis is y_1 , y_T and Σy_t .

The pattern of sequence	The contribution to the likelihood function
a) $y_1 = 1;\ y_6 = 0;\ \Sigma y_t=2;$	
1 1 0 0 0 0	$e^{\gamma} / (e^{\gamma}+ 3)$
1 0 1 0 0 0	$1 / (e^{\gamma}+ 3)$
1 0 0 1 0 0	$1 / (e^{\gamma}+ 3)$
1 0 0 0 1 0	$1 / (e^{\gamma}+ 3)$
b) $y_1 = 1;\ y_6 = 0;\ \Sigma y_t=3;$	
1 1 1 0 0 0	$e^{2\gamma} / (e^{2\gamma} + 4 e^{\gamma} + 1)$
1 1 0 0 1 0	$e^{\gamma} / (e^{2\gamma} + 4 e^{\gamma} + 1)$
1 1 0 1 0 0	$e^{\gamma} / (e^{2\gamma} + 4 e^{\gamma} + 1)$
1 0 1 1 0 0	$e^{\gamma} / (e^{2\gamma} + 4 e^{\gamma} + 1)$
1 0 0 1 1 0	$e^{\gamma} / (e^{2\gamma} + 4 e^{\gamma} + 1)$
1 0 1 0 1 0	$1 / (e^{2\gamma} + 4 e^{\gamma} + 1)$
c) $y_1 = 1;\ y_6 = 0;\ \Sigma y_t=4;$	
1 1 1 1 0 0	$e^{\gamma} / (e^{\gamma} + 3)$
1 1 1 0 1 0	$1 / (e^{\gamma} + 3)$
1 1 0 1 1 0	$1 / (e^{\gamma} + 3)$
1 0 1 1 1 0	$1 / (e^{\gamma} + 3)$
d) $y_1 = 0;\ y_6 = 1;\ \Sigma y_t=2;$	
0 1 0 0 0 1	$1 / (e^{\gamma} + 3)$
0 0 1 0 0 1	$1 / (e^{\gamma} + 3)$
0 0 0 1 0 1	$1 / (e^{\gamma} + 3)$
0 0 0 0 1 1	$1 / (e^{\gamma} + 3)$
e) $y_1 = 0;\ y_6 = 1;\ \Sigma y_t=3;$	
0 1 1 0 0 1	$e^{\gamma} / (e^{2\gamma} + 4 e^{\gamma} + 1)$
0 1 0 1 0 1	$1 / (e^{2\gamma} + 4 e^{\gamma} + 1)$
0 1 0 0 1 1	$e^{\gamma} / (e^{2\gamma} + 4 e^{\gamma} + 1)$
0 0 1 1 0 1	$e^{\gamma} / (e^{2\gamma} + 4 e^{\gamma} + 1)$
0 0 1 0 1 1	$e^{\gamma} / (e^{2\gamma} + 4 e^{\gamma} + 1)$
0 0 0 1 1 1	$e^{2\gamma} / (e^{2\gamma} + 4 e^{\gamma} + 1)$

f) $y_1 = 0$; $y_6 = 1$; $\Sigma y_t=4$;

0 1 1 1 0 1	$1 \quad / (e^\gamma + 3)$
0 1 1 0 1 1	$1 \quad / (e^\gamma + 3)$
0 1 0 1 1 1	$1 \quad / (e^\gamma + 3)$
0 0 1 1 1 1	$e^\gamma \quad / (e^\gamma + 3)$

g) $y_1 = 0$; $y_6 = 0$; $\Sigma y_t=2$;

0 1 1 0 0 0	$e^\gamma \quad / 3 (e^\gamma + 1)$
0 1 0 1 0 0	$1 \quad / 3 (e^\gamma + 1)$
0 1 0 0 1 0	$1 \quad / 3 (e^\gamma + 1)$
0 0 1 1 0 0	$e^\gamma \quad / 3 (e^\gamma + 1)$
0 0 1 0 1 0	$1 \quad / 3 (e^\gamma + 1)$
0 0 0 1 1 0	$e^\gamma \quad / (3 e^\gamma + 3)$

h) $y_1 = 0$; $y_6 = 0$; $\Sigma y_t=3$;

0 1 1 1 0 0	$e^\gamma \quad / 2 (e^\gamma + 1)$
0 1 0 1 1 0	$1 \quad / 2 (e^\gamma + 1)$
0 1 1 0 1 0	$1 \quad / 2 (e^\gamma + 1)$
0 0 1 1 1 0	$e^\gamma \quad / 2 (e^\gamma + 1)$

i) $y_1 = 1$; $y_6 = 1$; $\Sigma y_t=3$;

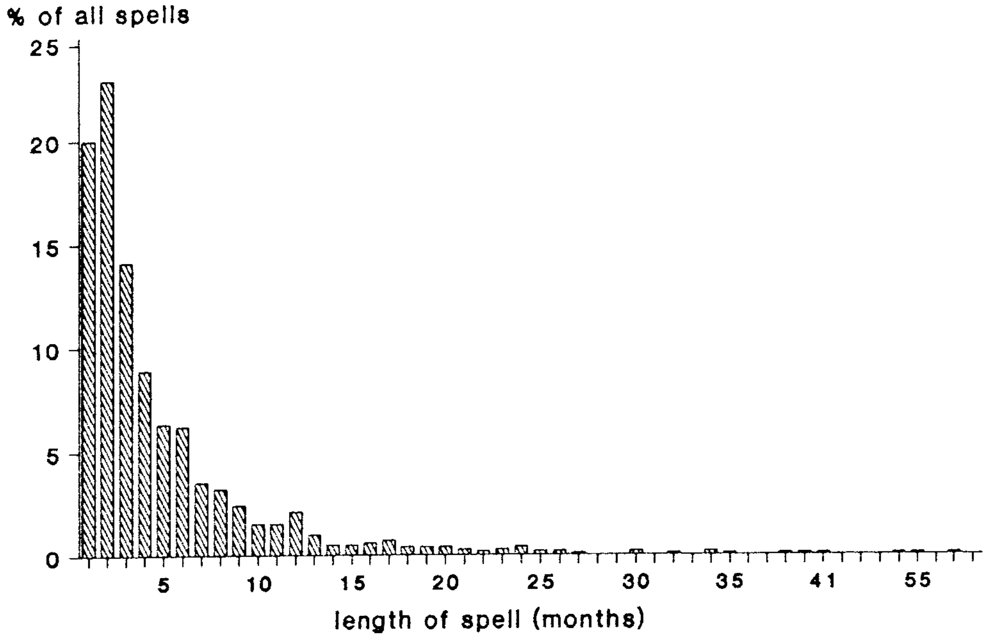
1 1 0 0 0 1	$e^\gamma \quad / 2 (e^\gamma + 1)$
1 0 1 0 0 1	$1 \quad / 2 (e^\gamma + 1)$
1 0 0 1 0 1	$1 \quad / 2 (e^\gamma + 1)$
1 0 0 0 1 1	$e^\gamma \quad / 2 (e^\gamma + 1)$

j) $y_1 = 1$; $y_6 = 1$; $\Sigma y_t=4$;

1 1 1 0 0 1	$e^\gamma \quad / 2 (e^\gamma + 2)$
1 1 0 1 0 1	$1 \quad / 2 (e^\gamma + 2)$
1 1 0 0 1 1	$e^\gamma \quad / 2 (e^\gamma + 2)$
1 0 1 1 0 1	$1 \quad / 2 (e^\gamma + 2)$
1 0 1 0 1 1	$1 \quad / 2 (e^\gamma + 2)$
1 0 0 1 1 0	$1 \quad / 2 (e^\gamma + 2)$

The distribution of durations of all completed spells of unemployment, 1974-81.

(Male minimum-age school leavers who left school in 1974)



The distribution of incidence of all completed spells of unemployment, 1974-81

(Male minimum-age school leavers who left school in 1974)

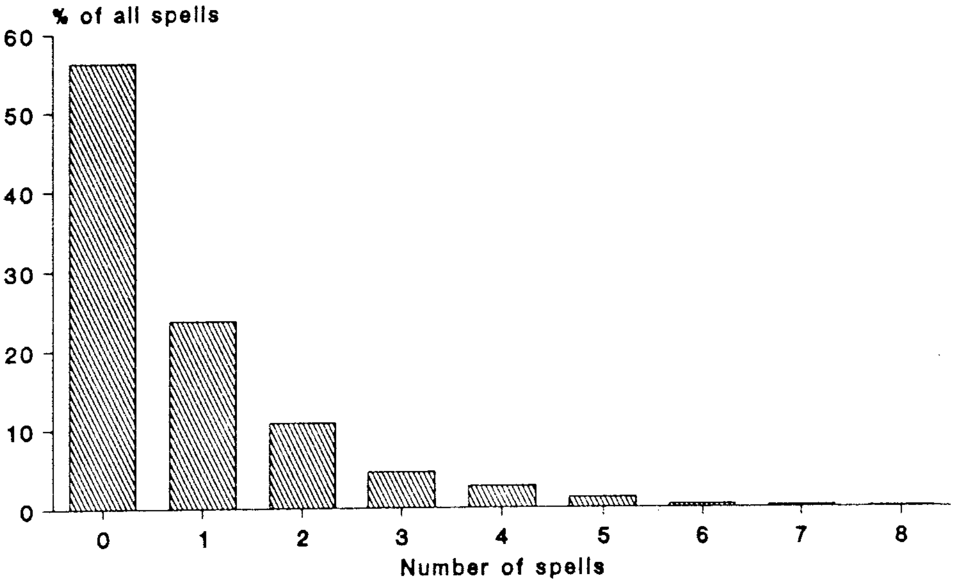


Table 1: Descriptive Statistics of the Sample

The Time Invariant Variables	Mean					
Father's Social Class:						
Professional, Managerial, Intermediate gps	0.12					
Skilled Non-Manual	0.05					
Skilled Manual	0.49					
Semi-Skilled Manual	0.19					
Unskilled Manual	0.15					
Maths Score below average	0.58					
Reading Score below average	0.65					
Birth order is 5 or more	0.08					
The Time varying variables	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Year 6**
Average Unemployment rate*	3.65	5.45	7.26	7.48	7.80	6.98
(standard deviation)	(1.92)	(2.26)	(2.53)	(2.76)	(2.96)	(2.88)
Marital Status	0.00	0.003	0.02	0.06	0.12	0.22
First Child	0.00	0.01	0.02	0.04	0.08	0.11
Second Child	0.00	0.00	0.001	0.003	0.01	0.02
Third Child	0.00	0.00	0.00	0.00	0.001	0.002
Proportion unemployed	0.11	0.05	0.06	0.06	0.07	0.07

* 1974 location travel-to-work area.

Table 2: Descriptive Statistics of the Sample Used in the Conditional Model

Percentages

Sampled months	years	Scored below average on reading test at age 11 yrs.	Scored below average on maths test at age 11 yrs.	Father/guardian was in non-manual occupation	Living in high unemp. area at age 21 yrs.	Birth order was 5 or higher	N
		Sample used in the estimation of $\hat{\gamma}_2$					
July/Aug.	74/79	87.5	87.5	12.5	62.5	18.8	16
Sept./Oct.	74/79	73.7	78.9	10.5	52.6	26.3	19
Nov./Dec.	74/79	69.6	65.2	13.0	39.1	26.1	23
Jan./Feb.	75/80	64.3	71.4	10.7	53.6	32.1	28
March/Apr.	75/80	80.0	86.7	13.3	53.3	30.0	30
May/June	75/80	73.3	70.0	10.0	63.3	23.3	30
		Sample used in the estimation of $\hat{\gamma}_1$					
July/Aug.	74/79	76.6	79.9	10.7	55.1	25.7	214
Sept./Oct.	74/79	76.0	79.7	9.9	57.8	27.6	192
Nov./Dec.	74/79	72.2	73.9	11.7	56.5	24.3	230
Jan./Feb.	75/80	75.5	78.2	11.1	53.7	25.9	216
Mar/Apr.	75/80	71.6	77.6	10.8	53.9	24.1	232
May/June	75/80	77.5	82.4	9.9	55.4	26.1	222
All 16 year old male school-leavers		58.3	65.0	17.0	37.0	9.8	4,067

Note: (i) High unemployment area is defined as unemployment rate in the Travel-To-Work-Area in 1981 being greater than or equal to 15%. The average male unemployment rate in the same period was 13.3%.

(ii) N is the number of individuals.

Table 3: Conditional Maximum Likelihood Estimates
(asymptotic standard errors)

Sampled Months	years	Model 1 $\hat{\gamma}_2$	Model 1 No. of indiv. contrib. to the lik. fn.	Model 2 $\hat{\gamma}_1$	Model 2 No. of indiv. contrib. to the lik. fn.
July/Aug	1974 - 79	0.79 (0.54)	16	1.20 (0.14)	214
Sep/Oct	1974 - 79	- 0.11 (0.46)	19	1.27 (0.15)	192
Nov/Dec	1974 - 79	- 0.44 (0.43)	23	1.20 (0.14)	230
Jan/Feb	1975 - 80	0.14 (0.38)	28	1.14 (0.14)	216
Mrch/April	1975 - 80	- 0.13 (0.37)	30	1.30 (0.14)	232
May/June	1975 - 80	0.13 (0.37)	30	1.24 (0.14)	222

Notes: (i) Total number of individuals in the sample 4,067.
(ii) γ_2 is estimated under the assumption that $\beta = 0$ in eq. (2).
(iii) γ_1 is estimated under the assumption that $\beta = \gamma_2 = 0$ in eq. (2).

Table 4: Marginal Maximum Likelihood Estimates using Normal Heterogeneity for the **July/August** Sample (asymptotic standard errors)

	Model 3	**Model 4**	**Model 5**	**Model 6**	**Model 7**
Intercept	- 4.23 (0.78)	- 3.59 (0.08)	- 4.31 (0.04)	- 4.34 (0.12)	- 5.96 (0.27)
The time-varying variables:					
Unemp. rate.		0.07 (0.01)	0.02 (0.01)	0.09 (0.01)	0.15 (0.02)
Married		-0.51 (0.12)	-0.50 (0.09)	-0.32 (0.13)	-0.43 (0.17)
First child		0.77 (0.13)	0.74 (0.10)	0.68 (0.14)	0.92 (0.20)
Second child		0.70 (0.25)	0.75 (0.21)	0.62 (0.27)	0.79 (0.35)
Third child		0.67 (0.73)	0.12 (0.63)	0.85 (0.76)	0.68 (0.98)
The time invariant variables:					
Father's Social class*					
1.Professional, Managerial or Intermediate gps (nonmanual)	- 0.53 (0.20)	- 0.42 (0.10)	- 0.50 (0.10)	- 0.38 (0.14)	- 0.54 (0.20)
2.Skilled non-manual	- 0.49 (0.28)	- 0.50 (0.15)	- 0.50 (0.15)	- 0.41 (0.20)	- 0.49 (0.29)
3.Semi-skilled manual	0.20 (0.14)	0.06 (0.07)	0.12 (0.07)	0.10 (0.09)	0.16 (0.14)
4.Unskilled manual	0.82 (0.15)	0.46 (0.07)	0.66 (0.07)	0.46 (0.09)	0.66 (0.15)
Maths Score below average	0.51 (0.15)	0.33 (0.07)	0.38 (0.04)	0.44 (0.10)	0.59 (0.14)
Reading Score below average	0.49 (0.13)	0.38 (0.07)	0.47 (0.04)	0.32 (0.09)	0.47 (0.13)
Birth order is 5 or more	0.10 (0.20)	0.14 (0.09)	0.24 (0.10)	-0.02 (0.13)	0.05 (0.20)
y_{it-1}	0.92 (0.15)			1.99 (0.09)	0.83 (0.15)
y_{it-2}	0.10 (0.15)			0.68 (0.10)	-0.01 (0.15)
σ	1.46 (0.38)		1.53 (0.03)		1.66 (0.14)
ψ_0	0.30 (0.46)		0.00		0.00
ψ_1	0.004 (0.003)		0.001 (0.0005)		0.003 (0.0015)
P_0	0.23		0.00		0.00
P_1	0.003		0.001		0.003
T	4	6	6	4	4
N	4067	4067	4067	4067	4067
- ln L	3246.8	5983.1	5538.1	3281.8	3192.3

Notes: (i) σ = the standard deviation of the Normal heterogeneity distribution.

(ii) P_0 = the estimated proportion of stayers in state 0 (i.e. non-unemployment).

(iii) P_1 = the estimated proportion of stayers in state 1 (i.e. unemployment).

* Omitted category is the skilled manual.

Table 5: Marginal Maximum Likelihood Estimates using Mass-points for the **July/August** Sample (asymptotic standard errors)

	Model 8	**Model 9**
The time-varying variables:		
Unemp. rate.	0.14 (0.02)	0.15 (0.02)
Married	- 0.36 (0.16)	- 0.43 (0.17)
First child	0.99 (0.18)	0.99 (0.20)
Second child	0.74 (0.29)	0.84 (0.30)
Third child	0.70 (0.94)	0.58 (0.95)
The time invariant variables:		
Father's Social class*		
1.Professional, Managerial or Intermediate gps (nonmanual)	- 0.53 (0.18)	- 0.60 (0.21)
2.Skilled non-manual	- 0.57 (0.28)	- 0.62 (0.31)
3.Semi-skilled manual	0.14 (0.13)	0.13 (0.15)
4.Unskilled manual	0.57 (0.12)	0.63 (0.14)
Maths Score		
below average	0.53 (0.14)	0.59 (0.07)
Reading Score		
below average	0.42 (0.13)	0.45 (0.14)
Birth order is 5 or more	- 0.01 (0.18)	0.04 (0.20)
y_{it-1}	1.04 (0.13)	0.85 (0.15)
y_{it-2}	0.10 (0.14)	- 0.02 (0.15)
Mass points		
Location 1	- 5.51 (0.19)	- 5.92 (0.30)
Probability	0.93 (fixed)	0.85 (0.35)
Location 2	- 2.31 (0.21)	- 3.33 (0.39)
Probability	0.07 (0.16)	0.14 (fixed)
Location 3		- 0.66 (0.60)
Probability		0.01 (0.43)
Mean of the Mixing Distrib.	- 5.19	- 5.49
Std. deviation of the Mix. dist.	0.81	1.06
T	4	4
N	4067	4067
- ln L	- 3199.9	- 3192.4

Notes: See Table 3 notes.