Working Paper 98-53
Economics Series 19
July 1998

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-98-75

SELF-EMPLOYMENT AND LABOUR MARKET TRANSITIONS: A MULTIPLE STATE MODEL

Maite Martínez-Granado *

Abstract

In this paper we estimate a multiple state transition model for the UK describing transitions in and out three possible labour market states: self-employment, paid employment and unemployment. This enables us to assess the effect of demographic characteristics as well as time changing economics conditions on the probabilities of exiting and entering each different state. A reduced form model is estimated allowing for unobservable individual heterogeneity.

The results are consistent with the hypothesis of a deterioration of the labour market conditions generating an increase in the self-employment rates in adverse economic conditions. However unemployment duration generates a loss on human capital that reduce the probabilities of switching to self-employment. It appears also that family background and education play and important role in determining the transition probabilities. Medium level educated individuals are the most likely to become self-employed.

Keywords: self-employment, duration, heterogeneity, economic conditions

* Martínez-Granado, Departamento de Economía, Universidad Carlos III de Madrid. E-mail: mmartine@eco.uc3m.es

I would like to thank Richard Blundell, Adriaan Kalwij, Costas Meghir, Juan Sanchis-Llopis and Frank Windmeijer for helpful comments. I am also grateful to the Bank of Spain for financial support. The data used in this paper were made available through tue ESRC Data Archive. The data were originally colledtec by the ESRC Centre on Micro-social Change at the University of Essex. Neither the original collectors of the data nor the Archive bear any responsability for the analyses or interpretatios presented here. This paper is part of the thesis "Analysing labour market mobility: some empirical aplications" presented as fulfilment of the PhD program at University College London.

1. Introduction

Self-employment in the UK (and in general in all OECD countries) has experienced a sharp increase during the late 70's and 80's. In March 1980, 10% of the workforce was self-employed. This figure increased to 15% in March 1989, according to the Department of Employment. Meanwhile, programs to promote and support start-up and expansion of small businesses have been carried out by the government. In spite of the growth of self-employment and of the policies to promote it, there is little empirical evidence, particularly for the UK, of the characteristics that motivate an individual to become self-employed.

The purpose of this paper is to add some empirical evidence of the characteristics and economic factors that determine self-employment decisions in UK. It has been recently argued that self-employment growth is the result of a labour market deterioration. Difficulties in finding a job generate the increases on self-employment figures rather than changes in personal characteristics, such as education, financial conditions or even unobserved skills. Consequently, an interesting question to answer is to analyse whether the probabilities of transition into self-employment are higher for unemployed people than for employees, after controlling for personal characteristics.

Most of the related empirical work relies on the hypothesis that capital markets are perfect and any individual can borrow and lend any amount of money at current interest rates.[1] In this context, self-employment is considered as an alternative to paid employment. These studies have usually been focused on the differential between expected earnings and wages. The econometrician observes whether an individual is self-employed or not. Thus, the usual approach is to estimate a structural or reduced form binary choice model (probit or logit). This is the approach followed by Ress and

[1] See Evans and Jovanovic (1989), Holtz-Eakin et al. (1994a, 1994b) or Blanchflower and Oswald (1990) for an analysis of self-employment decisions under liquidity constraints. The first two papers refer to the American labour market while the third one relates the UK.

Shah (1985) and more recently by Blundell et al. (1995) using British data. They also consider the influence of demographic characteristics on the probability of becoming self-employed, namely education, marital status, region, race, number of children and health. Evans and Leighton (1989) use the same approach for the United States.

The effect of unemployment on the probability of becoming self-employed is not clear in the literature, though. Two opposite results have been found. On the one side Blanchflower and Oswald (1991) or Taylor (1996) provide evidence for UK, supporting a negative relationship between unemployment rates and entering self-employment. Good economic conditions (lower risk of failure or higher probability of finding an alternative job in the event of failure) would encourage individuals to start their own business. On the other side, Alba-Ramirez and Freeman (1994) or Evans and Leighton (1989) find, that the longer one individual has been unemployed the higher is his probability of becoming self-employed for Spain and the US respectively. Individuals see self-employment as a way of avoiding unemployment when the probability of finding a paid job decreases[2]. Acs *et al.* (1994), using a panel of OECD countries, find that a 10% increase in the unemployment rate produces a 1.5% increase in the self-employment rate.

Previous empirical work has several limitations. On the one side, some authors, in estimating a binary choice model upon the stock of self-employed individuals, mix up entry and exit decisions (as Blanchflower and Oswald (1991) or Taylor (1996)). If unemployment leads individuals more likely to enter self-employment but also to exit that state, the final effect is going to be a mixture of both. Its sign will depend on the relative inflows and outflows.

To avoid this problem Evans and Leighton (1989) or Alba-Ramírez and Freeman (1994) constraint their studies to a particular group of individuals, wage workers, and look at exit rates towards self-employment. Only for this subgroup their results hold. The effects can be different for unemployed individuals; if working is an

[2] All these papers use the approach discussed above.

endogenous decision, as it is, results can be subjected to sample selection bias. In addition, some variables as previous unemployment experience, are likely to be endogenous. None of the previous studies addresses the analysis of the effect of general economic conditions jointly with individual unemployment spells.

A natural approach to model self-employment decisions consists of looking at individual work histories and considering self-employment an additional labour market state. Transitions among different states can be constructed and analysed[3] then. In this paper three possible labour market states would be considered: self-employment, paid employment and unemployment. We estimate reduced form parametric transition probabilities from and to any of the three possible states, using a multiple state transition econometric framework. Due to tastes' differences or ability or what can be called "entrepreneurial spirit", the presence of unobservable heterogeneity among individuals seems quite likely in this context. As a result we also estimate the model under this hypothesis in order to compare the results.

This approach helps us to overcome some of the limitations of previous literature. First, it allows us to consider entry and exit decisions separately. Unemployment duration arises naturally in this framework, avoiding possible endogeneity and selection bias. In addition it enables comparisons of the probability of becoming self-employed between unemployed and employed individuals.

Data used in this paper is a subsample of males drawn from the British Household Panel Survey (BHPS). This survey contains a retrospective work history questionnaire recording all job spells for each individual in the house since they left school.

The paper outline is as follows. Section 2 describes the model specification and discusses the estimation of the transition probabilities. Section 3 presents the data used

[3] Not many applications refer to transitions between labour market states including self-employment applications. Magnac and Robin (1991) estimate a reduced-form model of labour market transitions using discrete and tenure data for France. The aim of their paper is closely related to our paper, with the difference that continuous records are available for the UK, which allows us to construct a more complex model.

for the analysis and in section 4 the estimation results are discussed. Section 5 states the main conclusions of the analysis.

2. Model specification

We distinguish three different possible states in our model: unemployment, wage-work and self-employment. Self-employment is considered as an alternative to paid employment. Movements from and to "out of the labour force" are not considered here. An individual can move from any of these states (source state, denoted by the first subscript) to the others (destination state, denoted by the second subscript) at any time. Six types of transition can then be defined as shown in Table 1 below.

This specification of the model is coherent with on-the-job search theories. Unemployed individuals devote some of their time searching for jobs. But, once they accept a job and start working (as paid workers or self-employed) they continue to search for a better job. Two types of "jobs" are considered here: paid work and self-employment[4].

Transition intensities are defined as the probability of departure from state k to state l in the short interval $(t, t+\partial t)$ and are denoted as $\theta_{kl}(t|Z;\beta)$, where Z is a set of observable and unobservable individual characteristics (X and v respectively) and β is a set of unknown parameters to estimate. t, the elapsed duration, is measured in months. In particular, the following functional form represents the transition intensities:

$$\theta_{kl}(t|Z;\beta) = \exp\{g_{kl}(t) + X\beta_{kl} + \delta_{kl}v\} \qquad (1)$$

where $g_{kl}(t)$ is a function of time spent in state k, before departure towards l. This specification allows for a flexible and non-monotonic relation between elapsed duration and the hazard function. Its functional form will be discussed in Section 4.2.

[4] Blundell *et al.* (1995) use a similar approach to analyse upwards and self-employment transitions.

The set β, includes all parameters of interest in $g(.)$, β_{kl} and δ_{kl}, for all possible (k,l). An unobservable individual fixed effect is denoted by v which would be correlated with the time spent in each state. It can reflect differences in tastes for working or starting up a business. The estimation of parameters specific to every state allows state dependence along with duration dependence. Finally, X is a set of demand conditions and demographic variables.

Table 1: Possible transition intensities.

	Destination State		
Source State	Self-Employment	Employment	Unemployment
Self-Employment	------	$\theta_{see}(t\|Z;\beta)$	$\theta_{seu}(t\|Z;\beta)$
Employment	$\theta_{ese}(t\|Z;\beta)$	------	$\theta_{eu}(t\|Z;\beta)$
Unemployment	$\theta_{use}(t\|Z;\beta)$	$\theta_{ue}(t\|Z;\beta)$	------

Note: U denotes unemployment, E paid employment and SE self-employment.

Therefore the contribution to the likelihood function for each individual and completed spell is the probability of surviving in state k until t (survival function) times the probability of moving from k to l in t (transition intensity),

$$P_{kl}(t|Z;\beta) = \exp\{-\Theta_k(t|Z;\beta)\}\,\theta_{kl}(t|Z;\beta) \qquad (2)$$

where Θ_k is the corresponding integrated hazard function ($\Theta_{kl} = \int_0^{t_l} \sum_{l \neq k} \theta_{kl}(s|Z;\beta)\,ds$).

For each individual, the data consist of one or more spells in every state. Not every spell is complete by the time of the interview. Hence it is necessary to account for right censored spells. The contribution to the likelihood function of an incomplete spell is the survivor function, that is

$$\overline{F}_k(t|X;\beta) = \exp\{-\Theta_k(t|X;\beta)\} \tag{3}$$

Assuming that v equals zero for all individuals i. e. there is no unobserved heterogeneity, the likelihood function for an individual with a sequence of spells $\{t_1, t_2,....,t_{Ci}\}$, would be,

$$L_i(\beta|t_{i1},...,t_{iC_i}) = \left(\prod_{c=1}^{C_i}\prod_{k}\prod_{l\neq k} P_{kl}(t_c|X_i^c;\beta)^{d_{kl}^c}\right)\left(\prod_{c=1}^{C_i}\prod_{l}\overline{F}_k(t_c|X_i;\beta)^{s_k^c}\right) \tag{4}$$

where d_{kl}^c is an indicator variable which equals 1 if the individual exited state k towards state l in the cth spell; s_k^c is a dummy which equals one if the cth spell is incomplete and the individual did not move from state k.

Taking logs and considering a sample of N i.i.d. individuals the log-likelihood function is given by[5]

$$\log L = \sum_{i=1}^{N}\sum_{c=1}^{C_i}\sum_{k}\left[\left(\sum_{l\neq k} d_{kl}^c P_{kl}(t_c|X_i;\beta)\right) + s_k^c \overline{F}(t_c|X_i;\beta)\right] \tag{5}$$

In the presence of unobservable heterogeneity among the individuals the model becomes more complicated. The individual fixed effect, denoted by v_i, is an unobservable variable that varies over the population. Therefore, we cannot condition the individual probabilities on v_i and use it as an additional explanatory variable. To get the unconditional probabilities it is necessary to integrate v_i over all its possible values. In this case the individual likelihood takes the form

[5] For a step-by-step derivation of the likelihood function with and without unobserved heterogeneity, see Lancaster(1990)

$$L_i(\beta|t_{i1},\ldots,t_{iC_i},X_i) =$$

$$\int_{-\infty}^{\infty}\left\{\left(\prod_{c=1}^{C_i}\prod_{k}\prod_{l\neq k}P_{kl}(t_c|X_i^c;\beta)^{d_{kl}^c}\right)\left(\prod_{c=1}^{C_i}\prod_{l}\overline{F}_k(t_c|X_i;\beta)^{s_k^c}\right)\right\}h(v_i)\,dv_i \tag{6}$$

where $h(v_i)$ is the unknown distribution function of the individual effect. The log-likelihood function for all individuals would then be,

$$\log L = \sum_{i=1}^{N} L_i(\beta|t_{i1},\ldots,t_{iC_i},X_i) \tag{7}$$

The distribution of the unobserved heterogeneity could be fully specified and the previous equation estimated by maximum likelihood. However, Heckman and Singer (1984) pointed out that misleading results can be obtain by using these procedures when the chosen distribution for unobservables is not the right one. Therefore we alternatively use the Non-Parametric Maximum Likelihood Estimator (NPMLE) proposed by both authors which does not require any distributional assumption. This procedure approximates the distribution function of unobservables, $h(v)$, with a finite mixture distribution. The points of support of this finite distribution are the unknown values $v_1,\ldots,v_M$ to which the M unknown probabilities are attached. Then, the contribution to the likelihood of an individual becomes:

$$L_i(\beta,v,\pi|t_{i1},\ldots,t_{iC_i},X_i) =$$

$$\sum_{m=1}^{M}\left\{\left(\prod_{c=1}^{C_i}\prod_{k}\prod_{l\neq k}P_{kl}(t_c|X_i^c,v_m;\beta)^{d_{kl}^c}\right)\left(\prod_{c=1}^{C_i}\prod_{l}\overline{F}_k(t_c|X_i,v_m;\beta)^{s_k^c}\right)\right\}\pi_m \tag{8}$$

being the log-likelihood function its summation over all individuals. The points of support as well as the probabilities assigned to each of them are now parameters of interest to be estimated by the EM-algorithm (see Appendix B for description of

implementation). The function is maximised at different number of support points until the parameters of the criterion function relatively stable[6].

3. Data Description

The data used in this analysis is obtained from the British Household Panel Survey (BHPS). This is an annual survey carried out by the ESRC Research Centre on Micro-social Change since 1991. At the moment of starting the research, three data waves are available. The survey is conducted over a nationally representative sample of at least 5000 households, making a total of approximately 10000 individual interviews. Data is collected at an individual and household level including information about household organisation, labour market, income and wealth, housing, health and socio-economic values.

The Second Wave (1992) contains some additional records that do not appear in the First Wave relating individual's past history: marriage, cohabitation, children and employment status. In particular, it collects information about employment status spells since the respondent first left full time education. The dates at which each spell began and ended as well as its length are recorded. This information enables us to estimate the model proposed in Section 2. Demographic information can be obtained from the main record.

We select a subsample of working age males at the interview date, that is aged between 16 and 65 years old by the first of December of 1992. Males who were not directly interviewed (somebody else in the household answered the questionnaire on their behalf) are not considered here: no answers for the additional questionnaires are provided for these individuals. We also dropped men who were out of the labour market (full time students, retired or out of the labour market for other reasons) at some

[6] The basic specification includes two support points. A possible interpretation for this model specification is to think of two individual types: high ability and low ability ones. Those support points can also represent aggregate shocks with different effects over the population.

point in their work histories. That avoids initial condition problems and also mixing decisions of early retirement. No differences are made between full time and part time work: both are considered employment. Within paid employment, no job-to-job transitions are considered.

All the previous conditions are fulfilled by 1978 individuals providing 4227 complete or incomplete employment status spells. Table 2 reports the number of observations for each possible transition, being the last spell for each individual incomplete.

The variables used in the estimation can be classified in two groups: demographic variables relating the individual and demand side variables referring to general economic conditions. In the first group, we include age at the beginning of the spell, four educational dummies and two dummy variables reflecting the family background of the individual: whether his mother and his father were self-employed when the individual was fourteen years old. In the second group, the final specification includes the National Unemployment Rate at the beginning of the spell that accounts for business cycle changes[7]. Variables as vacancies or GDP were also tried but were not found significant so that they are not included in the final specification. Those figures have been taken from the Department of Employment. Table A.1 in Appendix A reports the mean and standard deviation for the relevant variables.

Table 2: Number of observations per possible transition.

	Destination State			
Source State	Self-Employment	Employment	Unemployment	Censored
Self-Employment	------	138	75	335
Employment	408	------	791	1467
Unemployment	91	746	------	176

[7] The probability of having a successful enterprise is not the same if it was started in a good or bad period of the economic cycle.

4. Empirical Results

4.1. Non-Parametric Analysis

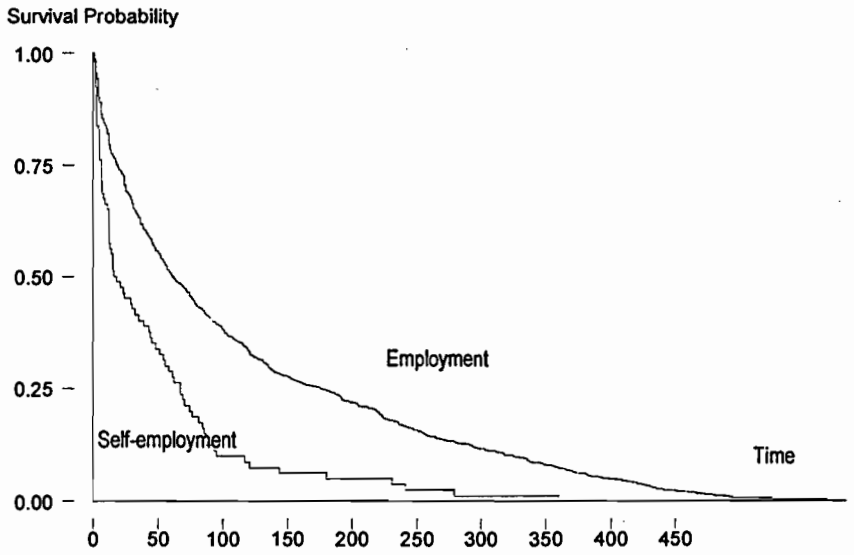
Before presenting the model estimates we describe survival probabilities in the labour market using non-parametric techniques. Figures 1 to 3 show Kaplan-Meier estimates of the survival probabilities for each possible source state conditioned on destination state.

Figure 1 shows survival probabilities in employment and unemployment for individuals who move to self-employment. Figure 2 presents the survival probabilities in self-employment and unemployment for individuals who move to employment. The pattern in both figures is similar: the probability of survival is lower if the origin state is unemployment, as expected. This just shows that unemployment is not individually considered as a definitive state. Nevertheless, it is worth to note that the survival probability decreases faster for self-employed individuals than for employees. There are different explanations for this finding.

Figure 1. Kaplan-Meier Survival curves: transition to self-employment from unemployment and paid employment

Figure 2. Kaplan-Meier Survival curves: transition to paid employment from unemployment and self-employment

Figure 3. Kaplan-Meier Survival curves: transition to unemployment from self-employment and paid employment



Note: the survival function is computed as $\overline{F}(t) = 1 - \frac{\text{number of individuals leaving before t}}{N}$.

First, it can reflect the higher risk of self-employment, being more difficult to survive. Second, survival in employment would be higher in the presence of liquidity constraints to allow the individual to earn the capital necessary to start a business. Finally, the idea of a deterioration of the labour market could be supported by this data, in the sense that self-employment is used as a temporary state better than being unemployed, before jumping again into paid employment.

Figure 3 shows the survival probabilities for unemployed people and displays a steeper curve for those who came from self-employment, suggesting again that self-employment is used to avoid unemployment when finding a job becomes difficult.

This analysis does not take into account either personal or demand side characteristics. Determining whether the stated differences can be explained by differences in characteristics is the next point to discuss.

4.2. Parametric estimation.

The previous analysis clearly shows that duration and state dependence are two important factors to explain mobility between different states. The longer an enterprise has been running the higher its probability of survival, through a reputation effect or because it has access to more resources than when it first started. The duration dependence comes through a tenure or experience effect for employess and through a loss in human capital for those unemployed. This is why in Section 2 our model includes a flexible function of elapsed duration, $h_{kl}(t)$, to control for duration dependence. We include the log and the log of tenure square[8] as regressors:

$$h_{kl}(t) = \alpha_{1kl} \ln(t) + \alpha_{2kl} (\ln(t))^2$$

[8] Alternative specifications were tried, including splines for the square term and dummies for duration shorter than 6 months or greater than one year, but they seemed to fit worse the data.

This specification generalises the traditional Weibull proportional hazard allowing non monotonic variation with respect to duration. So if $\alpha_{2kl}<0$, the transition intensity has a maximum level when $\ln(t)=-\alpha_{1kl}/\alpha_{2lk}$. If $\alpha_{2kl}>0$, this level corresponds to a minimum and if $\alpha_{2kl}=0$ for all k and l, the transition intensities are a monotonically increasing or decreasing function in t (Weibull specification).

The specification also includes three types of explanatory variables: demographic (four educational variables, age and age square), family background (two dummy variables taking value one if each of the individual parents were self-employed when he was 14) and demand side conditions (National Unemployment Rate, NUR). Education in this context can act jointly with duration as a proxy for the individual wage. The unemployment rate tries to pick up changes in the general economic conditions altering the probabilities of layoff and job arrival rates and therefore the individuals' possibility of choice.

Unobserved heterogeneity is explicitly modelled accounting for the differences in ability among individuals and it is uncorrelated with the rest of explanatory variables. Two types of individuals are considered and therefore two support points are used in the estimation of the model proposed in Section 2.

Tables 3 and 4 show the results of this basic specification without and with corrections for unobserved heterogeneity, respectively. Estimates of the parameters of interest are similar in both cases, with the biggest differences lying in the duration parameters. In general, they are overestimated when ignoring unobservable heterogeneity. In what follows we would refer to Table 4, although the same conclusions can be drawn from Table 3[9].

Education seems to play an important role in determining transitions between states. As it should be expected, the more educated an individual is the lower his probability of becoming unemployed. It also is interesting to point out that people with a medium level of education (A-levels and O-levels) are more likely to become self-

[9] A likelihood ration test of the joint hypothesis that all parameters related to heterogeneity are zero give us a value of 614, distributed as χ^2 with 10 d.o.f. The null hypothesis is clearly rejected.

employed, whether they come from unemployment or paid employment. Moreover they are the less likely to become unemployed once in self-employment. This contradicts to some extend previous findings that suggest that self-employed people are poor wage earners and misfit for paid work[10].

Table 3. Maximum Likelihood estimates for the transition equations; without controls for unobserved heterogeneity.

	E to U	E to SE	U to E	U to SE	SE to E	SE to U
High Degree	-0.686	-0.117	0.779	0.648	0.501	-0.424
	(0.157)	(0.208)	(0.172)	(0.529)	(0.378)	(0.456)
A-Levels	-0.384	0.367	0.769	0.977	0.422	-1.038
	(0.195)	(0.228)	(0.218)	(0.631)	(0.414)	(0.640)
O-Levels	-0.154	0.211	0.602	0.872	0.118	-0.420
	(0.156)	(0.207)	(0.193)	(0.561)	(0.386)	(0.453)
Other qualific.	-0.289	0.125	-0.122	-0.165	0.511	-0.851
	(0.306)	(0.328)	(0.404)	(1.392)	(0.643)	(3.293)
Mother SE	0.057	0.333	-0.157	0.503	0.549	-0.067
	(0.265)	(0.274)	(0.378)	(0.583)	(0.387)	(1.226)
Father SE	0.071	0.583	-0.126	0.549	-0.356	0.193
	(0.159)	(0.168)	(0.177)	(0.360)	(0.316)	(0.383)
Age	0.341	1.729	-0.241	1.969	-0.839	-0.404
	(0.382)	(0.771)	(0.395)	(1.329)	(0.988)	(1.437)
Age Squared	-0.050	-0.281	0.010	-0.242	0.080	0.018
	(0.058)	(0.144)	(0.055)	(0.178)	(0.156)	(0.230)
NUR	0.165	0.039	-0.028	0.159	0.025	0.177
	(0.020)	(0.028)	(0.026)	(0.074)	(0.038)	(0.064)
ln(duration/12)	-0.143	0.328	-0.617	-0.401	0.134	-0.364
	(0.084)	(0.127)	(0.114)	(0.252)	(0.201)	(0.159)
$(\ln(\text{duration}/12))^2$	0.050	-0.044	-0.246	-0.188	-0.262	-0.077
	(0.028)	(0.040)	(0.051)	(0.130)	(0.079)	(0.100)
Intercept	-4.604	-7.415	0.186	-7.729	-1.565	-3.311
	(0.552)	(0.941)	(0.680)	(2.379)	(1.496)	(2.052)
Log-likelihood			-7777			
Observations			4227			

Note: Standard errors (computed from the inverse of the information matrix) in brackets. SE denotes self-employment, E employment and U unemployment. Age is measured at the beginning of the spell. NUR is National Unemployment Rate.

[10] See Ress and Shah (1985) or Evans and Leighton (1989).

It is interesting to point out that the higher educated individuals are reluctant to become self-employed if they are actually employed. They are more willing to do so once they are unemployed (High Degree has a positive and significant effect on the transition probability from unemployment to self-employment). This can reflect the higher opportunity cost (in terms of wages) that this group of individuals face.

Table 4. Maximum Likelihood estimates for the transition equations; controlling for unobserved heterogeneity (NPMLE)

	E to U	E to SE	U to E	U to SE	SE to E	SE to U
High Degree	-0.620	-0.144	0.773	0.648	0.494	-0.318
	(0.114)	(0.156)	(0.108)	(0.312)	(0.264)	(0.374)
A-Levels	-0.340	0.340	0.775	0.982	0.411	-0.975
	(0.146)	(0.142)	(0.133)	(0.435)	(0.278)	(0.534)
O-Levels	-0.085	0.152	0.606	0.860	0.079	-0.148
	(0.109)	(0.155)	(0.109)	(0.291)	(0.271)	(0.418)
Other qualific.	-0.226	0.088	-0.115	-0.169	0.498	-0.835
	(0.218)	(0.288)	(0.215)	(0.818)	(0.414)	(1.539)
Mother SE	0.071	0.340	-0.187	0.501	0.563	-0.067
	(0.210)	(0.249)	(0.263)	(0.433)	(0.319)	(0.663)
Father SE	0.073	0.592	-0.116	0.522	-0.366	0.312
	(0.120)	(0.139)	(0.127)	(0.297)	(0.236)	(0.346)
Age	0.325	1.476	-0.214	1.982	-0.982	0.049
	(0.277)	(0.644)	(0.235)	(0.829)	(0.829)	(1.379)
Age Squared	-0.055	-0.224	0.007	-0.245	0.099	-0.047
	(0.043)	(0.117)	(0.034)	(0.116)	(0.132)	(0.222)
NUR	0.163	0.043	-0.029	0.164	0.030	0.143
	(0.014)	(0.021)	(0.015)	(0.061)	(0.30)	(0.059)
ln(duration/12)	-0.124	0.293	-0.662	-0.413	0.123	-0.325
	(0.040)	(0.120)	(0.062)	(0.154)	(0.169)	(0.151)
$(\ln(\text{duration}/12))^2$	0.048	-0.036	-0.246	-0.188	-0.262	-0.077
	(0.016)	(0.037)	(0.028)	(0.087)	(0.072)	(0.082)
Unobs.heter.	1.000	-0.715	-0.221	0.696	-0.290	2.176
	----	(0.268)	(0.153)	(1.279)	(0.249)	(1.427)
Intercept	-4.198	-7.791	0.093	-7.705	-1.656	-3.041
	(0.394)	(0.842)	(0.382)	(1.418)	(1.210)	(2.034)
Log-likelihood	-7470					
Observations	4227					

Notes: Standard errors (computed from the inverse of the information matrix) in brackets. SE denotes self-employment, E employment and U unemployment. Age is measured at the beginning of the spell. NUR is National Unemployment Rate.

Two support Points: $v_1=0$ with probability $p_1=0.57$ and $v_2=-1.691(0.363)$ with probability $p_2=0.43$. The heterogeneity coefficient for the transition from E to U is normalise to one for identification.

The effect of age differs depending on the transition considered. In general, it is not significant but it positively affects (a decreasing rate) the probability of becoming self-employed. This effect is higher for those individuals that come from unemployment. This result is again coherent with the theory of liquidity constraints. An individual needs to have some wealth before starting up a business. If he is unemployed he would need a longer period of time to achieve this goal.

Family background variables also have the expected sign. The probability of becoming self-employed is higher if one of the parents was self-employed (especially the father). Whether this happens because the individual keeps on running a family business or due to the effect of the environment in which he grew up, can not be separated given the characteristics of the data.

The effect of the national unemployment rate suggests that individuals are more likely to move towards self-employment when the economic situation worsens. This effect is much bigger if the individual is unemployed. The chances of getting a job decrease and the individual chooses to avoid unemployment by becoming self-employed. Therefore, the data seems to support to some extend the deterioration of the labour market theory as an explanation of the increase in self-employment rates.

Regarding elapsed duration variables, the null hypothesis of a monotone specification for the hazard rate is rejected (the Wald test statistic is 606.59 and distributed as a $\chi^2(6)$). The variables have also the expected sign. The longer an individual stays unemployed the lower the probabilities he has of moving towards employment or self-employment due to the loss of human capital. Employment duration has a positive effect on the probability of becoming self-employed at a decreasing rate although not significant. This result it is in line with the age effect (pointing to liquidity constraints). It could also show that people who have been longer in the labour market have more chances of picking up possibilities of starting a business. Tenure in self-employment increases the probability of moving towards employment though the effect is smaller and less precise than the previous one. On the

other hand, it reduces the probability of unemployment. We would expect that the longer an enterprise has been settled the higher its chances of survival and therefore the lower the probability that the individual ends up being unemployed.

Referring to the unobservable heterogeneity, one possible interpretation of the results (given the restrictions on the estimation procedure and the sign of the estimates) would be that if the individual effect is zero, the individual would have low ability, and if it is negative the individual would have high ability. Then, able individuals are more likely to become employed or self-employed and less likely to become unemployed.

An alternative specification is presented in Tables B.1 and B.2 allowing for lagged duration dependence[11]. The basic results hold though the precision of some of the estimates worsens a little. The most interesting result (with and without controls for heterogeneity) is that the previous self-employment experience pays, in the sense of increasing the probability of becoming self-employed, only if the individual is employed. If the individual is unemployed there is no effect, which again support the idea that individuals enter self-employment to avoid unemployment.

It is also important the fact that previous unemployment experience has a positive effect in the entry to self-employment if the individual is employed not if he is unemployed. This is exactly what Evans and Leighton (1989) found with a sample of workers from USA. Previous unemployment experience has in addition a positive effect on the transition from self-employment to employment. That suggests again that individuals coming from unemployment use self-employment as a temporary situation or they are less successful than the rest of self-employed workers.

Results from Tables 3 and 4 can be quite difficult to interpret in terms of transition probabilities from one state to other. Therefore Figures 4 to 8 highlight to what extent the transition intensities differ between individuals with different characteristics. Figure 4 shows the transitions into self-employment by source state and individual type. Employed are more likely to become self-employed than

[11] Flinn and Heckman (1982) point out that lagged spell values can be introduced as explanatory variables.

unemployed individuals with the same characteristics and in both cases this probability decreases with duration. This could be originated either by the presence of liquidity constraints (some capital is necessary to start a business and it is more difficult to earn the money if unemployed) or by the fact that employed people have more information about business opportunities. An interesting effect is that the more able individuals are more likely to enter self-employment but only if the source state is employment. If they are unemployed less able individuals are the ones with more chances to enter self-employment. Deterioration of labour market conditions could originate such an effect. Among the unemployed, the more able ones would find a job in paid employment, whereas the less able would not find any and therefore choose to be self-employed to avoid unemployment.

Figures 5 and 6 show transitions into self-employment for employed and unemployed people by education. Behaviour patterns are very different. Employed individuals have the lower probability of becoming self-employed during the first three years on their job. Afterwards their probability is the highest compared to the rest of the groups. For unemployed people, the transition probability curves are inversely related to the level of education. If the source state is unemployment, it is true that the less qualified individuals and therefore the ones who were probably receiving low wages before, are the ones switching to self-employment[12]. This finding does not hold for individuals whose source state is employment.

To conclude, Figures 7 and 8 show transition probabilities into self-employment by source state, ability of the individual and National Unemployment Rate (NUR). The purpose of these figures is to clarify the effect of general economic conditions on the probability of becoming self-employed. We take the NUR of late seventies (around 3%) and the corresponding to late eighties (around 10%) to see if that change can explain to some extent the high increase in self-employed population during the last twenty years. Figure 7 shows the transition probabilities for employed

[12] See Evans and Leighton(1989).

people. It is important to note that the probability of becoming self-employed increases in a higher proportion for the more able individuals, given an increase in the unemployment rate. On the contrary, for unemployed individuals (Figure 8) the increase in the probability of self-employment is much higher for those less able. For both groups the increment in the probability is quite high if compare for example with the effect of education.

5. Conclusions

This paper describes in some detail transitions from and to paid work, self-employment and unemployment. We use a multiple state transition model with unobservable heterogeneity to asses the importance of some demographic variables along with time varying economic conditions. Distribution of unobservables is approximated with a discrete function, whose support points and probabilities are computed using the Heckman and Singer approach through an EM algorithm.

The main purpose of the paper is to determine the effect of unemployment on the probability of becoming self-employed. The results from previous section show that worse economic conditions, that is, higher unemployment rates, push individuals towards self-employment. The mechanism driving unemployed and employed individuals to self-employment is anyhow different. Less able unemployed with lower chances to find a job choose self-employment to escape from unemployment. For employed people the pattern reverses: adverse economic conditions, as precariousness of their jobs or poor career perspectives, incentive more able individuals to start up a business.

However the longer an individual has been unemployed, the lower are his chances of switching to self-employment due to loss of human capital or lack of information about opportunities of starting an enterprise.

The model compounds the analysis of the effect of general economic conditions and individual unemployment experience. The results encompass previous findings. Evans and Leighton (1989) or Alba-Ramirez and Freeman (1994) find that previous experience of unemployment increases the probability of workers entering self-employment. Acs *et al.* (1994) find that unemployment rates increase the probability of becoming self-employed. Our model confirms these results and extends them in a natural way to unemployed individuals, considering also exit rates from self-employment for both groups.

Therefore, data in this analysis supports the theory of a deterioration of labour market conditions as fundamental to explain the growth in self-employment rates in the last two decades. Bad economic conditions have a positive effect on self-employment rates (through reduction of its opportunity cost). This encouraging effect dominates the negative effect implied by the reduction of the expected returns from self-employment.

Some other interesting results refer to family background variables effect in self-employment transitions. Although having a self-employed mother increases the probability of becoming self-employed, the effect of the father status is stronger and better defined; both prevent from becoming unemployed.

Data would also be consistent with the presence of liquidity constraints (through age and duration effects) although the simplicity of the model makes impossible to test this hypothesis. An interesting extension would be to allow for an explicit test of this hypothesis, besides some wage/earnings' effect, using a more structural model. This is far beyond the scope of the present paper and therefore left for future research.

Figure 4: Transition from Employment and Unemployment to Self-Employment (by individual type)

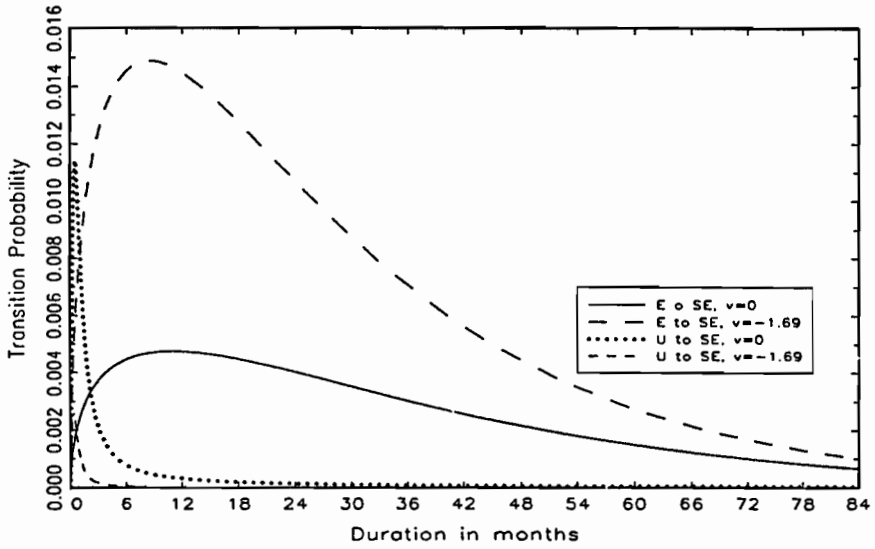
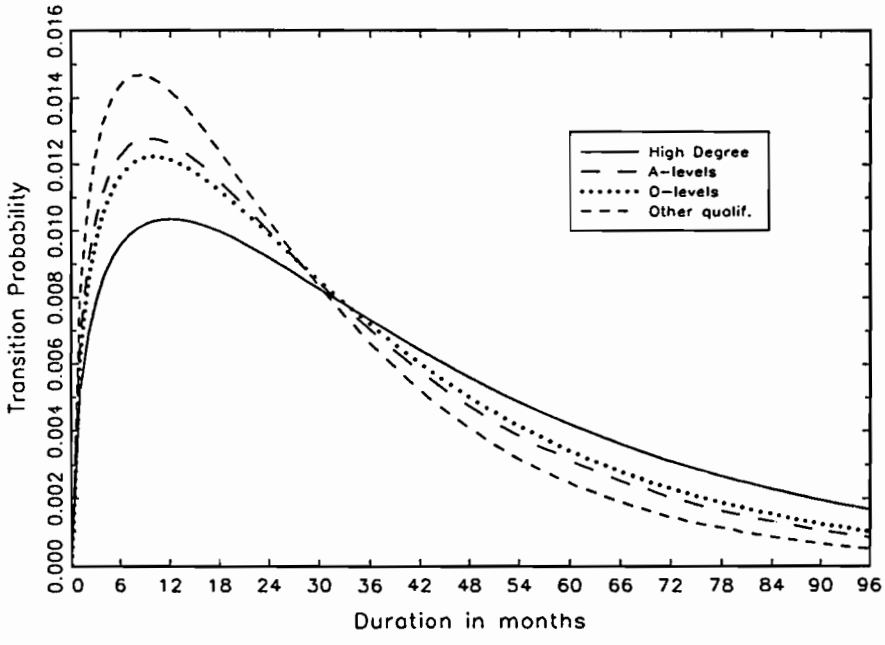


Figure 5: Transition from Employment to Self-Employment (by education).



Note: estimates used to compute transition probabilities from Table 4.

Figure 6: Transition from Unemployment to Self-Employment (by education).

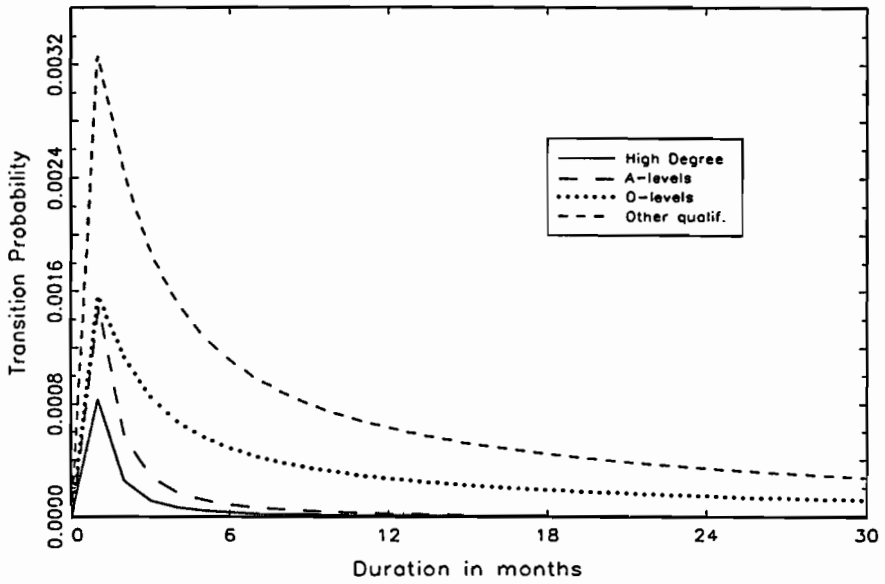
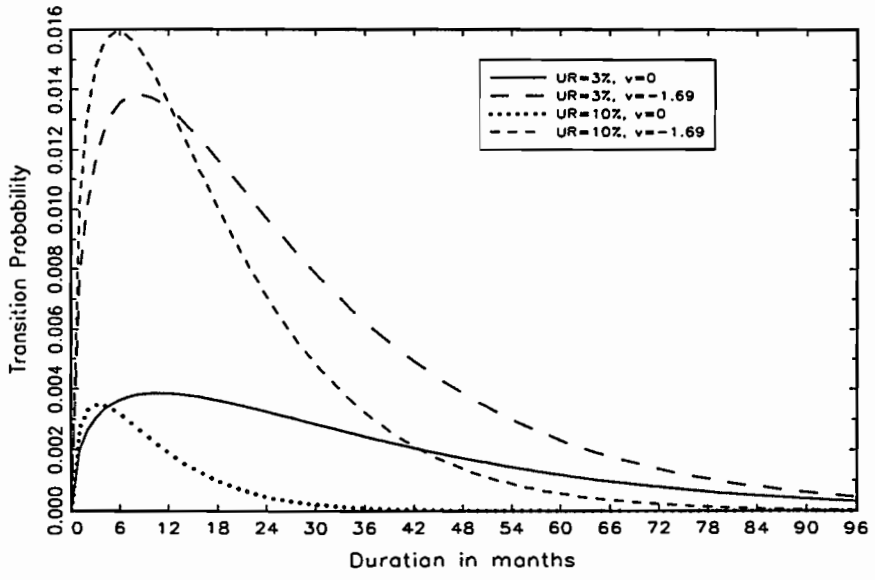
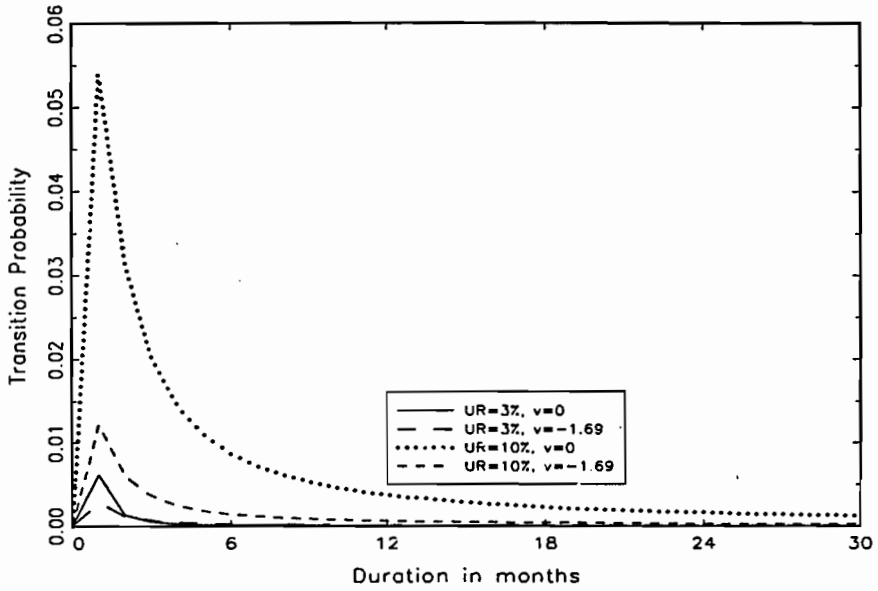


Figure 7: Transition from Employment to Self-Employment (increase in the National Unemployment Rate).



Note: estimates used to compute transition probabilities from Table 4.

Figure 8: Transition from Unemployment to Self-Employment (increase in the National Unemployment Rate)



Note: estimates used to compute transition probabilities from Table 4.

6. Appendix A: Variable description

Table A.1: Sample statistics for the relevant variables
(1978 individuals, 4227 spells).

	Observations	Mean (Std.dev.)
DURATION		
Self-Employment (SE)	213	55.038
		(61.866)
Employment (E)	1199	125.818
		(120.902)
Unemployment (U)	837	9.931
		(17.136)
AGE	1978	37.951
		(10.932)
AGE beginning spell	4227	24.668
Self-Employment		(9.807)
	548	30.757
Employment		(9.095)
	2666	21.590
Unemployment		(8.095)
	1013	29.473
		(10.780)
High Degree		0.341
		(0.474)
A levels		0.147
		(0.354)
O levels	1978	0.261
		(0.439)
Other qualif.		0.042
		(0.201)
No qualif.		0.209
		(0.407)
Mother SE		0.034
	1978	(0.182)
Father SE		0.131
		(0.337)
NUR	4227	6.113
		(3.509)

Notes: duration is measured in months. Right censored observations are not considered in computing its mean and standard deviation. Age is age at the interview date (around 12/92). NUR is National Unemployment Rate at the beginning of the spell.

7. Appendix B: EM algorithm

Section 2 introduces the likelihood function to estimate. Simplifying notation in equation (8), the final likelihood function for the whole sample is

$$L(\beta, v, \pi | \mathbf{t}, X_i) = \sum_i \left[\ln \sum_{m=1}^{M} f_i(\mathbf{t} | X_i, v_m, \beta) \pi_m \right] \tag{9}$$

where $f(.)$, is the contribution to the likelihood for each individual, conditional on v_m; $\mathbf{t}$ is a vector of duration in every spell, for each individual;β is the vector of all parameters of interest; X_i is a vector of individual characteristics and π_m is the probability attached to every mass point v_m.

Taking derivatives in (9) with respect to β and rearranging terms we obtain,

$$\frac{\partial L(.)}{\partial \beta} = \sum_i \frac{\partial \ln(f_i(.|v_m)}{\partial \beta} \hat{\pi}_m \tag{10}$$

where

$$\hat{\pi}_m = \frac{f_i(.|v_m)\pi_m}{\sum_m f_i(.|v_m)\pi_m} \tag{11}$$

The EM algorithm has two stages: expectation and maximisation. Giving initial values for all parameters of interest, including v_m and π_m, in the first stage we compute the probabilities $\hat{\pi}_m$ according to (11) and in the second stage we maximise the log likelihood function $L(.)$ with respect to β and v_m, obtaining $L_1(.)$; we will then update $\hat{\pi}_m$ recomputing (11) and so forth. This procedure produces a local optimum for $L(.)$ and the estimated values for the mass point probabilities are constrained to be in the unit interval (Heckman and Singer (1982,1984)).

To guard against failure to locate a global optimum a variety of starting values is used in the implementation of the EM algorithm.

8. Appendix C: Results using experience

Table C.1: Maximum Likelihood estimates for the transition equations; without controls for unobserved heterogeneity.

	E to U	E to SE	U to E	U to SE	SE to E	SE to U
High Degree	-0.390	-0.025	0.761	0.659	0.617	-0.144
	(0.172)	(0.205)	(0.172)	(0.540)	(0.386)	(0.469)
A-Levels	-0.126	0.384	0.750	0.982	0.445	-0.799
	(0.211)	(0.229)	(0.216)	(0.645)	(0.427)	(0.656)
O-Levels	0.018	0.275	0.612	0.977	0.167	-0.136
	(0.170)	(0.210)	(0.201)	(0.583)	(0.404)	(0.492)
Other Qualification	-0.024	0.349	-0.184	-0.169	0.393	-1.208
	(0.332)	(0.342)	(0.391)	(1.447)	(0.636)	(5.719)
Mother SE	-0.380	0.406	-0.140	0.661	0.549	-0.034
	(0.294)	(0.293)	(0.338)	(0.715)	(0.402)	(1.702)
Father SE	0.043	0.529	-0.086	0.480	-0.482	0.228
	(0.171)	(0.171)	(0.184)	(0.365)	(0.331)	(0.414)
Age	0.278	2.939	-0.266	1.926	-0.728	-0.226
	(0.390)	(0.723)	(0.385)	(1.411)	(1.037)	(1.589)
Age Squared	-0.090	-0.569	-0.010	-0.241	0.003	-0.012
	(0.060)	(0.135)	(0.053)	(0.185)	(0.162)	(0.260)
NUR	0.202	0.073	-0.020	0.165	0.051	0.173
	(0.021)	(0.027)	(0.025)	(0.077)	(0.040)	(0.070)
ln(duration/12)	-0.133	0.343	-0.639	-0.370	0.140	-0.350
	(0.085)	(0.124)	(0.116)	(0.268)	(0.206)	(0.163)
$(\ln(\text{duration}/12))^2$	0.052	-0.032	-0.244	-0.184	-0.249	-0.073
	(0.029)	(0.041)	(0.053)	(0.139)	(0.082)	(0.117)
Prev.SE.Exp.	0.153	2.409	-0.458	0.568	0.504	0.727
	(0.345)	(0.270)	(0.521)	(0.911)	(0.494)	(0.577)
Prev.Empl.Exp.	0.168	-0.149	-0.135	-0.106	-0.094	0.285
	(0.018)	(0.177)	(0.069)	(0.140)	(0.455)	(0.183)
Prev.Unemp.Exp.	0.042	0.031	0.023	0.0004	0.060	0.008
	(0.006)	(0.008)	(0.009)	(0.020)	(0.021)	(0.027)
Intercept	-4.963	-9.084	0.344	-7.609	-1.908	-3.975
	(0.570)	(0.917)	(0.667)	(2.517)	(1.593)	(2.267)
Log-likelihood				-7577		
Observations				4227		

Note: Standard errors (computed from the inverse of the information matrix) in brackets. SE denotes self-employment, E employment and U unemployment. Age is measured at the beginning of the spell. NUR is National Unemployment Rate.

Table C.2: Maximum Likelihood estimates for the transition equations controlling for unobserved heterogeneity (NPMLE).

	E to U	E to SE	U to E	U to SE	SE to E	SE to U
High Degree	-0.356	-0.042	0.762	0.655	0.620	-0.079
	(0.121)	(0.179)	(0.111)	(0.336)	(0.282)	(0.411)
A-Levels	-0.099	0.371	0.754	0.988	0.461	-0.750
	(0.151)	(0.198)	(0.135)	(0.476)	(0.299)	(0.572)
O-Levels	0.039	0.211	0.615	0.970	0.140	0.079
	(0.116)	(0.175)	(0.113)	(0.319)	(0.291)	(0.437)
Other qualific.	-0.001	0.327	-0.180	-0.189	0.360	-1.040
	(0.226)	(0.326)	(0.208)	(0.853)	(0.438)	(1.631)
Mother SE	-0.307	0.388	-0.142	0.685	0.596	-0.057
	(0.207)	(0.304)	(0.257)	(0.486)	(0.347)	(0.760)
Father SE	0.042	0.550	-0.086	0.438	-0.478	0.303
	(0.123)	(0.163)	(0.132)	(0.328)	(0.283)	(0.378)
Age	0.071	2.798	-0.236	1.953	-0.949	0.260
	(0.272)	(0.607)	(0.227)	(0.892)	(0.856)	(1.544)
Age Squared	-0.056	-0.551	-0.014	-0.245	0.036	-0.078
	(0.043)	(0.108)	(0.033)	(0.122)	(0.132)	(0.256)
NUR	0.199	0.083	-0.020	0.163	0.055	0.151
	(0.014)	(0.023)	(0.015)	(0.063)	(0.032)	(0.062)
ln(duration/12)	-0.123	0.318	-0.639	-0.359	0.138	-0.328
	(0.043)	(0.126)	(0.059)	(0.173)	(0.177)	(0.165)
$(\ln(\text{duration}/12))^2$	0.047	-0.013	-0.243	-0.181	-0.244	-0.082
	(0.012)	(0.040)	(0.028)	(0.091)	(0.080)	(0.096)
Prev.SE Exp.	0.341	2.571	-0.454	0.556	0.534	0.483
	(0.356)	(0.299)	(0.612)	(0.660)	(0.435)	(0.660)
Prev.Empl.Exp.	0.204	-0.052	-0.135	-0.121	0.030	0.152
	(0.018)	(0.087)	(0.033)	(0.101)	(0.325)	(0.147)
Prev.Unemp.Exp.	0.037	0.033	0.023	-0.001	0.061	0.002
	(0.005)	(0.008)	(0.006)	(0.017)	(0.017)	(0.026)
Unobs. heter.	1.000	-1.228	0.021	0.841	-0.489	2.358
	----	(0.440)	(0.231)	(1.072)	(0.380)	(1.222)
Intercept	-4.416	-9.787	0.300	-7.509	-2.041	-3.782
	(0.374)	(0.820)	(0.375)	(1.532)	(1.243)	(2.226)
Log-likelihood	-7315					
Observations	4227					

Notes: Standard errors (computed from the inverse of the information matrix) in brackets. SE denotes self-employment, E employment and U unemployment. Age is measured at the beginning of the spell. NUR is National Unemployment Rate.

Two support Points: v_1=0 with probability p_1=0.65 and v_2=-1.296(0.323) with probability p_2=0.35. The heterogeneity coefficient for the transition from E to U is normalise to one for identification.

REFERENCES

Alba-Ramirez, A. (1994), 'Self-Employment in the Midst of Unemployment: the case of Spain and the United States', *Applied Economics*, vol. 26, pp. 189-204.z

Blanchflower, D. and Oswald, A. (1990), 'What Makes a Young Entrepreneur?', *NBER Working Paper*, No. 3252.

Blanchflower, D. and Oswald, A. (1991), 'Self-Employment and Mrs. Thatcher's Enterprise Culture', CEP *Working Paper*, No. 30.

Blundell, R., Saleheen, J. and Walker, I. (1995), 'Job Mobility, Self Employment and Earnings Differentials', *mimeo*.

Evans, D. and Jovanovic, B. (1989), 'An Estimated Model of Entrepreneurial Choice under Liquidity Constraints', *Journal of Political Economy*, vol. 97, pp 808-27.

Evans, D. and Leighton, L. (1989), 'Some Empirical Aspects of Entrepreneurship', *The American Economic Review*, vol. 79.3, pp 519-535.

Heckman, J.J., and Singer, B. (1984), 'A Method for Minimizing the Impact Of Distributional Assumptions in Econometric Models for Duration Data', *Econometrica*, vol. 58, pp 1411-1441.

Holtz-Eakin, D., Joulfaian, D. and Rosen, H. (1994a), 'Entrepreneurial decisions and liquidity constraints', *RAND Journal of Economics*, vol. 25.2, pp334-346.

Holtz-Eakin, D., Joulfaian, D. and Rosen, H. (1994b), 'Sticking it Out: Entrepreneurial Survival and Liquidity Constraints', *Journal of Political Economy*, vol.102.

Lancaster, T. (1990), 'The Econometric Analysis of Transition Data', *Cambridge University Press*.

Magnac, T. and Robin, J.M. (1990), 'Econometric Analysis of the Transitions between Self-Employment and Wage-Work', *Delta Working Paper*, No. 9009.

Rees, H. and Shah, A. (1986), 'The Determinants of Self Employment in the UK', *Journal of Applied Econometrics*, vol.1, pp 95-108.

Taylor, M. P. (1996), 'Earnings, Independence or Unemployment: Why Become Self-Employed?', *Oxford Bulletin of Economics and Statistics*, vol. 58.2, pp 253-265.