

## ARTICLE

Katy Long  
Office for National Statistics

# Unemployment durations: evidence from the British Household Panel Survey

## SUMMARY

This article uses data from the British Household Panel Survey (BHPS), over the period 1991 to 2006, to examine the factors affecting the length of unemployment spells. The analysis is carried out with particular interest in the effect of regional labour market conditions on an individual's conditional probability of leaving unemployment. A discrete time proportional hazards model is estimated, controlling for a range of demographic, educational, occupational and regional characteristics.

## Introduction

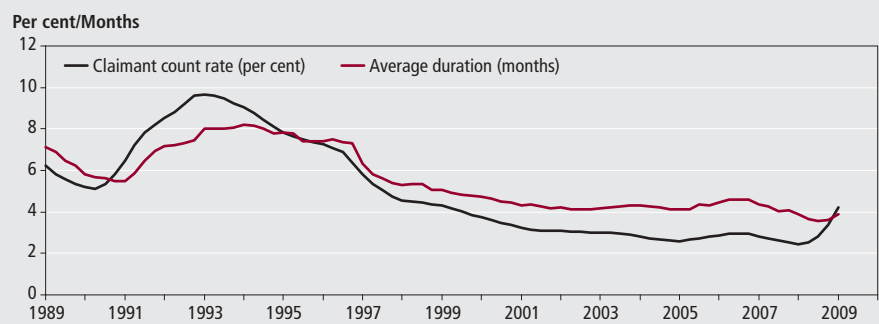
In the current economic climate, there is a heightened interest in the effect of the recession on the labour market. The downturn in the economy has led to an increase in unemployment. Of particular concern is the increase in long-term unemployment which is associated not only with a loss of current income but also inflicts longer term effects through increased future incidence of unemployment, lower job tenure and reduced earnings (Arulampalam (2001) and Gregory and Jukes (2001)). These 'scarring effects' can occur where the skills of unemployed individuals depreciate whilst they are unemployed or where potential employers view long spells of unemployment as a signal of a low quality worker.

Changes in the stock of unemployed workers are determined by the relative flows of individuals into and out of unemployment. Thus an increase in the

number unemployed workers can be attributed to either an increase in the rate that individuals become unemployed or a decrease in the rate that they leave unemployment. A lower outflow than inflow rate will mean that on average people remain in unemployment for longer.

Figure 1 illustrates the relationship between the stock and steady state average duration of the claimant count in the UK (which measures the number of people claiming unemployment related benefit). The steady state average duration of claimants in the UK was calculated using the method of Layard (2005). Over the past 20 years, the aggregate claimant count has shown substantial variability. The average duration of claims has followed a similar path, but lagged peaks and troughs of the aggregate count. This suggests that changes in the average duration of claimants contribute to explaining the claimant rate.

Figure 1  
UK claimant count rate and average duration<sup>1</sup>



Note:

Source: Office for National Statistics

<sup>1</sup> Claimant count rate = claimant count / (claimant count + workforce jobs). The average duration is calculated using ONS source data

Although the unemployment rate measured using International Labour Organisation (ILO) is different to the claimant count, the average duration of unemployment will influence the unemployment level and rate in a similar way.

It is therefore useful for policy design purposes to analyse the factors influencing the length of unemployment spells. For example, if individuals with specific characteristics are associated with longer unemployment durations, policy could be targeted at assisting these groups with job search. However, if wider macroeconomic conditions are found to be more relevant in explaining re-employment probabilities, the targeting policy described above, whilst changing the position of individuals in the conceptual queue of jobseekers, will have limited impact on the aggregate unemployment rate (Imbens and Lynch (2006)). The analysis used in this article controls for the effect of macroeconomic conditions using the regional claimant count rate (which is different but related to the unemployment rate) and furthermore investigates changes in the magnitude of any such effects over a spell of unemployment (as measured by the BHPS).

## Theoretical background

Job search theory is the primary theoretical framework used by economists for analysing the determinants of unemployment durations. At the most basic level job search theory holds that, when an individual becomes unemployed, their probability of re-employment is equal to their probability of receiving a job offer multiplied by the probability of the individual accepting it. Factors likely to determine the offer of a job include an individual's education and skills, search intensity and the demand conditions in the appropriate labour market they are searching in. The probability that an individual accepts a job offer is determined by their reservation wage. This is the minimum wage at which an individual is willing to supply their labour. Factors likely to affect this include the expected wage distribution in their segment of the labour market, family composition (i.e. whether they have children or their spouse works), unemployment income, for example job seekers allowance and job search costs.

The effect of the unemployment rate on the probability of re-employment is theoretically ambiguous. Whilst an increase in aggregate unemployment is likely to reduce an individual's probability

of receiving a job offer, it is also likely to reduce their reservation wage by lowering their expectation of the wage distribution. Determining the net effect is, therefore, a matter of empirical investigation.

An important feature of interest when analysing unemployment experiences is the nature of duration dependence. That is, how the probability of exiting unemployment changes over the length of an unemployment spell. Negative duration dependence occurs when the probability of exiting unemployment falls as the duration of the spell increases. The model developed for this article was primarily constructed to investigate this relationship.

The ranking model of Blanchard and Diamond (1994) predicts negative dependence duration. In this model, when firms receive multiple applications for a vacancy, they use the workers' unemployment spell length to rank individuals, assuming that it is a good proxy for unobserved differences in skills between individuals. In a depressed labour market, when the number of applications per vacancy is high, there is a greater likelihood of there being someone with a shorter spell length applying for the same job. Conversely, when the labour market is relatively tight, the likelihood of there being an individual applying for the same job with a shorter duration is lower. This generates the empirical prediction that individuals with long unemployment spells are more affected by an increase in the unemployment rate than those with relatively shorter durations. Put another way, those with lengthy unemployment durations are damaged more by an increase in the unemployment rate than those with relatively short unemployment spells.

## Description of the data

The primary data source used in this analysis is the British Household Panel Survey (BHPS). This, nationally representative sample, is a rich panel dataset of approximately 10,000 households, comprising 17 waves of information at both individual and household level. For more information on this source see Taylor et al (2009). The sample used in this study spans the period January 1991 to January 2006 and contains information on 3,959 spells of unemployment for 2,368 individuals. Additionally, for the relevant period, RPI and regional monthly claimant count data are used.

Only males are considered in this analysis due to the difficulties in constructing accurate work life histories for women. The

sample is further restricted to exclude males under the age of 18 and males who turn 60 during the sample period. This is because attachment to the labour market is typically weak for those aged below 18 and the need to abstract from the retirement decision, which may play a role for individuals turning 60. To illustrate the first of these points, consider the case of a 16 year old male that has recently left school. It is not clear to the researcher whether he has started his job search activity, is having a gap year or is living at home with his parents.

Alternative data sources on unemployment spells such as the Joint Unemployment and Vacancy Operating System used to produce the claimant count (JUVOS) and the Labour Force Survey (LFS) are also available. Both of these are preferable to the BHPS in terms of sample size. Additionally JUVOS, in being an administrative data source, provides more reliable data on income and industry variables. However, these alternative sources have their own drawbacks. The LFS lacks information on individuals' income whilst unemployed, and JUVOS lacks detailed information on individual characteristics such as educational attainment and housing tenure. Since survival analysis models are particularly sensitive to unobserved differences in individuals' characteristics, the BHPS is the preferred data source for this study.

A flow sample selection was used, so that an individual enters the sample when they become unemployed and remain in the sample until they exit to employment. This means that unemployment spells resulting in exit to inactivity, retirement or self-employment are not considered in this study. A weakness of the BHPS is that it is not administrative in nature. This means that the information gathered is based largely on self reporting. Since the definition of unemployment used in this study is based on the BHPS and a different subset of the male population to the LFS, it is not consistent with that of the International Labour Organisation (ILO). In addition, the regional rather than national claimant count rate is used to control for cyclical labour market effects. This is because it may not be possible to distinguish between the effect of the national claimant count rate and dependence duration.

A preliminary examination of the data reveals that of those individuals who entered unemployment over the period 1991 to 1992, 23 per cent remained unemployed after 12 months. In contrast,

the corresponding figure for those who entered unemployment between 2004 and 2005 is just 5 per cent. This is suggestive of a marked disparity in the unemployment experience of those who become unemployed in a recession and those who become unemployed during an economic upturn

## Modelling the unemployment durations

Central to modelling the determinants of the length of unemployment spells is the estimation of the conditional probability of an individual exiting unemployment. This is the probability that an individual will exit unemployment in the next period, conditional on being unemployed up until that period. This is commonly referred to as the hazard function. This is estimated using techniques from a branch of statistical methods known as survival analysis. Readers are referred to the technical note for a detailed exposition of this modelling technique.

Whilst other forms of the hazard function exist, this study adopts the discrete time proportional hazards model of Cox (1972) because of the flexibility and convenience of interpretation that it offers. The analogous continuous time hazard function for the  $i^{\text{th}}$  individual is parameterised as:

$$\theta_{it} = \lambda_i \exp[X_{it}'\beta] \quad (1)$$

where  $\theta_{it}$  is the conditional hazard,  $\lambda_i$  is the baseline hazard,  $X_{it}$  is a vector of individual specific explanatory variables (some of which may vary over time e.g. regional unemployment rate, marital status) and  $\beta$  is a vector of parameters to be estimated.

The baseline hazard is a function of the elapsed spell duration alone and can be thought of as the hazard function in the case where there are no covariates. Its parameter characterises the pattern of dependence duration i.e. how the conditional probability of exit changes with the elapsed length of the spell. A more detailed exposition of the model can be found in the technical note.

A proportionality assumption implies that the hazard function is multiplicatively separable in elapsed duration and the group of covariates. This essentially amounts to the assumption that the conditional hazard function is proportional to the baseline hazard by a scaling factor,  $\exp[X_{it}'\beta]$ . This representation is convenient since the estimated coefficients give the proportional change in the hazard associated with an absolute change in the corresponding

covariate. This proportional change is independent of time.

A further benefit of the proportional hazards model is that its consistent estimation does not require the parametric specification of a baseline hazard. Inconsistent parameter estimates may arise from a fully parametric model if any part of it is misspecified. Instead this study allows for a more flexible specification of the baseline hazard. Specifically, it takes a piecewise constant form so that the baseline hazard is assumed constant during each interval, but is allowed to vary between them. This allows for observation of the shape of the baseline hazard without constraining it to a specific functional form.

Unobserved differences between individuals pose a problem in survival modelling. If ignored, they can bias the results towards negative duration dependence. This is due to a composition effect. As an example consider a group of individuals who have a characteristic which lowers their employment probability. On average this group will have a lower probability of exit, relative to the rest the sample. As time goes on the sample will increasingly consist of individuals from this group and so the average probability of re-employment for the whole sample will diminish with time. If the characteristic is unobservable it will erroneously appear to the researcher that the probability of exit for a given individual is declining in the length of their unemployment spell. Additionally, Lancaster & Nickell (1980) showed that the presence of unobserved differences between individuals will artificially reduce the proportional effect of a change in a covariate. Moreover, the proportional effect will no longer be constant or independent of time. The model used in this study assumes unobserved differences can be represented using a normally distributed error term, the effects of which are

integrated out. Readers are directed to the technical note for a more thorough explanation of this technique.

## Results

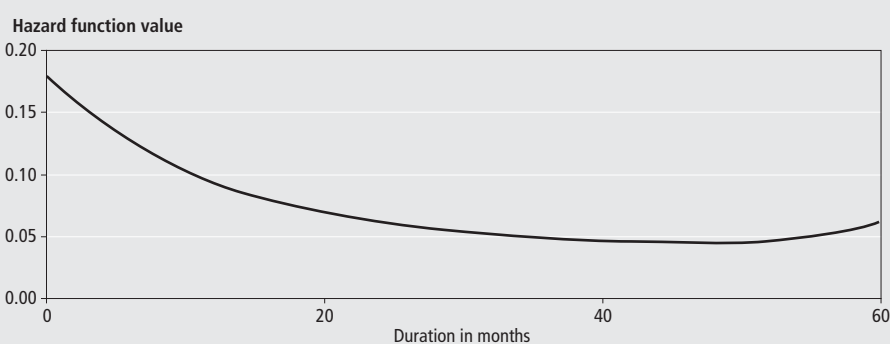
The results of the estimation are summarised in **Table 1** and the estimated baseline hazard function is illustrated in **Figure 2**. The hazard function shows probability of exit without taking into account the effect of any variables other than elapsed duration and for which no parametric form has been specified. Figure 2 shows a hazard function which is decreasing in elapsed spell duration until approximately the 50<sup>th</sup> month, after which the conditional probability of exiting unemployment is gently increasing in duration. The estimated hazard function is consistent with the hypothesis of duration dependence whereby the probably of exiting unemployment falls as the unemployment spell gets longer. The increasing probability of exit after the 50<sup>th</sup> month may be due to discouraged workers leaving the labour market, moving from unemployment to inactivity – the ‘discouraged worker’ effect.

Table 1 presents the results of the discrete-time analogue Cox’s proportional hazard model described. The coefficients associated with the explanatory variables are listed along with their degrees of significance and standard errors. It is mainly the statistically significant results at the one and five per cent level that are discussed in the section that follows.

### Personal characteristics

Age is found to have a negative effect on the conditional probability of leaving unemployment. When all other things are held equal, married men are 33 per cent more likely to exit unemployment than unmarried men. This could be rationalised in the context of job search theory as increasing an individual’s probability

**Figure 2**  
**Lowest smoothed estimate of interval hazard function**



Source: Author's estimates

**Table 1**  
**Results of discrete-time analogue Cox's proportional hazard model**

	Coefficient <sup>1</sup>	Standard Error
Elapsed Duration (months) < 7	0.353**	0.329
6 < Elapsed duration (months) < 13	0.522	0.343
12 < Elapsed duration (months) < 19	0.634	0.377
18 < Elapsed duration (months) < 25	0.981	0.412
24 < Elapsed duration (months) < 31	1.634	0.492
30 < Elapsed duration (months) < 37	0.840	0.527
36 < Elapsed duration (months) < 43	0.982	0.531
42 < Elapsed duration (months) < 49	1.022	0.542
48 < Elapsed duration (months) < 55	1.072	0.570
54 < Elapsed duration (months) < 60	1.330	0.599
<sup>2</sup> Regional claimant count rate × (13 > elapsed duration > 6 months)	0.929**	0.019
Regional claimant count rate × (19 > elapsed duration > 12 months)	0.905**	0.029
Regional claimant count rate × (25 > elapsed duration > 18 months)	0.878**	0.039
Regional claimant count rate × (31 > elapsed duration > 24 months)	0.805**	0.058
Regional claimant count rate × (elapsed duration > 30 months)	0.885*	0.056
Regional claimant count rate	0.993	0.023
Age (in months at start of spell)	0.998**	0.000
Married	1.327**	0.075
Has dependent children	0.877	0.069
Member of ethnic minority	1.028	(0.101)
Post introduction of New Deal	0.624**	0.131
Post introduction of National Minimum Wage	0.642**	0.132
No. of previous unemployment spells	1.025	0.017
Replacement ratio	0.670**	0.127
Highest Educational Qualification:		
A Level	1.215**	0.075
Further education	1.037	0.150
Degree (or higher)	1.266*	0.102
Housing tenure:		
Home owned outright	1.117	0.257
Home owned with mortgage	1.273**	0.001
Council rented	0.682**	0.000
Occupational Group of Previous Job:		
Managers & administrators	1.275*	0.115
Professional occupations	1.268	0.152
Associate professional & technical occupations	1.152	0.133
Clerical & secretarial occupations	1.215*	0.095
Craft & related occupations	1.144	0.076
Personal & protective service occupations	1.280*	0.098
Sales occupations	1.156	0.104
Plant and machine operatives	1.154	0.077
Father's Occupational Group:		
Managers & administrators	0.937	0.659
Professional occupations	0.955	0.827
Associate professional & technical occupations	1.245	0.358
Clerical & secretarial occupations	1.319	0.245
Craft & related occupations	0.790*	0.029
Personal & protective service occupations	0.813	0.370
Sales occupations	1.240	0.441
Plant and machine operatives	0.900	0.399
Region:		
Inner & Outer London	0.956	0.137
South West	0.817	0.131
East Anglia	0.723*	0.161
East Midlands	0.852	0.119
West Midlands	0.889	0.132
Tyne & Wear	0.764	0.182
North West	0.673**	0.120
Yorkshire & Humberside	0.908	0.131
Region of the North	0.810	0.182
Wales	0.547**	0.120
Scotland	0.655**	0.113
Northern Ireland	0.337**	0.170

**Notes:**

- Exponentiated coefficients are reported with standard errors in parentheses.
  - Elasticities are calculated using the logged values of these coefficients.
- \* Denotes significance at the 5% level.  
\*\* Denotes significance at the 1% level.

Source: Author's estimates

of receiving an offer, insofar as positive attributes such as reliability are associated with being married. All other things equal, those with A Levels as their highest educational qualification have a hazard that is 22 per cent higher than those who left school aged 16. The corresponding figure for those with a degree is 27 per cent. This is unsurprising in light of the existence of positive returns to education and training. Ethnicity and the number of dependant children are not found to have a statistically significant effect. This latter result may be explained by the fact that unemployment benefits are generally adjusted for family size.

**Labour market policy**

When all other factors are kept equal, having an unemployment spell after the introduction of the national minimum wage is associated with a hazard that is 36 per cent lower than that of individuals who experience an unemployment spell before its introduction. Whilst this is consistent with the economic theory of a perfectly competitive labour market model, caution should be applied when interpreting this result. The variable used to control for the national minimum wage is relatively crude in its design and is merely comparing the labour market between two, approximately, eight year time spans. The New Deal was introduced just a year earlier and the variable which controls for it is constructed in the same way. These variables are therefore likely to be capturing wider structural changes in the UK labour market. To identify the true effect of the national minimum wage on re-employment probabilities a difference-in-difference approach such as that in Stewart (2004) would be more appropriate. A further point to consider is that the approach used in the current study only considers transitions from unemployment to employment. Where the effect of the minimum wage is concerned, analysing transitions between states such as low pay employment, high pay employment, unemployment and inactivity would be a more instructive exercise. For example, it may well be that the introduction of the national minimum wage has resulted in individuals who were previously discouraged re-entering the labour market. Analysis of this type is carried out using competing risks models.

An incremental previous unemployment spell is found to increase the hazard of leaving unemployment by three per cent, all other things equal. This previous unemployment spell variable is likely to be

capturing attachment to the labour market. For example, individuals who have just left education, and therefore have no previous unemployment spells, are likely to have lower attachment than those individuals who have 20 years work experience and, therefore, have likely experienced a spell of unemployment before.

A person's income whilst unemployed as a proportion of their income whilst in employment is defined as their replacement ratio. The results show that a unit increase in the replacement ratio is associated with a hazard rate that is 33 per cent lower than its previous value. Evaluated at its mean, a 10 per cent increase in the replacement ratio is associated with a 3.5 per cent decrease in the probability of exiting unemployment. To help conceptualise this, when all variables are at their means a £3 increase in income whilst unemployed is associated with a decrease in the hazard of 0.35 units. Therefore, whilst there is a negative impact associated with unemployment income (which includes benefits, investments and pensions income), the probability of exit is relatively inelastic to it. This is likely to be reflecting that fact that the replacement ratio does not take account of sources of wealth such as savings which an individual may live off whilst unemployed.

### Region and access

When all other factors are held equal, those residing in the East Anglia, the North West, Wales, Scotland and Northern Ireland are all found to have a significantly lower hazard relative to those living in the South East. The greatest of these effects is for those people living in Northern Ireland who have a hazard rate that is just 34 per cent of that of identical individuals living in the South East. For Wales and Scotland the corresponding figures are 55 and 66 per cent, respectively. Living in local authority accommodation is associated with a hazard that is 32 per cent lower than that of individuals living in privately rented accommodation, all other things equal. Conversely owning a home with a mortgage is associated with a hazard 27 per cent higher than that of private renters, all else equal.

### Occupation

Individuals whose previous job was in "Managers and Administrators", "Personal and Protective Services" and "Clerical and Secretarial" occupations all have a hazard that is higher than the lowest skilled occupational group and

statistically significant at the 5 per cent level. Individuals who work in "Personal and Protective Services" and "Managers and Administrators" have a hazard rate 28 per cent higher than individuals in the lowest skilled group with the same characteristics.

An individual's father's occupation is not found to have a statistically significant effect, with the exception of those individuals whose father's were in the "Craft and Related" occupations. These individuals, all other things equal, have a hazard that is 21 per cent lower than that of individual's whose fathers were in the lowest skilled occupational group.

### Labour market conditions

Looking at the labour market variables, the regional claimant count rate is found to have a negative effect on the conditional probability of exit that is increasing in magnitude with elapsed duration. The response of the conditional probability of exit to a proportionate change in the regional claimant count rate is given by the elasticity of the hazard to the regional claimant count rate. When the regional claimant count rate is at its mean value, the elasticity of the hazard is constructed as follows: the logged value of the regional claimant count coefficient is added to the logged value of the coefficient variable for the relevant months of duration. This is then multiplied by the mean value of the regional claimant count to get the corresponding elasticity.

The results from this analysis show that when the regional claimant count is at its mean level, all other things equal, a 10 per cent increase in the regional claimant count rate is associated with a 2.1 per cent decrease in the hazard for individuals with durations between 7 and 12 months. In contrast for those with durations between 13 and 18 months, the corresponding reduction is 2.7 percent. The corresponding figures for those with elapsed durations between 19 and 24 months and 25 and 30 months are 3.5 per cent and 5.7 per cent respectively.

The implication of this result is that, all other things equal, individuals with long spells of unemployment are damaged more by increases in the claimant count rate than individuals with shorter spells. This is consistent with the ranking model in which potential employers rank applicants by their elapsed duration. Hence, for a given individual, an increase in the claimant count rate (as would an increase in the unemployment rate) increases

the likelihood that someone else applies for the same job. As stated previously, a rival applicant is more likely to have a shorter unemployment duration when the reference individual is in a long spell of unemployment than when they are in a relatively short one. In this way the reference individual is damaged more by an increase in the claimant rate when they are in a long term spell of unemployment. These results echo the empirical findings of Dynarski and Sheffrin (1990) for the US labour market.

### Conclusions

The results of this study indicate that re-employment prospects are positively influenced by: being married, having A Levels or a degree, and owning a home with a mortgage. This is shown by the higher conditional probabilities of re-employment associated with these factors.

In contrast, living in Scotland, Wales, Northern Ireland, the North West and East Anglia is associated with a lower hazard of exit from unemployment, relative to those in the South East. Those in local authority accommodation also experience a lower conditional probability of exit.

The effect of a labour market downturn is not found to affect all individuals symmetrically. Specifically, the negative effect of an increase in the unemployment rate is amplified as the length of an individual's unemployment duration increases. This suggests a role for government policy in preventing individuals from losing contact with the labour market. This is because it is these long term unemployed individuals whose re-employment prospects suffer most from a labour market downturn.

### ACKNOWLEDGEMENTS

The author would like Richard Jones for his help and advice.

### CONTACT

✉ [elmr@ons.gsi.gov.uk](mailto:elmr@ons.gsi.gov.uk)

### REFERENCES

- Arulampalam, W. (2001), "Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages", *Economic Journal*, 111 (475), pp.F585-F606.
- Blanchard, O. J. and Diamond, P. (1994), "Ranking, Unemployment Duration and Wages", *Review of Economic Studies*, vol. 61, no. 3, pp. 417-434.

- Cox, D., R. (1972), "Regression Models and Life-Tables", *Journal of the Royal Statistical Society, Series B*, vol. 34, no. 2, pp. 187-220.
- Dynarski, M. and Sheffrin, S. M. (1990), "The Behaviour of Unemployment Durations over the Cycle", *The Review of Economics and Statistics*, vol. 72, pp. 350-356.
- Gregory, M. and R. Jukes (2001), "Unemployment and Subsequent Earnings: Estimating Scarring among British Men 1984-1994", *Economic Journal*, 111 (475), pp. F607-F625.
- Imbens, G. W. and Lynch, L. M. (2006), "Re-employment Probabilities over the Business Cycle", *Portuguese Economic Journal*, vol. 5, pp. 111-134.
- Jenkins, S., P. (1998), "Discrete Time Proportional Hazards Regression", *Stata Technical Bulletin*, vol. 7, no. 39, pp. 22-32.
- Lancaster, T. (1979), "Econometric Methods for the Duration of Unemployment", *Econometrica*, vol. 47, no. 4, pp. 940-956.
- Layard, R., Nickell, S. and Jackman, R. (2005), *Unemployment: Macroeconomic Performance and the Labour Market*, Oxford University Press, New York.
- Lancaster, T. and Nickell, S. (1980), "The Analysis of Re-employment Probabilities for the unemployed", *Journal of the Royal Statistical Society, Series A*, vol. 143, part. 2, pp. 141-165.
- Stewart, M. B. (2004). "The Impact of the Introduction of the U.K Minimum Wage on the Employment Probabilities of Low Wage Workers", *Journal of the European Economic Association*, volume 2, no. 1, pp.67-97.
- Taylor, Marcia Freed (ed). with John Brice, Nick Buck and Elaine Prentice-Lane (2009) *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*. Colchester: University of Essex.

## TECHNICAL NOTE

**Survival analysis**

The statistical technique of survival analysis is used to model individuals' conditional probability of exit to employment. The approach was originally developed in the biomedical literature and was used to measure a patient's probability of survival, conditional on their treatment type and other individual specific factors. It has since been adapted to many other contexts in which the probability of transitions between states is analysed.

The model used in this study is the discrete time analogue of the continuous time proportional hazards model. The discrete time version is adopted owing to the inability to observe the exact time of exit for the individuals in the sample. Instead unemployment durations are measured in terms of the month the spell began and the month in which the spell ended.

The following specification of the continuous time model is based on that of Jenkins (1998). In the basic model the continuous time hazard rate, at time  $t > 0$  for the  $i^{\text{th}}$  individual is parameterised as:

$$\theta_{it} = \lambda_t \exp[X_{it}'\beta]$$

where  $i = 1, \dots, N$  indexes individuals who enter the state of unemployment at time  $t=0$ ,  $\lambda_t$  denotes the baseline hazard at time  $t$ ,  $\beta$  is a vector of parameters to be estimated and  $X_{it}$  is a vector of explanatory variables for the  $i^{\text{th}}$  individual, some of which are time variant. The corresponding survivor function is given by:

$$S(t | X_{it}) = \exp\{-\exp[X_{it}'\beta + \rho_t]\}$$

where  $\rho_t = \ln \int_0^t \lambda(u) du$  (which is the integrated baseline hazard at time  $t$ ).

**Unobserved heterogeneity**

As mentioned in the main body of this article, unobserved differences between individuals, if left unaccounted for, can bias results. Lancaster (1979) provides a full exposition of these effects. In order to guard against this, the hazard rate in the continuous time context is altered as follows. Unobserved heterogeneity is assumed to be analogous to omitted variables and is represented by a multiplicative random error term,  $v$ . This is assumed to take on only positive values, have finite variance  $\sigma^2$  and, for identification purposes, have a mean normalised to unity. We additionally require  $v$  to be independent of both  $x_i$  and  $t$ . In this case the parameterisation of the continuous time hazard for the  $i^{\text{th}}$  individual is given by:

$$\begin{aligned} \Omega_{it} &= \lambda_t \exp[X_{it}'\beta]v_i \\ &= \lambda_t \exp[X_{it}'\beta + \varepsilon_i] \end{aligned}$$

where  $\varepsilon = \ln(v)$  is a random error term with mean zero.

We are unable to write down the likelihood contribution of each individual because  $v$  is, by definition, unobservable. In order to estimate the model it is therefore necessary to specify a form for the distribution of  $v$  in terms of parameters to be estimated. This allows the survival and density functions to be written in a way that doesn't condition on  $\varepsilon$ . Theoretically, any distribution with a positive support, finite variance and unit mean is appropriate (Jenkins (1998)). In this study it is assumed that  $v$  is normally distributed. Since there is no closed form expression available for the survivor and density functions and, therefore, the likelihood contributions the effects of unobserved heterogeneity are integrated out using numerical quadrature techniques.