

IZA DP No. 754

Separability of Duration Dependence and Unobserved Heterogeneity

Hélène Turon

April 2003

Separability of Duration Dependence and Unobserved Heterogeneity

Hélène Turon

University of Bristol and IZA Bonn

Discussion Paper No. 754
April 2003

IZA

P.O. Box 7240
D-53072 Bonn
Germany

Tel.: +49-228-3894-0
Fax: +49-228-3894-210
Email: iza@iza.org

This Discussion Paper is issued within the framework of IZA's research area *Mobility and Flexibility of Labor*. Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent, nonprofit limited liability company (Gesellschaft mit beschränkter Haftung) supported by the Deutsche Post AG. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public. The current research program deals with (1) mobility and flexibility of labor, (2) internationalization of labor markets, (3) welfare state and labor market, (4) labor markets in transition countries, (5) the future of labor, (6) evaluation of labor market policies and projects and (7) general labor economics.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website (www.iza.org) or directly from the author.

ABSTRACT

Separability of Duration Dependence and Unobserved Heterogeneity*

Mixed proportional hazard models are commonly used to estimate duration dependence and unobserved heterogeneity in unemployment exit rates. Some strong assumptions are made in this framework, i.e. that the various influences on the individual unemployment exit rate are separable. The model we use in this paper allows for both the individual duration dependence pattern and the inflow composition to exhibit cyclical variations, thereby relaxing two of the three separability assumptions. The aim of this paper is to assess the validity of the third separability assumption, namely that the duration dependence pattern is the same for all individuals.

JEL Classification: J64

Keywords: unemployment outflow rate, regional and age group data, mixed proportional hazard

Corresponding author:

Hélène Turon
Department of Economics
University of Bristol
8 Woodland Road
Bristol BS81TN
United Kingdom
Email: Helene.Turon-Lacarrieu@bristol.ac.uk

* This paper is based on Chapter 4 of my PhD thesis. I would like to thank Simon Burgess and David Demery for their very helpful supervision and Gerard van den Berg for useful comments. The data used comes from NOMIS, the National On-line Manpower Information System for the UK, based in Durham University (<http://www.nomisweb.co.uk>).

1 Introduction

There is a large literature on the issue of duration dependence of the exit rate out of unemployment and on the difficulty of separating its impact from the effects of unobserved heterogeneity. There is a widely documented decrease in the rate at which individuals leave unemployment as the duration of their unemployment spell increases. This decline can reveal ‘individual duration dependence’ if all individuals see their chances of exiting unemployment decrease when they have been unemployed for a long period. Alternatively, this decline can be the result of a ‘weeding out’ process whereby individuals exhibit some heterogeneity in their ability to leave unemployment and those with high ability tend to leave unemployment relatively early, hence decreasing the average ability of leaving unemployment amongst those remaining unemployed. Mixed proportional hazard models are most commonly used to estimate both phenomena. It has been shown that results, especially results on the duration dependence phenomenon, are sensitive to any parametric assumption made in the model¹. Although techniques have been developed to get round this problem (Heckman and Singer, 1984) it has proved very difficult to carry them out with cross-section data (Meyer, 1990). The use of panel data, however, allows the estimation of both duration depen-

¹See Lancaster (1990) for a survey of the econometric issues relating to the estimation of mixed proportional hazard models.

dence and unobserved heterogeneity without parametric assumption (Aru-lampalam, Booth and Taylor, 2002, Kalwij, 2001). With aggregate data², van den Berg and van Ours (1994, 1996) developed a technique allowing the identification of both phenomena without any parametric assumption.

However, some strong assumptions are still made in the mixed proportional hazard framework, i.e. that the various influences on the individual unemployment exit rate are separable. These influences are typically the duration of the elapsed unemployment spell, the individual characteristics or the unobserved heterogeneity, and the calendar time, embodying the influence of the business cycle on unemployment exit rates. Separability between duration dependence and calendar time, and between unobserved heterogeneity and calendar time have recently been investigated by Abbring, van den Berg and van Ours (2001, 2002), Kalwij (2001) and Turon (2003). The model we use in this paper indeed allows for both the individual duration dependence pattern and the composition of the inflow (in other words the heterogeneity distribution) to exhibit cyclical variations, thereby relaxing two of the three separability assumptions inherent in the mixed proportional hazard model. The aim of this paper is to assess the validity of the third separability assumption, namely that duration dependence and heterogeneity

²Layard and Jackman (1991) also found some evidence of duration dependence with British aggregate data using a qualitative approach.

are separable, i.e. that the duration dependence pattern is the same for all individuals. We do this by estimating the model on two sets of British disaggregated data, by age and by region, and by comparing our results across the different groups. This also allows us to compare the cyclical sensitivity of the inflow composition across age and regional groups. Using a model with fixed duration dependence and inflow composition and US data, van den Berg and van Ours (1996) found that the duration dependence pattern was different between groups of male or female, black or white individuals.

The paper is organised as follows. The model used is described in section 2. In section 3, we discuss the optimal size of the disaggregated data groups and describe the data. Estimation results are reported in sections 4 and 5 for age groups and regions respectively. Section 6 concludes.

2 Model

We use a modified version of the models of van den Berg and van Ours (1994, 1996). They use a Mixed Proportional Hazard specification in a discrete time setting:

$$h(i, t, d) = \Psi_2(t) \cdot \Psi_3(d) \cdot v_i \tag{1}$$

where h is the hazard of exiting unemployment at calendar time t for individual i whose unemployment spell duration to date is d , v_i is the unobserved heterogeneity. v follows a distribution $G(v)$ with positive support. If $\Psi_3(d)$ is decreasing in d , there is negative individual duration dependence³. $\Psi_2(t)$ is the impact of the business cycle on the individual exit rate. The strength of this approach is that it does not need any parametric assumption on either the shape of the duration dependence pattern $\Psi_3(d)$ or on the shape of the heterogeneity distribution $G(v)$. This is made possible by our use of ‘aggregate’ data. Details of the method can be found in van den Berg and van Ours (1994, 1996).

As in Turon (2003), the model specification we estimate also allows for cyclical variations in inflow composition and in individual duration dependence. It hence relaxes two separability assumptions inherent in the mixed proportional hazard specification. The time at which an individual entered unemployment will be one influence on the hazard rate (embodied in the term $\Psi_4(t - d)$ below) as it is an indication of the quality of the inflow pool from which he comes. The cyclicity of the individual duration dependence

³It should be noted that, as studies on duration dependence often use data on exit rates from unemployment which include the number of people who exit unemployment to drop out of the labour force, the estimated duration dependence phenomenon might relate to some labour supply decisions as well as to job-finding behaviour.

is embodied in the term $\Psi_3(d, t)$. So our specification of the hazard rate is:

$$h(i, t, d) = \Psi_2(t) \cdot \Psi_3(d, t) \cdot \Psi_4(t - d) \cdot v_i \quad (2)$$

This allows the estimation of features of the duration dependence pattern and the unobserved heterogeneity distribution in terms of the following parameters. Three η coefficients represent the duration dependence pattern:

$$\eta_i = \frac{\Psi_3(i)}{\Psi_3(i-1)}, \text{ for } i = 1, 2, 3$$

and three γ coefficients represent the first moments of the heterogeneity distribution ($E_v \{v\} = 1$ without loss of generality):

$$\gamma_i = \frac{E_v \{v^i\}}{[E_v \{v\}]^i}, \text{ for } i = 1, 2, 3$$

The functional forms assumed for inflow composition and duration dependence are:

$$\Psi_4(t - d) = \lambda \cdot [bc(t - d)]^\alpha \quad (3)$$

and

$$\eta_i = \eta_i^0 + \beta_i \cdot bc(t) \quad (4)$$

hence

$$\Psi_3(d, t) = \prod_{s=1}^d [\eta_s^0 + \beta_s \cdot bc(t)] \quad \text{for } d = 1..3 \quad (5)$$

and $\Psi_3(0, t) = 1$

It has to be noted that these functional forms are a departure from the non-parametric setting of van den Berg and van Ours's model and is less flexible than Abbring, van den Berg and van Ours's (2002) polynomial specification⁴. The shape of the duration dependence pattern and of the heterogeneity distribution are still fully flexible, which ensures reliability of parameter estimates.

Here all heterogeneity is unobserved, and is fully represented by the term

⁴We did experiment with various functional forms for $\Psi_4(t-d)$ including the linear, logarithmic, exponential and polynomial forms, but the best results in terms of specification tests were obtained with the expression in (3)

v_i . A seasonal factor is introduced as follows:

$$G_s(v) = G_{s-1}(v \cdot w_s) \text{ where } \prod_s w_s = 1$$

$$E_v \{v\}_{s-1} = w_s \cdot E_v \{v\}_s$$

where s is the season of inflow.

Identification issues have been thoroughly examined by Abbring, van den Berg and van Ours (2001, 2002). With the above specifications for $\Psi_3(d, t)$ and $\Psi_4(t - d)$, the model in (2) will be identified⁵. The coefficient α will be estimated but λ is not identified as, in the expressions of the ratio of average exit rates from different duration bands d_1 and d_2 at the same calendar time t , only the ratio $\frac{\Psi_4(t-d_1)}{\Psi_4(t-d_2)}$ appears. Two different measures of the business cycle $bc(t)$ have been used: the tightness of the labour market (V/U , where V and U are respectively the stocks of vacancies and unemployment) or the unemployment rate. The best results were obtained when using the tightness rate in (4) and the unemployment rate in (3).

⁵The model is identified as long as the specification of $\Psi_4(t - d)$ does not contain a multiplicative exponential term in $t - d$, such as $\exp\{\rho \cdot (t - d)\}$, and the specification of $\Psi_3(d, t)$ does not include a multiplicative term in either t or $t - d$. For more details, see Abbring, van den Berg and van Ours (2001, Appendix C).

3 Data

3.1 Minimum size of group

Although the data are available at the level of very narrow age bands, and very small areas (travel-to-work areas), we use only three age bands, and 11 regions of Great Britain. This is to make sure that the observed hazard rates that we use are close enough to the actual ('theoretical') hazard rates. Hazard rates being probabilities, we cannot really observe them. If all individuals were facing the same hazard rate h of exiting unemployment, and if the size of the unemployment pool was infinite, then the fraction of individuals who exit unemployment would equal h . However, over a finite sample, or finite size of unemployment pool, all we can observe is an average of the realisations of a random variable X_i . For each individual, $X_i = 0$ if the individual stays unemployed, with probability $1 - h$, and $X_i = 1$ if the individual finds a job, with probability h . The expected value of X equals h , and its variance equals $h \cdot (1 - h)$. The observed hazard rate that we observe from an unemployment pool of size N is:

$$\hat{h} = \frac{1}{N} \cdot \sum_{i=1}^N X_i \quad (6)$$

which is just the sample mean of the realisations of X_i . As N increases, \hat{h} tends to its expected value, h . However, as we use the observed \hat{h} as data for the actual hazard rates h , we want \hat{h} to be ‘reasonably’ close to h , not to introduce too much measurement error (added to other types of measurement errors related to the measurement of unemployment). The variance of \hat{h} equals $\frac{h \cdot (1-h)}{N}$. From the central limit theorem, we know that \hat{h} tends in probability to a normally distributed random variable of mean h and variance $\frac{h \cdot (1-h)}{N}$. So, with 90% confidence, we can rely on the true h to lie within $1.7\sqrt{\frac{h \cdot (1-h)}{N}}$ of the observed \hat{h} .

The data we use for estimation are hazard rates up to two decimal places. So if \hat{h} lies within 0.01 of h , we can consider that it is ‘reasonably’ close to the true value in terms of the precision of our data. For a value of h of 0.5⁶, this means:

$$1.7\sqrt{\frac{h \cdot (1-h)}{N}} \leq 0.01 \tag{7}$$

$$\Leftrightarrow N \geq 7,225 \tag{8}$$

Hence we consider that 7,225 individuals is the minimum size of unemployment pool that we wish to use, in order for the observed hazard rate to reflect the actual hazard rate accurately enough. As we have data on five

⁶For smaller values of h , the condition on N is less restrictive.

quarterly duration bands, and that the size of these unemployment pools decreases with duration (as more and more workers have found a job and left the pool), we must make sure that the size of the unemployment pool of individuals in their fourth quarter of unemployment exceeds 7,225. This criterion guided our choice of age bands and geographical breakdown.

We have chosen to use an age breakdown comprising only three age bands: under 30 years old, 30 to 50 years old, and over 50 years old. The average size of the total unemployment in each age band, and of the number of unemployed in the fourth quarter of their unemployment spell are reported in Table 1. Our criterion on the size of the pool is satisfied for all age bands. The geographical breakdown used comprises 11 regions over Great Britain. The corresponding statistics are shown in Table 2. The size of regional unemployment varies from 45,349 for East Anglia to 258,217 for London. All regions but Wales and East Anglia satisfy the criterion on the value of N for the size of the fourth duration band. Wales is borderline with 7,046 individuals in the fourth duration band. East Anglia is much below the limit size we set ourselves. However, as this regional breakdown is the one that splits Great Britain into the largest possible areas, we are going to use this data - bearing in mind that data on hazard rates in East Anglia are less accurate than in other regions. We use male data only.

3.2 Description of the data

The averages of hazard rates for each duration band over 1985:IV to 1997:II are shown in Tables 3 and 4. For the regional data sets, as we can consider that, to some extent, different regions represent different labour markets, we will use the regional unemployment rate and the regional tightness rate to represent the state of the business cycle. In each individual data set, we observe a decline of average hazard rates with duration: this decline reflects the combination of the weeding out process and the duration dependence pattern. One exception is the increase in hazard rate in the data set of males age over 50, between the third and fourth duration band. As suggested below, this could be linked with some labour force participation decision. Apart from the hazard rate of older men in their fourth duration band, hazard rates decline with age, for a given duration.

4 Estimation on age group data

Results of the estimation with age group data are reported in Table 5. We carried out the regression on each data set separately. In the specifications (3) and (4) of the cyclical variations of the inflow composition and the duration dependence pattern, we keep the aggregate unemployment rate and tightness rate as indicators of the business cycle. This is because we do not

consider the different age groups as looking for jobs in separate labour markets. For the three data sets, the inflow composition varies significantly over the business cycle as the α coefficient is significant. Its positive sign suggests that these variations are countercyclical, in that individuals entering unemployment when unemployment is high have on average a better ability to exit unemployment than those who enter unemployment when it is low. The estimated α varies across groups: it is highest for middle-age men, and smallest for the older age group. This seems to indicate that the average quality of the inflow is more sensitive to the unemployment level in the middle-age group, and less sensitive in the older age group. The β coefficients are jointly insignificant for the under 30 group, but jointly significant for the other two groups. The duration dependence between the third and the fourth quarter of unemployment is smoother when the tightness is higher, i.e. in booms. This goes along the insights of Blanchard and Diamond's (1994) ranking model. These two findings suggest that the first two separability assumptions⁷ of the standard mixed proportional hazard model are not satisfied in our data.

For the three age group data sets the estimation results exhibit the same main features as the results obtained with aggregate data in Turon (2003): significantly negative duration dependence over the first year of unemploy-

⁷i.e. that the duration influence is separable from the business cycle influence, and that the heterogeneity distribution is invariant over the business cycle

ment, small or insignificant variance of the heterogeneity distribution, same seasonal pattern, and countercyclical inflow composition.

Males under 30 and males aged 30 to 50 experience very similar depreciation of their chances to find a job during their first year of unemployment. Males over 50 have a smoother decrease of their hazard rate in the first three quarters of unemployment, and their chances to exit unemployment in the fourth quarter is 30 percent higher than in the third quarter. Like in the female aggregate data set⁸, this suggests that some of them just drop out of the labour force, maybe to some kind of retirement. So for the older group, η_3 cannot be interpreted purely as a job search parameter, but rather a combination of this with labour force participation behaviour. One implicit assumption of the mixed proportional hazard specification of the model is that the duration dependence pattern does not depend on individual characteristics. If this separability assumption is right, the η coefficients should be the same for all age groups. Most of the differences described above are not significant, which gives support for this assumption: η_1 and η_2 do not vary significantly across groups, but η_3 is significantly higher in the over 50 group than in the other two. This could be attributed to the fact that η_3 is affected by labour participation decisions.

For the age groups of under 30 and 30 to 50, the γ_2 estimate is not signifi-

⁸See Turon (2003).

cantly different from 1, which means that the variance of v is not significantly different from zero. The absence of heterogeneity that this suggests is quite surprising, but is in line with van den Berg and van Ours's and our findings with aggregate data. With the data set of males aged over 50, however, the variance of the heterogeneity distribution is significantly positive. The seasonal pattern of the w coefficients is similar in all groups. w_{jan} is always significantly smaller than 1, while w_{apr} is always significantly higher than 1. This suggests that individuals entering unemployment in January are more successful than average in finding a job, but that people entering unemployment in the three other seasons have got very similar average ability to find a job at the start of their unemployment spell.

Specification tests have been carried out to check the validity of the model: first of all, the γ coefficients cannot theoretically take any value. Three constraints⁹ on their values have to hold in order for them to plausibly represent the first moments of a distribution with positive support. These constraints do not reject the model specification for any group. There is evidence of serial correlation in the third equation of the models, for all age groups. For

⁹These constraints are the following (see Shohat and Tamarkin, 1970):

$$\begin{aligned} \gamma_2 &\geq 1 \\ \gamma_3 &\geq \gamma_2^2 \\ [\gamma_4 - \gamma_2^2] (\gamma_2 - 1) - (\gamma_3 - \gamma_2)^2 &\geq 0 \end{aligned}$$

the first two equations, no serial correlation is observed in the data sets of males aged under 30, and aged 30 to 50. In the data sets of males aged over 50, however, serial correlation is present in all three equations.

5 Estimation on regional data

The model has also been estimated with the 11 regional data sets¹⁰. Two features of the regional data should enable us to test cross-influences between duration dependence and other factors. Firstly, the distribution of heterogeneity is likely to vary across regions, because the demographic characteristics as well as the distribution of skills in the workforce might not be the same between London and Yorkshire, for example. Similarly, the demand for various skills may well vary across regions, leading to regional variations in the ability to find a job. Secondly, as seen in Table 4, the state of the labour market varies significantly across regions at the same point in time. The average unemployment rate over the period covered by the data varies from around 7% in East Anglia and the South East to 15% in the Northern region. If the assumption of separability between duration dependence and heterogeneity is correct, we should obtain similar estimates of the η coeffi-

¹⁰Ideally, we would have liked to carry out a joint estimation of the model on the 11 regions, in order to test for the equality of coefficients across regions. However, attempts at a joint estimation have proved unsuccessful. We therefore estimate the model separately for each region and compare the coefficients obtained.

coefficients in all regions in spite of these differences. On the other hand, if the duration dependence pattern is sensitive to the state of the labour market, this should be revealed in the estimation of the model.

Estimated coefficients for each region are shown in Tables 6 and 7. Allowing for cyclical variations in inflow composition improves the fit of the model significantly compared to a standard mixed proportional hazard model as in (1) as the likelihood ratio test statistics LR_1 reported in Tables 6 and 7 show. This confirms that the influences of heterogeneity and calendar time on the hazard rates are not separable since the mean of the heterogeneity distribution fluctuates with the business cycle. On the other hand, allowing for cyclical variations in the duration dependence pattern does not improve the fit of the model significantly for most regions, as the likelihood ratio test statistics LR_2 show. This supports the separability assumption between the duration dependence and business cycle. For most regions, the duration dependence pattern can be considered constant over the business cycle. For Scotland, the South East and the West Midlands however, we do find significant variations of the duration dependence pattern, although of varying sign across regions.

The duration dependence pattern looks similar across regions and suggests significantly negative duration dependence over the first year of unemployment. As a joint estimation over the 11 regions was not possible, we

cannot test for the joint equality of the η parameters across regions. Point estimates of the η coefficients, together with their 5% confidence interval are shown in Figures 1 to 3. All the confidence intervals are overlapping. The estimated value of η_1 obtained with the aggregate male dataset 0.85 (line drawn on Figure 1), belongs to all confidence intervals of the η_1 parameter estimate across regions, except the South East. The corresponding values for η_2 and η_3 are 0.91 and 0.93 respectively (lines drawn on Figure 2 and 3), and belong to all the confidence intervals. This suggests that the duration dependence estimates across regions are ‘fairly close’ and offers some support for the separability assumption. As mentioned above, for most regions, the η coefficients are found to be constant over the business cycle. So this rejects the potential explanation that the observed variations of duration dependence across regions might be due to the different regions being at different stages of the business cycle. Another potential cause for different duration dependence lies in the differences in heterogeneity distribution across regions. If the differences in duration dependence pattern were significant, we would have to reject the assumption of separability between duration dependence and heterogeneity.

In his model of stigma, Berkovitch (1990) predicts that the greater the heterogeneity in a group, the stronger the negative duration dependence. Similarly, Lockwood (1991) shows that when the variance of the productivity

distribution among workers increases, the screening on duration on the firms' side becomes tighter, because firms make larger losses from low-productivity workers, so the incentive to detect them increases. So we should find lower η coefficients in regions where γ_2 is higher. As seen in Tables 6 and 7, the variance of the heterogeneity distribution, represented by $\gamma_2 - 1$, varies across regions. These variations are not large, as most variances are not significantly positive, although all point estimates (except the one for South East) are very near or greater than 1. This suggests insignificant or small presence of heterogeneity in the distribution of ability to find a job across individuals entering unemployment in a given region, which is a bit surprising, but again, mirrors our results with aggregate data. The product of the three duration dependence coefficients, i.e. the decrease in individual exit rate over the first year of unemployment, has been plotted against γ_2 (Figure 4). A clear positive relationship appears. The slope of the regression line is positive (0.52 with a standard error of 0.12), and this regression explains 68% of the observed variations in the duration dependence pattern.

This positive relationship does not confirm Berkovitch's or Lockwood's predictions. In their 1996 article, van den Berg and van Ours used data decomposed according to race (black or white) and gender. The group of black males, which exhibited the highest variance of unobserved heterogeneity, had a positive duration dependence pattern over the first months of

unemployment, whereas white males, with a lower variance, exhibited a negative duration dependence. So their results did not confirm Berkovitch's idea either. Our results across regions suggest that duration dependence is smoother in regions where the unobserved heterogeneity is greater. The intuitive explanation for this relationship is not clear. Also, as the variance of the heterogeneity distribution exhibits little variation across regions, our result should be interpreted with caution.

The effect of the season of inflow on exit rates, represented by the w 's is remarkably similar across model specifications and across regions. January is the season where most successful people enter unemployment.

For all regions, the estimated α coefficient is significantly positive, suggesting countercyclical variations in inflow composition. The estimated values of α vary from 0.35 to 0.99. The low values are obtained for London and the South East, while the high values are obtained for Scotland, the North West, and Yorkshire and Humberside. As with the η coefficients, we have plotted the estimated values of α across regions, together with their confidence interval (Figure 5). This plot suggests that there are some significant differences between the α coefficients across regions. To see whether the variations in α could possibly reflect variations in the labour market state, the estimated α 's have been plotted against the average unemployment rate in each region (Figure 6). The figure shows a positive relationship between

α and unemployment rate. The slope of the regression line between α and the unemployment rate is significantly positive (4.7 with a standard error of 2.1) and this regression explains 35% of the variations in α . In the regions that combine high unemployment and high α , like Scotland, Northern, North West and Yorkshire and Humberside, the average quality of the inflow into unemployment is much more sensitive to the unemployment rate than in the regions where both α and the unemployment rate are quite low, like the South East, South West, and West Midlands. In London, average unemployment is quite high (13.2%), but as the estimate of α is one of the lowest, inflow composition is not very sensitive to unemployment rate.

The test on the validity of the constraints on the γ coefficients does not reject these constraints, with the only exception of the South East. Serial correlation tests were carried out. While no evidence of serial correlation was found in the first two equations of the model (A.1) and (A.2) for most regions, there is some evidence of serial correlation in the third equation (A.3) for nearly all regions.

6 Conclusion

In this paper, estimation of a model allowing for cyclical variations in both the inflow composition and the duration dependence pattern has been carried

out with data disaggregated by age and by region. The duration dependence pattern is quite similar between age groups, except for the coefficient representing the change in individual exit rate between the third and fourth quarter of unemployment, which is much higher and greater than one for males over 50. This difference might be due to labour participation behaviour. The results with regional data show similar duration dependence patterns for all regions. This gives some support to the remaining separability assumption inherent to our model, which is that the influences of duration and of heterogeneity are separable. The variance of the heterogeneity distribution exhibits some small variation across age groups and regions but, as with the aggregate data, it is not significant in most cases. The duration dependence phenomenon over the first year of unemployment is smoother in regions where the variance of the heterogeneity distribution is greater.

Results obtained also allow us to compare the cyclical sensitivity of the inflow composition across groups, and relate it to the state of the local labour market in the case of regional data. This sensitivity seems to be positively correlated with the regional unemployment rate.

Appendix

The three equations to be estimated are in logarithmic form as follows¹¹:

$$\ln \left\{ \frac{h(1 | t)}{h(0 | t)} \right\} = \ln \{ \eta_1 \} + \ln \{ W_t \} + \ln \left\{ \frac{\Psi_4(t-1)}{\Psi_4(t)} \right\} \quad (\text{A.1})$$

$$+ \ln \{ 1 - \gamma_2 h(0, t-1) \} - \ln \{ 1 - h(0, t-1) \}$$

$$\ln \left\{ \frac{h(2 | t)}{h(0 | t)} \right\} = \ln \{ \eta_1 \eta_2 \} + \ln \{ W_t W_{t-1} \} + \ln \left\{ \frac{\Psi_4(t-2)}{\Psi_4(t)} \right\} \quad (\text{A.2})$$

$$+ \ln \left\{ \begin{array}{l} 1 - \gamma_2 h(0, t-2) \\ -\eta_1 W_{t-1} \frac{\Psi_4(t-2)}{\Psi_4(t-1)} h(0, t-1) \cdot [\gamma_2 - \gamma_3 h(0, t-2)] \end{array} \right\}$$

$$- \ln \left\{ \begin{array}{l} 1 - h(0, t-2) \\ -\eta_1 W_{t-1} \frac{\Psi_4(t-2)}{\Psi_4(t-1)} h(0, t-1) \cdot \{ 1 - \gamma_2 h(0, t-2) \} \end{array} \right\}$$

$$\ln \left\{ \frac{h(3 | t)}{h(0 | t)} \right\} = \ln \{ \eta_1 \eta_2 \eta_3 \} + \ln \{ W_t W_{t-1} W_{t-2} \} + \ln \left\{ \frac{\Psi_4(t-3)}{\Psi_4(t)} \right\} \quad (\text{A.3})$$

$$+ \ln \left\{ \begin{array}{l} 1 - \gamma_2 h(0, t-3) + \\ \eta_1 W_{t-2} \frac{\Psi_4(t-3)}{\Psi_4(t-2)} \cdot [\gamma_2 - \gamma_3 h(0, t-3)] \cdot \\ \left[\eta_2 W_{t-1} \frac{\Psi_4(t-2)}{\Psi_4(t-1)} h(0, t-1) + h(0, t-2) \right] \\ + [\gamma_3 - \gamma_4 h(0, t-3)] \cdot \\ \eta_1^2 \eta_2 W_{t-2}^2 W_{t-1} \frac{\Psi_4(t-3)^2}{\Psi_4(t-2) \Psi_4(t-1)} h(0, t-1) h(0, t-2) \end{array} \right\}$$

$$- \ln \left\{ \begin{array}{l} 1 - h(0, t-3) + \\ \eta_1 W_{t-2} \frac{\Psi_4(t-3)}{\Psi_4(t-2)} \cdot [1 - \gamma_2 h(0, t-3)] \cdot \\ \left[\eta_2 W_{t-1} \frac{\Psi_4(t-2)}{\Psi_4(t-1)} h(0, t-1) + h(0, t-2) \right] \\ + [\gamma_2 - \gamma_3 h(0, t-3)] \cdot \\ \eta_1^2 \eta_2 W_{t-2}^2 W_{t-1} \frac{\Psi_4(t-3)^2}{\Psi_4(t-2) \Psi_4(t-1)} h(0, t-1) h(0, t-2) \end{array} \right\}$$

where

$$W_t = \sum_{s=oct}^{jul} w_s I_{t,s}, \text{ with } I_{t,s} = \begin{cases} 1 & \text{if season at } t = oct \\ 0 & \text{otherwise} \end{cases}$$

¹¹Although obtained with van den Berg and van Ours's (1994) method, the derivation of these equations is ours.

References

- [1] Abbring, J., Van den Berg, G. and J. Van Ours, 2002, The anatomy of unemployment dynamics, *European Economic Review* 46, 1785-1824.
- [2] Abbring, J., Van den Berg, G. and J. Van Ours, 2001, Business cycles and compositional variation in US unemployment, *Journal of Business and Economic Statistics* 19, 436-448.
- [3] Arulampalam, W., Booth, A. and M. Taylor, 2000, Unemployment persistence, *Oxford Economic Papers* 52, 24-50.
- [4] Baker, M., 1992, Unemployment duration: Compositional effects and cyclical variability, *American Economic Review* 82, 313-321.
- [5] Berkovitch, E., 1990, A stigma theory of unemployment duration in: Weiss, Y. and Fishelson, G., eds., *Search unemployment: Theory and measurement*, 20-56.
- [6] Blanchard, O. and P. Diamond, 1994, Ranking, unemployment duration, and wages, *Review of Economic Studies* 61, 417-434.
- [7] Devine, T. and N. Kiefer, 1991, *Empirical labor economics* (Oxford University Press, New-York).
- [8] Heckman, J. and B. Singer, 1984, A method for minimizing the impact of distributional assumptions in econometric models for duration data, *Econometrica* 52, 271-320.
- [9] Kalwij, A., 2001, Individuals' unemployment durations over the business cycle, IZA Discussion Paper No 369.
- [10] Lancaster, T., 1990, *The econometric analysis of transition data* (Cambridge University Press, Cambridge).
- [11] Jackman, R. and R. Layard, 1991, Does long-term unemployment reduce a person's chance of a job? A time-series test, *Economica* 58, 93-106.
- [12] Layard, R., Nickell, S., and R. Jackman, 1991, *Unemployment* (Oxford University Press, Oxford).
- [13] Lockwood, B., 1991, Information externalities in the labour market and the duration of unemployment, *Review of Economic Studies* 58, 733-753.
- [14] Machin, S. and A. Manning, 1999, The causes and consequences of longterm unemployment in Europe, in O. Ashenfelter and R. Layard, eds., *Handbook of Labor Economics*, Vol. 3 (North Holland, Amsterdam) 3085-3139.
- [15] Meyer, B., 1990, Unemployment insurance and unemployment spells, *Econometrica* 58, 757-782.

- [16] Shohat, J., and J. Tamarkin, 1970, The problem of moments (American Mathematical Society, Providence).
- [17] Turon, H., 2003, Inflow composition, duration dependence and their impact on the unemployment outflow rate, Oxford Bulletin of Economics and Statistics, forthcoming.
- [18] Van den Berg, G. and Van Ours, J., 1994, Unemployment dynamics and duration dependence in France, The Netherlands, and the United Kingdom, Economic Journal 104, 432-443.
- [19] Van den Berg, G. and Van Ours, J., 1996, Unemployment dynamics and duration dependence, Journal of Labor Economics 14, 100-125.

Table 1: Average size of age data sets

	Males under 30	Males aged 30-50	Males over 50
Average U	759,249	657,386	322,469
Average $U(Q4)$	60,641	48,135	24,532

Averages over the sample period: 1985.IV to 1997.II.

$U(Q4)$ is the size of the unemployment pool for the fourth duration band.

Table 2: Average size of regional data sets

	Average U	Average $U(Q4)$
East Anglia	45,349	3,437
East Midlands	111,865	8,525
London	258,217	21,150
North West	230,329	17,138
Northern	126,089	9,203
Scotland	186,392	13,727
South East	225,906	18,032
South West	117,641	9,161
Wales	92,854	7,046
West Midlands	174,709	13,207
Yorkshire and Humberside	169,753	12,682

Averages over the sample period: 1985.IV to 1997.II.

$U(Q4)$ is the size of the unemployment pool for the fourth duration band.

Table 3: Average hazard rates - Age data sets

	Males under 30	Males aged 30-50	Males over 50
$h(0)$	0.45	0.40	0.35
$h(1)$	0.36	0.31	0.27
$h(2)$	0.33	0.28	0.24
$h(3)$	0.28	0.23	0.29

Averages over the sample period: 1985.IV to 1997.II.

Table 4: Average hazard rates, unemployment rates, tightness rates -
Regional data sets

	$h(0)$	$h(1)$	$h(2)$	$h(3)$	$Urate$	V/U
East Anglia	0.47	0.37	0.33	0.30	0.08	0.16
East Midlands	0.43	0.34	0.30	0.27	0.10	0.11
London	0.38	0.30	0.27	0.24	0.13	0.09
North West	0.41	0.33	0.29	0.26	0.13	0.10
Northern	0.42	0.33	0.30	0.26	0.15	0.07
Scotland	0.43	0.34	0.31	0.27	0.13	0.12
South East	0.46	0.36	0.32	0.29	0.07	0.20
South West	0.46	0.36	0.33	0.30	0.09	0.15
Wales	0.42	0.34	0.31	0.27	0.13	0.13
West Midlands	0.40	0.31	0.29	0.26	0.12	0.10
Yorkshire and Humberside	0.43	0.33	0.30	0.27	0.12	0.08

Averages over the sample period: 1985.IV to 1997.II.

Table 5: Results by male age group, for the three model specifications

	under 30	30 to 50	over 50	Aggregate
η_1	0.85 (0.04)	0.83 (0.05)	0.87 (0.04)	0.85 (0.04)
η_2	0.91 (0.04)	0.88 (0.04)	0.94 (0.04)	0.91 (0.04)
η_3	0.84 (0.03)	0.90 (0.03)	1.34 (0.08)	0.93 (0.04)
γ_2	1.08 (0.07)	1.09 (0.10)	1.26 (0.09)	1.12 (0.08)
γ_3	1.34 (0.26)	1.16 (0.40)	1.79 (0.41)	1.38 (0.32)
γ_4	2.28 (0.78)	1.04 (1.31)	2.86 (1.50)	2.07 (1.01)
w_{oct}	1.01 (0.02)	0.97 (0.01)	1.00 (0.01)	1.00 (0.01)
w_{jan}	0.97 (0.01)	0.95 (0.02)	0.92 (0.01)	0.96 (0.01)
w_{apr}	1.05 (0.01)	1.08 (0.01)	1.11 (0.01)	1.07 (0.01)
α	0.48 (0.07)	0.74 (0.09)	0.23 (0.07)	0.56 (0.07)
β_1	0.03 (0.32)	-0.21 (0.25)	0.16 (0.40)	-0.03 (0.29)
β_2	0.11 (0.44)	1.22 (0.43)	1.01 (0.44)	0.55 (0.46)
β_3	0.96 (0.63)	-0.17 (0.42)	-0.14 (0.99)	0.39 (0.56)
SC1	√	√	X	√
SC2	√	√	X	√
SC3	X	X	X	X

Standard errors in brackets. 42 observations.

SC1, 2, and 3: Result of the test on residuals' serial correlation. X = evidence of serial correlation, √ = no serial correlation.

Table 6: Results by region (1/2)

	EA	EM	L	NW	N	Sc
η_1	0.79 (0.03)	0.82 (0.04)	0.91 (0.05)	0.87 (0.05)	0.83 (0.04)	0.81 (0.04)
η_2	0.89 (0.02)	0.89 (0.03)	0.95 (0.06)	0.95 (0.05)	0.95 (0.04)	0.99 (0.04)
η_3	0.94 (0.03)	0.94 (0.04)	0.87 (0.04)	0.92 (0.04)	0.88 (0.05)	1.00 (0.04)
γ_2	0.99 (0.06)	1.10 (0.07)	1.25 (0.11)	1.17 (0.09)	1.05 (0.09)	1.11 (0.07)
γ_3	0.86 (0.28)	1.36 (0.30)	2.18 (0.34)	1.65 (0.38)	1.07 (0.40)	1.21 (0.28)
γ_4	0.15 (0.95)	2.11 (0.96)	5.31 (1.00)	3.22 (1.16)	1.22 (1.35)	1.01 (0.92)
w_{oct}	0.97 (0.01)	0.99 (0.01)	1.01 (0.02)	1.00 (0.01)	0.99 (0.02)	0.98 (0.01)
w_{jan}	1.01 (0.01)	0.98 (0.01)	0.97 (0.02)	0.98 (0.01)	0.98 (0.02)	1.00 (0.01)
w_{apr}	1.07 (0.01)	1.08 (0.01)	1.04 (0.01)	1.07 (0.01)	1.07 (0.01)	1.06 (0.01)
α	0.56 (0.05)	0.55 (0.06)	0.40 (0.09)	0.85 (0.12)	0.75 (0.18)	0.99 (0.14)
β_1	-0.08 (0.11)	0.22 (0.23)	0.25 (0.38)	0.21 (0.32)	-0.28 (0.32)	0.39 (0.20)
β_2	0.37 (0.19)	0.15 (0.30)	-0.59 (0.32)	-0.08 (0.40)	0.24 (0.59)	0.31 (0.27)
β_3	-0.33 (0.21)	-0.07 (0.42)	0.74 (0.41)	0.98 (0.54)	1.22 (0.78)	-0.50 (0.32)
LR_1	82.8	67.6	24.2	48.4	26.1	42.7
LR_2	5.0	2.4	5.4	5.2	4.2	9.4
SC1	✓	✓	✓	X	✓	✓
SC2	✓	✓	X	✓	✓	✓
SC3	X	✓	X	X	X	X

Standard errors in brackets. 44 observations.

EA: East Anglia, EM: East Midlands, L: London, NW: North West, N: Northern, Sc: Scotland.

LR_1 : Likelihood ratio test between the model with varying inflow composition and the standard MPH specification. LR_1 should be compared with $\chi_1^2 = 3.8$ (one variable added)

LR_2 : Likelihood ratio test between the model with varying inflow composition and duration dependence and the model with varying inflow composition only.

LR_2 should be compared with $\chi_1^2 = 7.8$ (3 variables added).

SC1, 2, and 3: Result of the test on residuals' serial correlation. X = evidence of serial correlation, ✓ = no serial correlation.

Table 7: Results by region (2/2)

	SE	SW	Wa	WM	YH	Aggregate
η_1	0.78 (0.03)	0.80 (0.04)	0.80 (0.04)	0.79 (0.04)	0.81 (0.04)	0.85 (0.04)
η_2	0.88 (0.02)	0.88 (0.03)	0.89 (0.04)	0.91 (0.03)	0.93 (0.04)	0.91 (0.04)
η_3	0.92 (0.02)	0.91 (0.03)	0.88 (0.04)	0.93 (0.03)	0.92 (0.04)	0.93 (0.04)
γ_2	0.96 (0.06)	0.98 (0.08)	0.99 (0.08)	1.02 (0.09)	1.07 (0.07)	1.12 (0.08)
γ_3	0.72 (0.25)	0.88 (0.34)	1.08 (0.29)	0.90 (0.48)	1.16 (0.31)	1.38 (0.32)
γ_4	-0.18 (0.85)	0.49 (1.09)	1.83 (1.06)	0.78 (1.83)	1.45 (1.10)	2.07 (1.01)
w_{oct}	0.98 (0.01)	0.99 (0.01)	0.98 (0.01)	1.00 (0.01)	0.99 (0.01)	1.00 (0.01)
w_{jan}	1.00 (0.01)	0.99 (0.02)	1.02 (0.02)	0.96 (0.01)	0.98 (0.01)	0.96 (0.01)
w_{apr}	1.05 (0.01)	1.07 (0.01)	1.05 (0.01)	1.06 (0.01)	1.07 (0.01)	1.07 (0.01)
α	0.35 (0.04)	0.48 (0.05)	0.70 (0.09)	0.42 (0.07)	0.72 (0.11)	0.56 (0.07)
β_1	-0.14 (0.06)	-0.21 (0.14)	-0.07 (0.14)	-0.08 (0.21)	0.04 (0.28)	-0.03 (0.29)
β_2	0.39 (0.12)	0.26 (0.20)	-0.14 (0.23)	0.62 (0.41)	0.32 (0.50)	0.55 (0.46)
β_3	0.11 (0.13)	-0.04 (0.21)	-0.21 (0.30)	1.18 (0.43)	0.65 (0.52)	0.39 (0.56)
LR_1	63.2	67.7	48.0	37.9	52.2	54.4
LR_2	17.8	2.4	2.2	15.2	4.4	3.5
SC1	✓	✓	✓	X	✓	✓
SC2	✓	X	X	✓	✓	✓
SC3	X	✓	X	✓	✓	X

Standard errors in brackets. 44 observations.

SE: South East, SW: South West, Wa: Wales, WM: West Midlands, YH: Yorkshire and Humberside.

LR_1 : Likelihood ratio test between the model with varying inflow composition and the standard MPH specification. LR_1 should be compared with $\chi_1^2 = 3.8$ (one variable added)

LR_2 : Likelihood ratio test between the model with varying inflow composition and duration dependence and the model with varying inflow composition only.

LR_2 should be compared with $\chi_1^2 = 7.8$ (3 variables added).

SC1, 2, and 3: Result of the test on residuals' serial correlation. X = evidence of serial correlation, ✓ = no serial correlation.

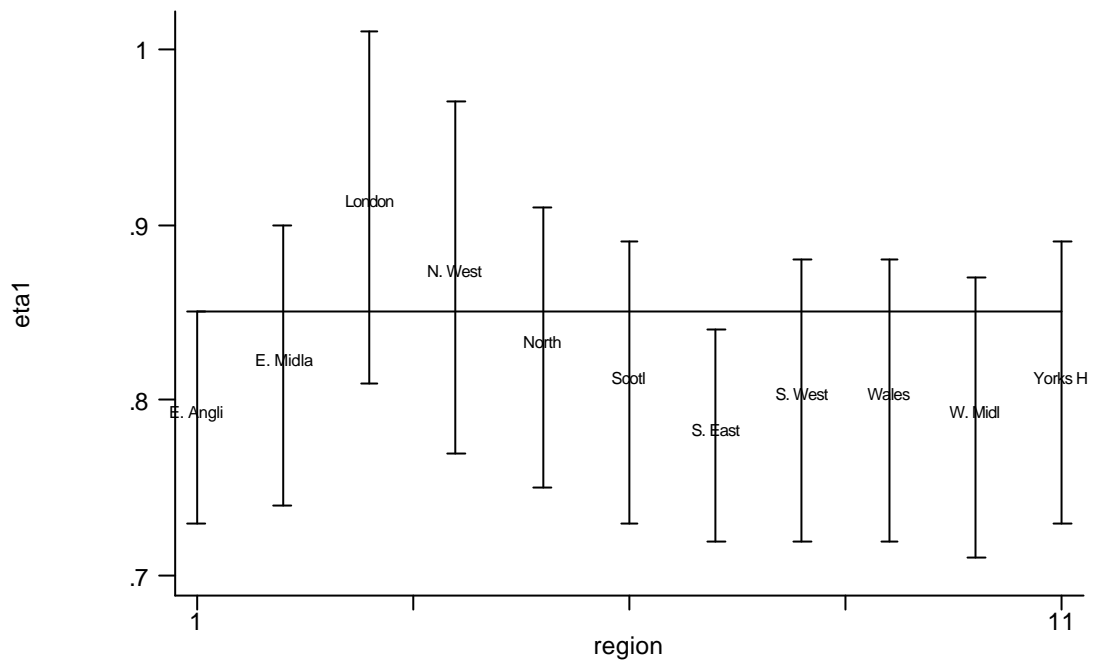


Figure 1: Estimates of h_1 and confidence intervals

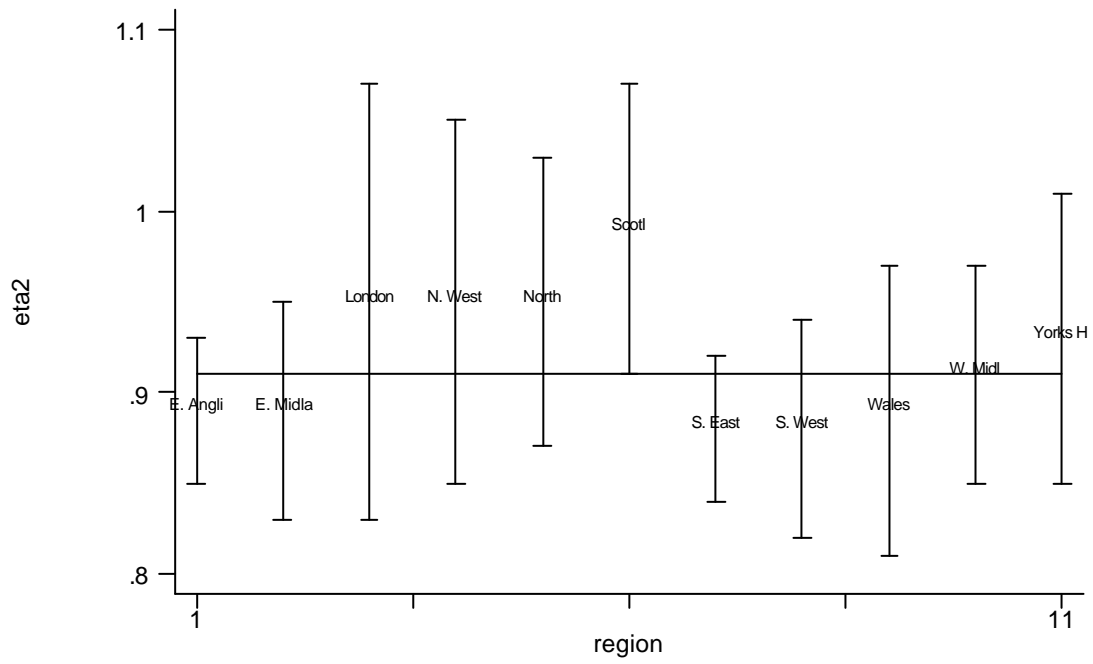


Figure 2: Estimates of h_2 and confidence intervals

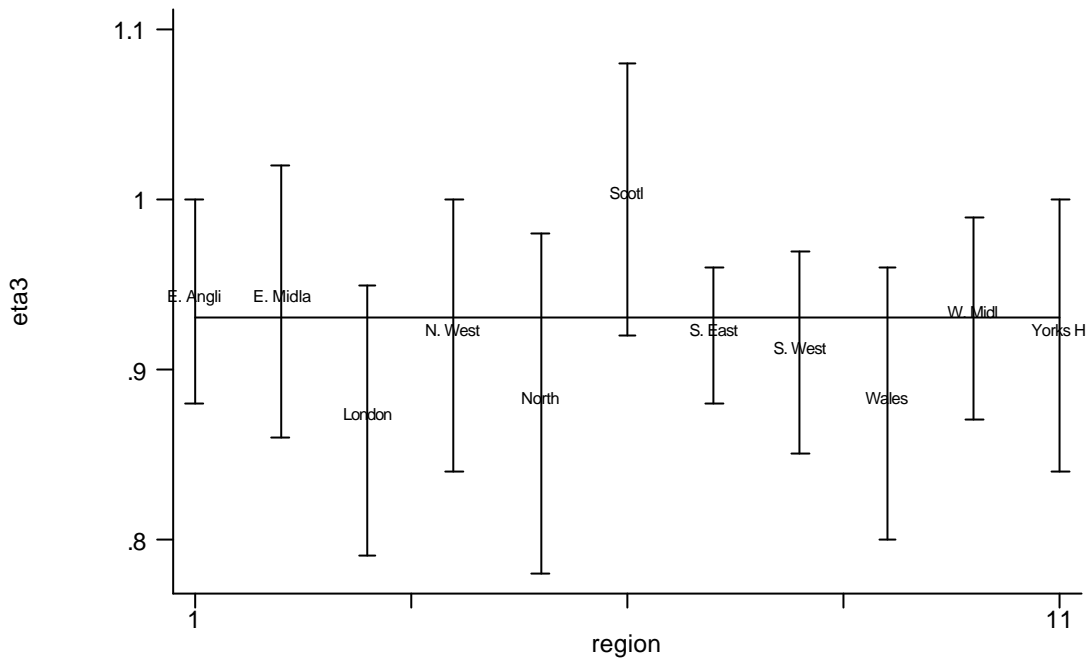


Figure 3: Estimates of h_3 and confidence intervals

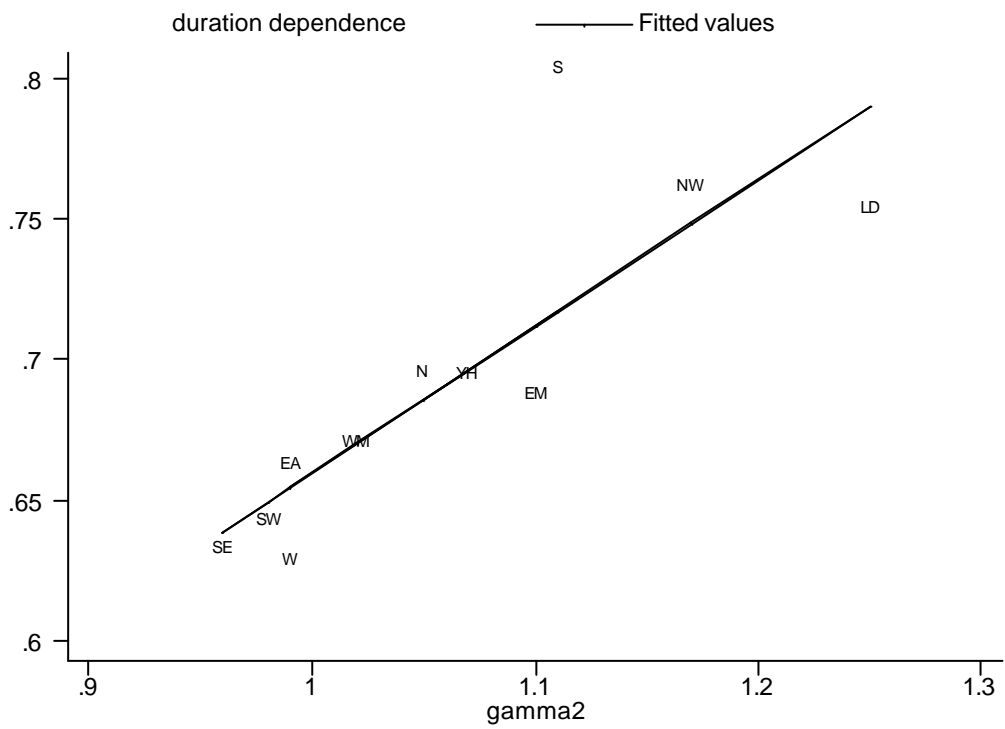


Figure 4: Duration dependence and unobserved heterogeneity

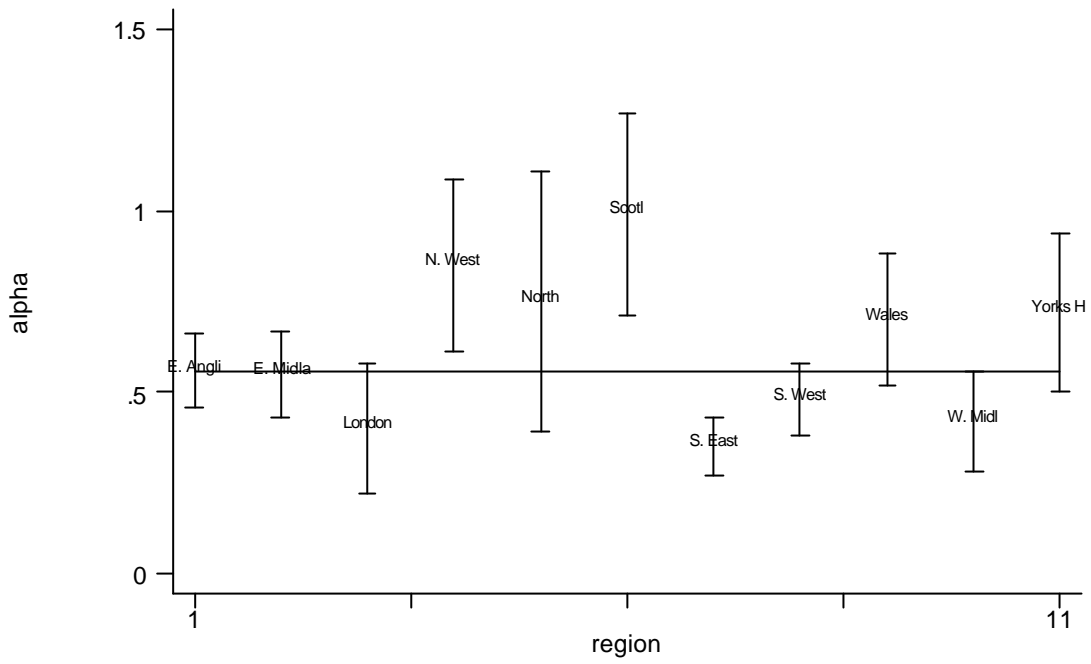


Figure 5: Estimates of α and confidence intervals

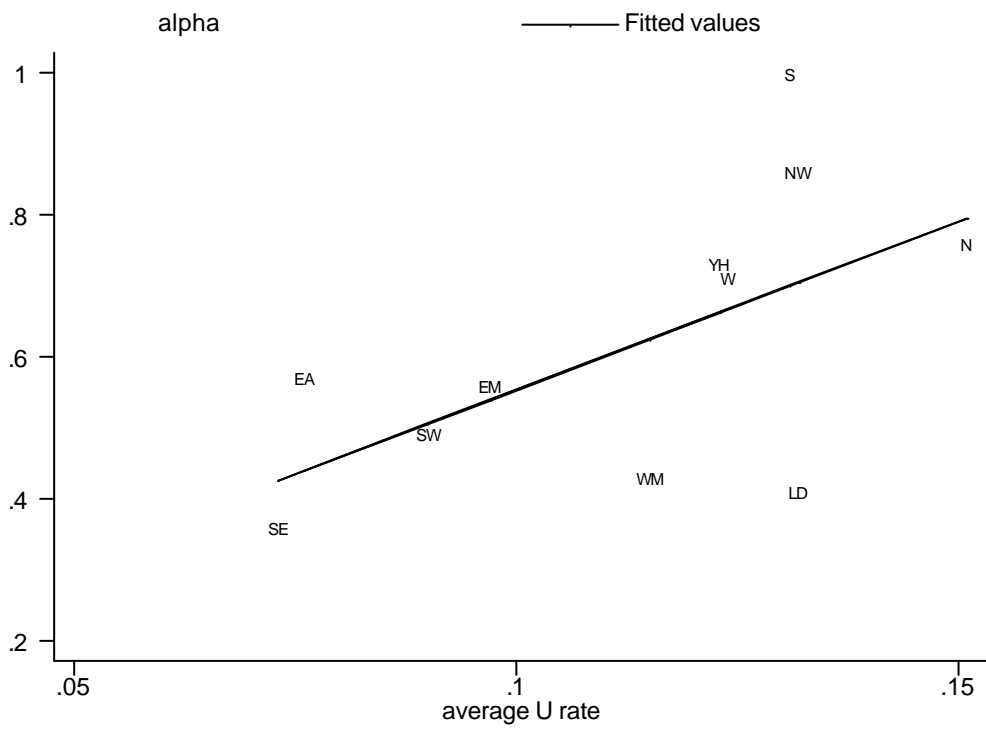


Figure 6: Estimates of α and unemployment rate

IZA Discussion Papers

No.	Author(s)	Title	Area	Date
738	V. Grossmann	Managerial Job Assignment and Imperfect Competition in Asymmetric Equilibrium	1	03/03
739	M. Fertig	Who's to Blame? The Determinants of German Students' Achievement in the PISA 2000 Study	1	03/03
740	B. T. Hirsch S. L. Mehay	Evaluating the Labor Market Performance of Veterans Using a Matched Comparison Group Design	6	03/03
741	B. T. Hirsch D. A. Macpherson	Wages, Sorting on Skill, and the Racial Composition of Jobs	5	03/03
742	R. A. Easterlin	Building a Better Theory of Well-Being	3	03/03
743	G. Heineck J. Schwarze	Substance Use and Earnings: The Case of Smokers in Germany	5	03/03
744	M. R. West L. Woessmann	Which School Systems Sort Weaker Students into Smaller Classes? International Evidence	2	03/03
745	L. Woessmann	Educational Production in East Asia: The Impact of Family Background and Schooling Policies on Student Performance	2	03/03
746	A. Ammermueller H. Heijke L. Woessmann	Schooling Quality in Eastern Europe: Educational Production During Transition	4	03/03
747	J. Messina	Sectoral Structure and Entry Regulations	3	03/03
748	M. Pflüger	Economic Integration, Wage Policies and Social Policies	2	04/03
749	V. Jakobsen N. Smith	The Educational Attainment of the Children of the Danish 'Guest Worker' Immigrants	1	04/03
750	R. Hujer M. Caliendo S. Thomsen	New Evidence on the Effects of Job Creation Schemes in Germany - A Matching Approach with Threefold Heterogeneity	6	04/03
751	L. Borghans B. ter Weel	Are Computer Skills the New Basic Skills? The Returns to Computer, Writing and Math Skills in Britain	5	04/03
752	N. Malchow-Møller J. R. Skaksen	Skill-Biased Technological Change in Denmark: A Disaggregate Perspective	5	04/03
753	S. Burgess H. Turon	Unemployment Equilibrium and On-the-Job Search	1	04/03
754	H. Turon	Separability of Duration Dependence and Unobserved Heterogeneity	1	04/03

An updated list of IZA Discussion Papers is available on the center's homepage www.iza.org.