The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Discussion Papers in Economics

Discussion Paper
No. 09/03

Job seeker's allowance in Great Britain: How does the regional labour market affect the duration until job finding?

Philip Ball & Ralf A. Wilke

September 2009
Reposted January 2011

2003 DP 09/03

Job seeker's allowance in Great Britain: How does the regional labour market affect the duration until job finding?*

Philip Ball†
Ralf A. Wilke‡

September 2009

Abstract

Employing a large individual-level administrative dataset from Great Britain, covering the period 1999-2007, we analyse the factors influencing the length of unemployment benefits claimant periods with subsequent transition to re-employment. To this end, this individual-level data is merged with a group of regional indicators to control for relevant regional labour market characteristics. From a methodological point of view, we adopt a flexible censored quantile regression approach to estimating conditional re-employment hazards. Our results indicate that the individual characteristics of an unemployed person are generally more important than the regional labour market conditions. However, there are important differences between re-employment hazards across several regions. Large cities such as London and Birmingham provide the worst local labour market conditions for job seekers allowance recipients, while remote regions like the Shetland islands perform among the best.

Keywords: benefit duration, quantile regression, hazard rate

JEL: C41, J64, J65

---

*This work is supported by the Economic and Social Research Council through the grant *Bounds for Competing Risks Duration Models using Administrative Unemployment Duration Data* (RES-061-25-0059). P. Wright, B. Fitzenberger, C. Dustmann, P. Dolton, as well as seminar participants at the ZEW Summer School 2008, the NIESR, the PSI London, the SES Annual Conference 2009, and at the DWP Annual Conference 2009 are thanked for their feedback.

†University of Nottingham, E–mail: lexpbb1@nottingham.ac.uk

‡University of Nottingham, PSI London, ZEW Mannheim, E–mail: ralf.wilke@nottingham.ac.uk

1 Introduction

The UK Government's initiative to boost employment sustainability, through 'Welfare to Work', highlights the need for accurate and rigorous analysis into the impact of national employment schemes at the regional and individual level in an integrated framework. For this reason, we develop a comprehensive database matching individual-level unemployment benefit claimant periods from the Joint Unemployment & Vacancies Operating System (JUVOS) to a rich set of regional indicators from sources such as NOMIS and the Department of Work & Pensions (DWP) on a monthly basis. This data is then mapped to the UK geography using the National Statistics Postcode Directory (NSPD), available from UK Borders, allowing the spatial characteristics of regions to be identified. Data on further education institutions and unemployment benefit office locations is included to capture the relevant supply, demand, as well as structural, social and institutional factors of interest. This database allows one to conduct research at a highly disaggregated local authority level in order to answer policy relevant questions. In this paper we link the individual-level JUVOS data to the regional context in which unemployment benefit claimants reside, rather than simply parameterizing regional heterogeneity through fixed effects.

Differences in labour market institutions are cited as a major explanation of unemployment disparities between countries. However, although institutions do not vary markedly between regions, there is considerable variation in UK regional unemployment rates (incidence) and individuals' experiences (durations). For instance, regional -local authority- ILO[1] unemployment rates varied from 3.3% to 14% over the year 2005. Furthermore, the greater spread in unemployment rates at lower levels of aggregation in the UK is well documented in the literature (Brown and Sessions, 1997; Collier, 2005). Figure 1 illustrates the unconditional distribution of median unemployment durations across Great Britain over the period of investigation: 1999 to 2005.

The importance of the job offer arrival rate in explaining average unemployment durations has been highlighted in the theoretical literature (Cahuc and Zylberberg, 2004). This phenomenon has been extensively studied using individual-level unemployment duration data. However, Collier's results suggest the regional context to be significant. Despite the vast unemployment duration literature, there are surprisingly few studies which explicitly take into account the regional context. Most use parametric approaches, and regional effects are implicitly accounted for in some studies (for the UK see Kalwij, 2004; Brown and Sessions, 1997; for the Netherlands see Folmer and van Dijk, 1988) via fixed effects.

[1]International Labour Organisation.

Some studies have looked at the impact of regional-level indicators like local unemployment rates and local labour market tightness on individuals' unemployment experiences (e.g. Meyer, 1990 for the US; Petrongolo, 2001 for the UK) however very few studies have analysed individual unemployment duration at the UK regional level. We are only aware of Collier's study which focusses exclusively on the county of Kent (Collier, 2005). Adopting a structural job search model and using detailed (unique) individual-level survey data, the author concludes that differences in regional labour market characteristics (notably regional variation in job offer arrival rates) may matter more than individual heterogeneity for unemployment experiences. This result is in contrast to more recent results for other countries. Using detailed individual-level administrative data, Arntz and Wilke (2009) do not observe a strong effect of the regional labour market on unemployment duration in Germany. They conclude that regional policies may have a smaller effect than commonly thought.

Theoretical job search literature models the individual job finding probability as a function of the job offer arrival probability as well as the probability of job offer acceptance. The former will be influenced by individual productivity, human capital accumulated, and local demand conditions whereas the latter will be influenced by individuals' reservation wages as well as local demand conditions (Petrongolo, 2001)[2]. Given the attempt to model the regional environment in which individuals live and conduct their job search, the relevant local demand conditions will be those of self contained local labour markets, which Petrongolo (2001) approximates by using regional indicators at the 'Travel-To-Work-Areas' (TTWA) level of aggregation. Her study reaches the conclusion that regional labour market tightness is negatively related to, whereas the stock of jobseekers in the region of residence impacts positively on, individual re-employment probability (Petrongolo, 2001). This result is found to be insignificant for females, which the author suggest could be an artifact of the data source (unemployment benefits offices) and heterogeneity in job search strategies by gender.

Meyer's (1990) results suggest that whilst higher local unemployment rates may have a significant spell lengthening effect, *ceteris paribus*, over time an increase in local unemployment rates could actually shorten spells as layoffs increase during economic downturns and these job separations are precisely the type that carry the least 'stigma' in terms of future re-employment probability. Furthermore, the advantage of the legislated requirement of two weeks written notice before termination of contract should give laid off individuals the added advantage of an early

[2]Lancaster (1979) proxies the former by the local unemployment rate in the region where the individual resides and the latter by the individual reservation wage.

start to job search (Arulampalam, 2001), relative to other job separation types.

Looking at regional unemployment in the UK, Martin (1997) provides suggestive evidence, via cointegration analysis, that the pattern of regional unemployment disparities exhibited significant geographical persistence since the 1960s. This is of great concern given that Gregg's (2001) results suggest that individuals experiencing unemployment earlier on in their life are more likely to experience it later on. Furthermore, Atkinson and Micklewright (1991) highlight that being unemployed at time t makes it more likely to be unemployed at time $t+1$ ('negative duration dependence'). Petrongolo (2001) finds strong evidence of negative duration dependence in the UK. However, it is important to distinguish between spurious & genuine state dependency (Collier, 2005), as both genuine state dependency and unobserved ability of the unemployed can explain the observation of 'negative duration dependence'. Controlling for unobserved heterogeneity will avoid spurious correlations between the probability of leaving unemployment and elapsed duration (Lancaster, 1979). Using a Mixed Proportional Hazard model, van den Berg and van Ours (1994) found evidence of negative duration dependence for UK men, whereas heterogeneity was insignificant. This result accords with that of Petrongolo (2001) for both UK men and women. Table 1 summarises the aforementioned literature in the context of regional effects.

For these reasons we see scope to exceed the previous work in two aspects. Our data set is richer, using individual-level administrative unemployment benefit claim periods linked with to institutional & regional variables at a low level of aggregation. As an empirical strategy Figure 1 suggests that our approach is informative as, after conditioning on observed factors in the Cox proportional hazard model, we observe quite a different distribution of unemployment durations relative to the unconditional distribution. From a methodological point of view, we adopt a flexible censored quantile regression approach to estimating conditional re-employment hazards. The quantile regression framework allows us to capture different effects on short- and long-term claimant periods in the same model. In addition, this approach is more flexible than standard techniques, as even in the case of the semi-parametric Cox proportional hazards (PH) model (Cox, 1972) the sign of the effect of a regressor is restricted to be the same across all quantiles of the conditional distribution. Rather than the usual conditional mean, our approach employs a conditional quantile function which is unaffected by outlier observations. This implies that results are also robust to the shape of the error distribution.

Non-parametric conditional hazard rates are estimated from the quantile regression estimates using a resampling method similar to Machado et al. (2006). Since this econometric model imposes less structure, the resulting conditional hazard rates can be disproportional and they can even cross. Our estimation results obtained by the censored quantile regressions provide evidence of several violations of the proportional hazard assumption.

The structure of the paper is as follows. The next section provides a detailed account of the relevant institutional setup. Following this, we briefly cover the data set construction, variable selection[3], as well as the individual and regional level data included[4]. The methodology exploited as well as the empirical results are considered in the following sections. Subsequently, relevant policy implications are detailed in light of the analysis.

[3]For a detailed exposition of the data preparation steps, see Ball (2009).

[4]The procedure for linking the individual & regional levels is documented in appendix A2.

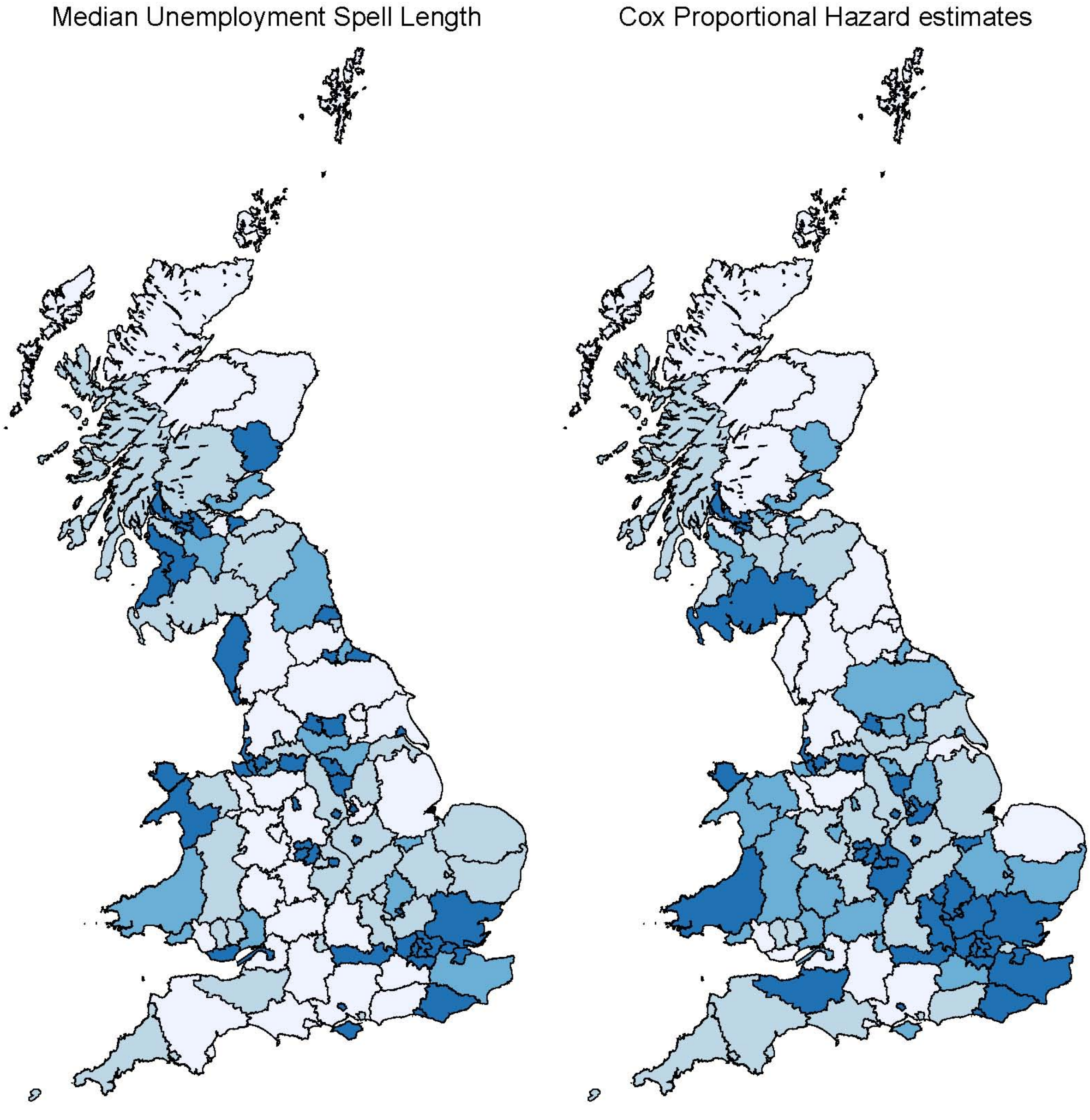


Figure 1: Distribution of median unemployment spell length, versus conditional Cox Proportional Hazard estimates*, by NUTS3 region over the period 1999-2005. (*Darker = Worse in terms of unemployment experiences)

Table 1: Summary of existing literature.

Author(s)	Country	Sample	Regional Controls?	Individual vs. Regional level?
Meyer (1990)	US	CWBH (Admin. data), 1978-83	Unemployment rate & fixed effects	Regional effects = significant
Lancaster (1979)	UK	Cross-section, Representative Survey, 1973.	Unemployment rate	Significant -VE effect, however significance is sensitive to specification.
Petrongolo (2001)	UK	Cross-section, Representative Survey, 1987	Unemployment/ Vacancies & Stock of jobseekers	Regional effects = Significant (for men)
Folmer/Van Dijk (1988)	Netherlands	Cross-section, Representative Survey, 1979	Fixed effects	Regional effects = insignificant
Brown/Sessions (1997)	UK	Non-random Survey, 1986-91.	Fixed effects	Regional effects = significant
Kalwij (2004)	UK	JUVOS (admin. data), Males, 18-34, 1983-98	Fixed effects	Regional effects = insignificant
Arntz/Wilke (2009)	Germany	IEBS (individual merged admin. data), 2000-02	Regional-level variables & fixed effects	Individual level = more important
Collier(2005)	England (Kent)	Cross-section†, Survey, 1992.	Regional-level variables & fixed effects	Regional level = more important

CWBH: Continuous Wage and Benefit History Unemployment Insurance Records;
JUVOS = Joint Unemployment & Vacancies Operating System;

2 Institutional setup

Unemployment benefits (Job Seeker's Allowance, JSA) are administered by the Jobcentre Plus which is a part of the Department for Work and Pensions (DWP). As in many other countries, the number of people on unemployment benefits in the UK and the number of people unemployed according to the International Labour Organisation's (ILO) definition do not necessarily coincide.

Jobseeker's Allowance is the main benefit for people who are out of work. In order to get Job Seeker's Allowance, an individual must be able to work for at least 40 hours a week and have been actively looking for work. There are two types of JSA: The first is called 'Contribution-based Jobseeker's Allowance' and lasts for up to six months (182 days), subject to eligibility. An unemployed person gets Contribution-based Jobseeker's Allowance if he or she paid or was credited with class 1 National Insurance (NI) contributions in the preceding 2 tax years. The other is based on a family Means test, which includes personal and/or family income and savings, whichever is relevant given an individual's circumstances (single/married/cohabiting). Unlike the Contributions-based JSA, this Means-tested JSA can be granted for an indefinite period. This is called 'Income-based Jobseeker's Allowance'. Thus, type one requires that the individual has paid enough national Insurance on income and the second requires that current household income and savings are below a certain threshold.

Independent of the type of JSA, the level of unemployment benefits are the same and do not depend on the pre-unemployment wage. Since April 2008 the weekly level has been set at £47.95 for individuals aged 16 - 24 and £60.50 for those aged 25 or over. However, the level of benefits can increase, depending on household size. This implies that the JSA wage replacement rate is in general very low for previous high earners, an important difference when compared to many other European countries with more generous income-related benefits.

The receipt of other benefits may make an individual ineligible for JSA. Quitting a job voluntarily may lead to a benefit sanction of up to 26 weeks. In order to remain eligible for entitlements, the unemployed must visit the Jobcentre at least once every two weeks, and provide evidence that they have been actively looking for employment[5] and are ready to work. In the UK, eligible individuals must be normally between 18-65 years and have a jobseekers agreement with the jobcentre. For more details on the institutional setup see Jobcentre (2008).

[5] There are various ways of providing this evidence, as highlighted in the JSA brochure: "You should do at least 3 things every week. This could include writing a CV or speaking to employers (Jobcentre, 2008, pg.10)."

The Jobcentre Plus operates across 8 major regions which cover the whole of Britain (excluding Northern Ireland). It maintains over 1,000 offices, amongst which includes back office branches and call centres. The administrative and institutional structure is generally the same across the country, whilst intermittent internal restructuring and the introduction of nationwide policies may lead to temporary regional disparities. For example, the New Deal Programme was first introduced in pilot regions before being implemented in the rest of the country. However, we are not aware of permanent regional differences in the institutional setup of the programme.

Jobcentres administer the main active labour market policy programme: the New Deal Programme. This is a programme that gives people on benefits additional support, including training and preparing for work, in order to improve their employment prospects. Whilst eligibility for this programme is the same nationwide, there are considerable regional and local disparities in the share of eligible individuals starting the scheme. The New Deal for Young People programme is compulsory for JSA claimants aged 18-24 after 6 months. Investigating the impact of RESTART, Dolton and O'Neill (1996) highlight that self-selection on to the scheme may be an issue due to perceived re-employment prospects. Tighter monitoring restrictions, as well as poor re-employment prospects, make exits to alternative labour market states, e.g. Income Support, a more attractive proposition.

3 Data

Our analysis is mainly based on individual-level administrative data from the United Kingdom which we merge with several regional labour market indicators.

Individual data. We use the JUVOS (Joint Unemployment and Vacancies Operating System) cohort, which is a randomised 5% sample of all benefit claimants. This data is organised into daily spells relating to individual unemployment benefit claim periods. See Ward and Bird (1995) for a general description of the JUVOS. Our version covers the period 1982 to June 2007. The data is available as a scientific use file from the Office for National Statistics. We restrict our sample to spells starting from the 1st of January 1999 to 31st of December 2005.

It is well known that the claimant count-based and ILO-defined unemployment measures diverge, notably following the 1996 introduction of JSA. Wilke (2009) proposes ways to deal with the limitation in the JUVOS of not being able to identify the true length of unemployment periods, as well as the gaps in individuals' employment histories due to lack of matched administrative data. In Wilke (2009)'s study, the author suggests several implementations of unemployment duration in the JUVOS as, in many cases, single claim spells will not coincide with the true duration of unemployment. By using the reason for leaving markers at the end of claim periods, it develops bounds for the true level of unemployment as well as enabling the use of a competing risks approach with respect to destination state.

In this paper we consider durations of continuous receipt of unemployment benefits (Concept 1 of Wilke, 2009). This is a lower bound for the true unemployment duration and should not contain periods other than claimant unemployment. By restricting the sample to spells with sufficient foregoing employment duration, this should ensure that, in most cases, the start of an unemployment spell equals the start of a claimant spell. Since JSA is means tested after six months, we face the problem of attrition in particular for individuals with an employed spouse[6]. We restrict our sample to single individuals only since, in the case of singles, the benefit duration bears a closer resemblance to the true unemployment duration[7]. Our duration analysis is therefore not an analysis of ILO unemployment duration. However, our sample of claim periods should be comparable to unemployment durations, as individuals are likely to be entitled for JSA for the duration of unemployment. We right censor observations with exits to states other than employment and at the end of the observation period. Based on this definition, we have 39.1% right censored observations in the final sample.

As with most administrative individual data, the JUVOS is handicapped by a limited covariate set. Information contained in the JUVOS includes: start & end date of claims, gender, age and marital/cohabiting status. Following Wilke (2009), we refer to the 2000 Standard Occupational Classification (SOC2000) aggregating the 4-digit occupational codes in the JUVOS to the 1-digit level. We then further group these into 5 representative categories: elementary. manufacturing, trade/services, technical and senior/professional (see Wilke, 2009). van den Berg and van Ours (1994) find that the season of entry onto the unemployment register impacts significantly on exit probabilities, leading us to control for seasonal influences by including quarterly fixed effects. Calendar time effects are indirectly controlled for via year dummies, which should also account for business cycle effects over the period of observation (Lüdemann et al., 2006).

[6]Due to the family Means test

[7]Using German data, Arntz and Wilke (2009) show that empirical results for single males and females are quite similar while married males and females possess different result patterns. This is likely due to the well documented labour market attachment differences between married males and females (Kalwij, 2004).

Table 2: Work History Variables

Variable Name	Description
Active Labour Market Participation	Individual engaged in at least one past Active Labour Market Programme participation.
Long-Term Unemployment	Individual experienced at least one period of long-term unemployment in the past (>365 days).
Incapacity Benefits	Individual claimed incapacity benefits on at least one occasion, on exiting claimant unemployment in the past.
Income Support	Individual claimed income support on at least one occasion, on exiting claimant unemployment in the past.

Using linked German administrative data, Arntz and Wilke (2009) found a strong influence of individual heterogeneity - notably work history - on unemployment durations, whereas regional factors were found to be less important. The importance of work history is also highlighted in the individual-level study by Lüdemann et al. (2006). Collier's (2005) results suggest the opposite, using unique individual-level survey data for the English county of Kent, finding individual characteristics to be less important than regional macroeconomic environment. Taking this into account, we control for individual work history, using the measures defined in Table 2. Age and gender are also included in order to control for socio demographic factors.

Our final sample consists of about 187,000 spells, a descriptive summary of which can be found in the appendix, Table 6.

Regional data. Öberg and Oscarsson (1979) observe that individuals with similar labour market characteristics tend to gravitate to specific regions. This suggests that the evolution of compositional changes in relative regional demographics may be of interest in determining what is influencing individuals' unemployment experiences in these geographies.

Regional-level data was sourced from the quarterly Local Area Labour Force Survey, available from UK Data Archive. Regional data was also sourced from other providers, however missing values limited the final covariate set (e.g. NOMIS censors all observations with values less than 500, implying that small area data is likely to be affected). Continuous variables at the regional level were standardised across regions by month, the shortest interval in the regional dataset. We link the regional-level to the individual-level data by claimant spell start month, since we lack continuous daily data on regional characteristics[8]. The final data set consisted of 60 possible covariates at the individual and 160 at the regional level of aggregation.

In order to make the model tractable, we implemented cluster analysis techniques to class regional variables into representative groups. The Clustering of Variables Around Latent Variables (CLV) routine by Vigneau and Qannari (2003) was used. This is a two-stage routine which implements hierarchical clustering analysis followed by a partitioning algorithm, thus capturing the benefits of both approaches. This method clusters highly correlated variables together, regardless of the direction of this correlation. This allows us to select a variable to represent the information captured by the other neighbouring covariates. Given data availability issues, certain variables would be more attractive than others. This approach implies that this selection is not arbitrary and based on economic and statistical criteria.

Following Arntz and Wilke (2009), regional data was sourced in order to characterise the local environment in which individuals reside. Regional variables were clustered into 5 representative groups, capturing the relevant supply, demand, as well as structural, social and institutional factors of interest.

Supply & Demand: Local ILO unemployment rates were used in order to indirectly capture regional 'labour market tightness'. An alternative proxy for 'labour market tightness' is the unemployment/vacancy ratio. However, this indicator is plagued by data quality issues due to significant changes to Jobcentre Plus procedures for handling vacancies in 2001[9]. The retrospective average 12 quarter change in ILO unemployment is included as a proxy for the medium-term evolution of local supply and demand imbalance.

Local economic performance: Regional Gross Domestic Product (GDP) and change in GDP proxy for the level and change in economic activity in a region. In addition, the rate of new business startups is a further indicator of economic activity. Prosperous areas are likely to have high and positive values for these measures, respectively, implying greater job finding prospects. Since GDP data is unavailable at the aggregation level of interest, we use unadjusted quarterly Gross Value Added (GVA) as a proxy. This workplace-based measure, allocated to the region in which commuters live, is reported at basic prices[10]. The retrospective 3 year average change in

[8]For a detailed exposition of the data preparation steps, see Ball (2009). For details on how this link was created see Appendix A2.

[9] The effect being that vacancy statistics are not comparable over time (Bentley, 2005).

[10]Deflated for changes in prices over time and across regions (Office for National Statistics, 2007).

GDP per head is used as a medium-term measure of this phenomenon, as annual changes are likely to be picking up the effects of transitory shocks at the national level. The rate of business startups is proxied by the number of new businesses registering for VAT each year as a proportion of the resident population. Due to their size, this indicator will not include sole proprietors. However, since the majority of VAT-registered businesses employ less than 50 employees, this indicator is capturing small business activity. Less than half of UK businesses are registered for VAT (NOMIS, 2009).

Social Structure: We define Skill Intensity as the proportion of all employees aged 16 & over working in the following occupational classifications: Managers & Senior Officials; Professionals; Associated Professionals & Technical; Admin. & Secretarial; & Skilled Trades [11]. This measure proved to be highly correlated with educational attainment rates, constructed from the same source. Education attainment and income levels are assumed to be linked through productivity by Human Capital Theory (Becker, 1964). Since educational attainment and skill intensity are highly correlated, it would then be expected that individuals living in skill intensive areas would experience higher job offer arrival rates. Their unemployment spells would thus be expected to be shorter[12]. However, the impact of skill level on unemployment duration is likely to be endogenous due to the fact that higher job offer arrival rates are likely to push up reservation wages. The effect of this would lead affected individuals to be more selective about the job offers they accept and in turn lengthening unemployment periods (Mortensen, 1970). Furthermore, the institutional context (section 2) and monitoring restrictions that the UK unemployment benefits system places on job offer acceptance/rejection, suggest that the impact of this covariate is an empirical question.

Institutional Organization: We tried to collect any kind of information about the internal structure of the jobcentre branches but our requests were rejected by the DWP. Given the shortage of information and given the nationwide identical entitlements for participation in the New Deal Programme, it is therefore difficult to control for the institutional organization. However, we have constructed one indicator, the New Deal for Young People Starters as a proportion of the eligible claimant count. In our analysis this variable is interacted with individuals being aged 18-24. Note that we do not include the base effect of this variable due to multicollinearity. A

[11]It is acknowledged that this measure is likely to suffer from measurement error due to heterogeneity of skill-intensities within detailed occupational categories.

[12]The implicit assumption is one of *perfect information*, that an individual's education level an accurate signal of true productivity and does not pick up unobserved heterogeneity, *viz.* Signal Theory (Spence, 1973; Silles, 2008). In support of the assumed link between income level and educational attainment, Silles (2008) finds that higher levels of education are always associated with higher earnings in the UK, however whether Human Capital or Search Theory can explain this as a causal relationship is a debatable given the influence of confounding factors like family background (Angrist and Krueger, 1999).

negative -shortening- effect of this variable would indicate that local jobcentres are more likely to assign eligible individuals to the New Deal Programme if the local labour market offers better re-employment opportunities.

Structural indicators are included in order to characterise the type of region.

Unemployment Dynamics: Regions with high levels of seasonal employment, proxied by the 'flow of unemployed as a proportion of the resident population', are more likely to be characterised by longer unemployment spells as the sample median unemployment duration is around two months.

Urban/Rural indicator: Two versions of this variable were sourced. One from the National Statistics Postcode Directory (NSPD) and one from the Department of Environment, Food & Rural Affairs (DEFRA). For England & Wales, the NSPD indicator, a population density-based measure, is derived using the 21st of July 2004 release of the National Statistics Rural & Urban Classification of Output Areas (NSPD, 2007). This Output Area-based indicator is not valid for higher levels of aggregation which may include a mixture of rural and urban output areas based on the definitions used. For Scotland, areas are defined as rural if they have less than 3,000 inhabitants (NSPD, 2007). The DEFRA classification is based on local authorities, but is only available for England[13]. The correlation between these two measures is low, .56 by our calculations. Given the superiority of the DEFRA classification, where it was available it was implemented, and where not the NSPD definition was used, implying that this indicator involves some measurement error for Scotland and Wales.

Accessibility: Exploiting the rich data available in the NSPD, the sparsity of the surrounding area was used in order to define whether a local authority was accessible or remote in the case of England and Wales. Driving distance to the nearest large settlement is used as a proxy in the case of Scotland[14]. One would expect that, on average, individuals' labour market outcomes would be better in regions that are urban and/or near large urban conurbations due to the positive job-prospect spill-overs as a result of higher levels of economic activity.

University Present: Information on Higher Education institution location was sourced from the Higher Education Statistics Authority (HESA). As a policy relevant variable, one would ex-

[13]See DEFRA, 2007.

[14]Since this indicator is output area-based, this may be subject to some error. We assume that this error is small, given the lack of alternative local authority-based measures.

pect that the presence of higher education institutions would be a force for improved employment prospects for the local population, given the support services needed to run such an institution as well as the influx of young consumers into the local market. However, as pointed out by Arntz and Wilke (2009), the increased availability of a young flexible workforce willing to work at minimum wage rates may impact negatively on the labour market participation on a section of the local population. The overall impact of this indicator is likely to be an empirical question, given these confounding factors.

Table 3: Regional Indicators.

No.	Group	Indicator	Source	Mean	SD
I	**Demand/supply**	• Quarterly *ILO Unemployment rate* (%)	A	.048	.021
		• Average retrospective 12 quarter *change in ILO Unemployment rate* (%)	A	.047	.023
II	**Economic Performance**	• Annual *GDP per head (GDPPH)* (£)	B	14971.3	7604.48
		• Average retrospective 3 year *change in GDP per head (GDPPH)* (%)	B	.076	.139
		• New Small Business Startups/ Resident Population	E	.003	.001
III	**Social Structure**	• Quarterly *Skill Intensity*	A	.622	.072
IV	**Institutional Organisation**	• Monthly *18-24 New Deal Starters* (% of eligible claimant count)	A/E	.024	.051
V	**Structural Indicators**	• Accessibility	C	.933	.249
		Type of region: (ref. rural)			
		• Urban	C/D	.634	.212
		• University Present	F	.135	.342
		• Flow of U/ Resident Population	E	.007	.003
Number of obs = 187,032					

A: Local Area Labour Force Survey; B: Office of National Statistics; C: National Statistics Postcode Directory;
D: DEFRA (Department of Environment, Food and Rural Affairs; E: NOMIS; F: HESA (Higher Education Statistics Agency)

The linked data set matching the individual- and regional-level data to the UK geography is conditioned on the start of claimant spells. In order to match the continuous individual-level data to the regional information, individual spells were matched to the regional information pertaining to the month in which they started (see also appendix A2).

The final data set contains the information of 963 unemployment benefit office (UBO) locations (full postcodes and postcode districts). This is then mapped to the existing data via the NSPD. Given the self-reported nature of the JUVOS postcode information, data quality issues were present with postcode information missing or wrongly imputed at times. In order to maintain some regional variation we only replaced this self-reported variable with the UBO postcode district when this variable was missing and no information could be obtained from previous spells (implemented in 2% of cases). If the postcode information was missing, the initial strategy was to replace this with the postcode reported in a previous spell if this existed, i.e. assuming that the individual did not move location between the spells. This was implemented in 2.8% of cases.

We omit Northern Ireland from proceedings, due to lack of coverage for some major regional indicators of interest at all levels. Our analysis thus focusses on Great Britain. The City of London and Isles of Scilly local authorities are dropped from the analysis, as data for these geographies is systematically missing at the aggregation level of interest (local authority level). However, in the case of randomly missing values we impute values for the variables of interest given the number of missings is so low for the selected variables. For each variable affected, the imputation method was to replace the variable by the data in the preceding period.

Due to the creation of 46 unitary authorities[15] over the period 1995 to 1998 in the regional data, including 13 extra units, we restricted our observation period to after 1998. This was due to a restructuring of local governments over the period, from a one-tier to two-tier (lower level) system in some areas. The resulting geography is a mixture of Local Authority Districts, Unitary Authorities and Metropolitan Districts. Restricting ourselves to the 1999 to 2005 period also avoids a concordance issue between the 1990 Standard Occupational Classification (SOC90) and the 2000 update (SOC2000), as the Local Area Labour Force Survey is available according to the SOC2000 methodology from the first quarter of 1999 (Beerten et al., 2001).

Due to data limitations, we are unable to distinguish between New Deal participants on govern-

[15]"Single-tier administrations with responsibility for all areas of local government (Office for National Statistics, 2004)"

ment supported training initiatives and those partaking in subsidised work-placements. Individual-level studies on the Swedish and the German labour markets highlight fundamental differences in the re-employment probabilities of these two population sub-groups (Adda et al., 2007; Arntz and Wilke, 2009). Using individual-level Slovakian administrative data, van Ours (2004) found a significant locking-in effect of government subsidized jobs. Given the regional context of our study, this suggests that where these jobs occur may be of importance.

4 Econometric Model

We analyse the determinants of unemployment duration by means of censored quantile regression and the Cox proportional hazard model. Censored quantile regression is recently emerging as an attractive and powerful alternative to proportional hazard models (see for example Koenker and Geling, 2001). The linear quantile regression model, introduced by Koenker and Bassett (1978) models the conditional quantile function of the dependent variable as a linear functional of the regressors x_i, where x_i is $k \times 1$ with $x_{1i} = 1$ for all $i = 1, \ldots, N$. Let the dependent variable lny_i be the logarithm of the ith duration of unemployment y_i. Then the θth conditional quantile of the dependent variable given x is given by

$$\begin{aligned} Quant_\theta(lny_i|x_i) &= x_i'\beta^\theta \quad \text{or} \\ Quant_\theta(y_i|x_i) &= exp(x_i'\beta^\theta) \end{aligned}$$

where β^θ is a $k \times 1$ vector of unknown coefficients. Note that these coefficients are allowed to vary over the quantile $\theta \in (0, 1)$. This means that the framework is flexible enough to allow for different effects of the regressors at different quantiles of the conditional distribution of unemployment duration. In particular, as the sign of the coefficients can change, a regressor can have a shortening effect for a lower quantile θ_1 ($\beta_j^{\theta_1} < 0$) and a prolonging effect for a higher quantile θ_2 ($\beta_j^{\theta_2} > 0$) with $\theta_1 < \theta_2$. Since our sample of unemployment duration is partly right-censored, we apply censored quantile regression. Our sample is (lny_i, x_i, yc_i), $i = 1, \ldots, N$, where $yc_i = lny_i$ if the unemployment duration is not censored and $yc_i = \infty$ when it is right censored. We apply the censored quantile regression estimator of Powell (1984) and Powell (1986) and obtain $\hat{\beta}^\theta$ by minimising the following distance function

$$\frac{1}{N}\sum_{i=1}^{N}\rho_\theta(lny_i - min(x_i'\beta^\theta, yc_i)) \tag{1}$$

with,

$$\rho_\theta(u) = \begin{cases} \theta \cdot |u| & \text{for } u \geq 0 \\ (1-\theta)\cdot|u| & \text{for } u < 0. \end{cases} \tag{2}$$

For more details on censored quantile regression see the recent survey by Koenker (2008). We use the censored LAD procedure of TSP 5.0 to estimate the unknown coefficients at three quantiles $\theta = 0.1, 0.5$ and 0.7. We bootstrap the full sample 100 times to approximate the distribution of the estimator and therefore to obtain inference statistics. [16]

The Cox proportional hazard model is based on the idea that the conditional hazard rate is proportional for different values of the regressors x. For the ith observation let $\lambda_i(y|x) = f_i(y)/P(Y_i \geq y) = exp(x_i'\tilde{\beta})\lambda_0(y)$ be the hazard rate and $f_i(y)$ the conditional density of Y_i given x_i. λ_0 is the so called baseline hazard which is nonparametric. The Cox model gains its popularity from the fact that it is relatively simple to estimate.

We estimate the Cox model by using the implementation in STATA 10 and report hazard ratios, i.e. the proportionate change in the hazard rate relative to a reference group with $x_i = 0$ rather than the estimated coefficients itself. The Cox model has also several drawbacks. It ignores individual specific error terms, which can lead to a systematic bias of estimated coefficients even if the error is uncorrelated with the regressors. Moreover, the estimated baseline hazard is usually downward biased in the presence of unobserved heterogeneity in particular for longer durations.

While the marginal effect of a regressor on the conditional distribution of unemployment duration can vary over the quantiles, the Cox model implies a unique sign of this effect (see Lüdemann et al., 2006). Therefore, the censored quantile regression model offers an attractive alternative as it is robust with respect to the unobserved heterogeneity and it does not restrict the effect of the regressors over the distribution of unemployment duration. Note that there is no one-to-one correspondence between the quantile regression model and the Cox proportional hazard model, the coefficients $\tilde{\beta}$ and β are not the same. We focus our comparison of estimation results therefore

[16] We do not bootstrap more often because of the extensive computational effort and we do not apply the bootstrap method of Bilias et al. (2000) as the degree of censoring in our data is rather high. Since our sample consists of dummy variables and standardised continuous variables only, we do not report marginal effects as interpretation is straightforward in this case.

on the sign and relative importance of the regressors and whether we can observe different signs of the estimated quantile regression coefficients for different quantiles.

In order to provide a more complete insight in the effects of various regressors on unemployment duration, we also investigate conditional hazard rates. Since the nonparametric baseline hazard of the Cox model is likely to be biased, we estimate nonparametric conditional hazard rates based on quantile regression estimates. We apply the resampling procedure of Fitzenberger and Wilke (2006) for right censored duration data which is a modification of the approach by Machado et al. (2006) (henceforth denoted as MPG). The main idea of the MPG is to simulate data based on the estimated quantile regression coefficients given the regressors and to estimate the conditional density and the conditional distribution function of the dependent variable directly from the simulated data.

In detail the procedure is as follows:

1. Generate M independent random draws $\theta_m, m = 1, ..., M$ from a uniform distribution on (θ_l, θ_u), i.e. extreme quantiles with $\theta < \theta_l$ or $\theta > \theta_u$ are not considered here. θ_l and θ_u are chosen in light of the type and the degree of censoring in the data. Additional concerns relate to the fact that quantile regression estimates at extreme quantiles are typically statistically less reliable, and that duration data might exhibit a mass point at zero or other extreme values. The benchmark case with the entire distribution is given by $\theta_l = 0$ and $\theta_u = 1$.[17]

2. For each θ_m, estimate the censored regression model obtaining M vectors β^{θ_m}.

3. For a given value of the covariates x_0, the sample of size M with the simulated durations is obtained as,

$$Y_m^* \equiv \hat{q}_{\theta_m}(Y|x_0) = exp(x_0'\beta^{\theta_m}) \quad \text{with} \quad m = 1, ..., M \; .$$

4. Based on the sample $\{Y_m^*, m = 1, ..., M\}$, estimate the conditional density $f^*(y|x_0)$ and the conditional distribution function $F^*(y|x_0)$.

5. The hazard rate conditional on x_0 and conditional on the durations drawn in the interval (θ_l, θ_u)[18] is estimated by

[17] In our application, $\theta_l = 0.05$ and $\theta_u = 0.7$. Random numbers are then drawn from a discrete uniform distribution which has the quantile grid points as support points. This increases computation time significantly at the cost of small approximation errors.

[18] Simulating the full distribution ($\theta_l = 0$ and $\theta_u = 1$), it follows by definition: $\hat{\lambda}_0(y) = f^*(y|x_0)/[1 - F^*(y|x_0)]$.

$$\hat{\lambda}_0(y) = \frac{(\theta_u - \theta_l) f^*(y|x_0)}{1 - \theta_l - (\theta_u - \theta_l) F^*(y|x_0)} \quad .$$

MPG uses a kernel estimator for the conditional density

$$f^*(y|x_0) = \frac{1}{M\,b} \sum_{m=1}^{M} K\left(\frac{y - Y_i^*}{b}\right)$$

where b is the bandwidth and $K(.)$ the kernel function. Based on this density estimate, the distribution function estimator is

$$F^*(y|x_0) = \frac{1}{M} \sum_{m=1}^{M} \mathcal{K}\left(\frac{y - Y_i^*}{b}\right) \quad \text{with} \quad \mathcal{K}(u) = \int_a^y K(v)\, dv \; .$$

We follow Fitzenberger and Wilke (2006) and use a kernel density estimator based on log durations. The estimates for density and distribution function for the duration itself are easily derived from the density estimates for log duration by applying an appropriate transformation.

5 Empirical Results

Table 4 reports estimation results for the duration models as described in the previous section. It shows the estimated coefficients of the censored quantile regression model and the estimated hazard ratios for the Cox model. Cox model A is a model which includes the reported variables only while model B also contains dummy variables for the 128 NUTS3 regions in Great Britain. Estimated conditional hazard rates based on the resampling procedure are presented in Figures 2 and 3.

Table 4: Estimated coefficients of the censored quantile regression model and estimated hazard ratios of the Cox proportional hazard model.

	Censored Quantile Regression			Cox Model	
	Quantile 0.1	Quantile 0.5	Quantile 0.7	A	B
Intercept	2.315***	4.088***	4.755***		
Socio-demographics					
age< 25	0.244***	0.148***	0.073***	0.894***	0.894***
age> 56	-0.094	0.086*	0.140***	0.85***	0.847***
female	-0.042***	-0.048***	-0.048***	1.025***	1.033***
Occupation(ref:Elementary)					
Manufacturing	-0.148***	-0.190***	-0.167***	1.165***	1.157***
Trade, services	-0.061***	-0.178***	-0.190***	1.157***	1.168***
Technical	0.016	-0.112***	-0.146***	1.108***	1.128***
Senior, professional	-0.010***	-0.280***	-0.305***	1.267***	1.288***
Unknown	-0.177***	-0.333***	-0.347***	1.341***	1.348***
Work History variables					
Active Labour Market Programme Participation	0.071***	0.261***	0.276***	0.8***	0.802***
Long-Term Unemployment	0.310***	0.501***	0.518***	0.646***	0.654***
Incapacity Benefits	-0.051**	0.009	0.008	0.972**	0.955***
Income Support	0.024	0.070	0.067***	0.933***	0.944**
Calendar time(ref: 1999q1)					
y2000	0.010	-0.009	-0.021**	1.017*	1.018*
y2001	0.072***	-0.030***	-0.054***	1.022**	1.029***
y2002	0.120***	0.016	-0.010	0.974**	0.985
y2003	0.209***	0.070***	0.044***	0.926***	0.934***
y2004	0.230***	0.119***	0.065***	0.878***	0.889***
y2005	0.438***	0.312***	0.273***	0.746***	0.754***
q2	-0.029*	-0.015	-0.016	0.991	0.991
q3	-0.026	-0.041***	0.028**	0.975***	0.972***
q4	0.118***	0.193***	0.111***	0.908***	0.902***

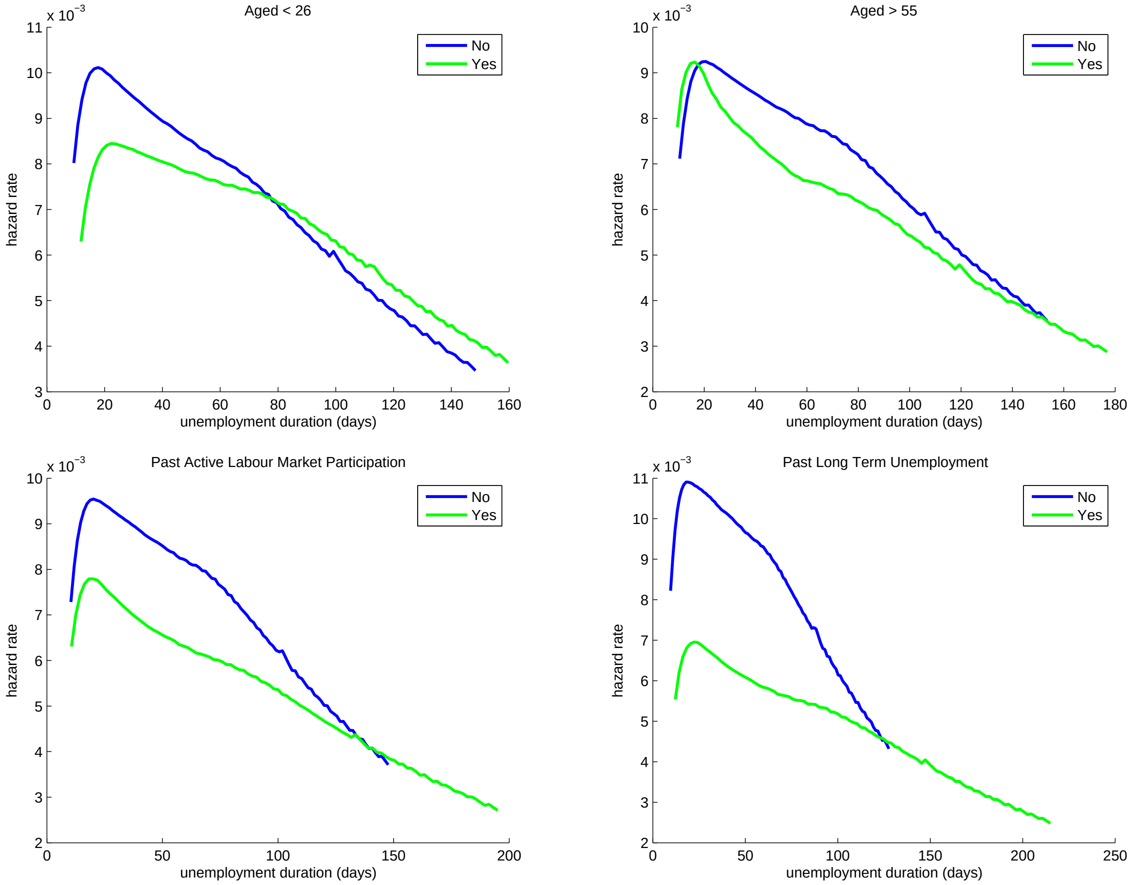
Continued on next page

Table 4 – continued from previous page

	Censored Quantile Regression			Cox Model	
	Quantile 0.1	Quantile 0.5	Quantile 0.7	A	B
Regional variables					
Accessible	0.110***	0.088***	0.082***	0.902***	0.985
Urban	-0.026	-0.013	-0.034	1.023	0.939*
University Present	0.023	0.049***	0.061***	0.953***	0.917***
Skill Intensity	0.016**	0.024***	0.028***	0.973***	1
GDPPH	0.013**	0.025***	0.022***	0.978***	0.972
ILO unemployment rate	0.056***	0.072***	0.068***	0.938***	0.98***
Change in GDPPH	0.010*	0.015***	0.017***	0.987***	1.003
Change in ILO unemployment rate	-0.018**	-0.035***	-0.039***	1.032***	0.999
18-24 New Deal Starters	-0.012**	-0.029***	-0.028***	1.023***	1.012***
Flow of Unemployed/ Resident Population	0.014*	0.023***	0.038***	0.98***	0.975***
New Small Business Startups/ Resident Population	0.098***	0.091***	0.081***	0.92***	0.98***
NUTS3 fixed effects					✓
Number of obs = 187,032					
Significance levels: ***: 1% **: 5% *: 10%					
Note: for regional dummy results see Figure 1(Cox model B only)					

In what follows we discuss and compare the estimation results in more detail. When we compare the Cox model estimates with the Cox estimates of Wilke (2009), we observe that they are similar. The presence of region dummies in the Cox model (B) does not change estimates much but several regional variables become insignificant. Many region dummies in this model are also insignificant, suggesting that the regional variables capture important parts of the regional variation in the data. This is further supported by the significance of the region dummies in a model without regional controls. These observations suggest that our regional controls are able to capture regional heterogeneity sufficiently well. This result is robust to the inclusion of Travel-To-Work Area fixed effects, as well as to changes in specification (Exponential, Weibull, Gompertz).

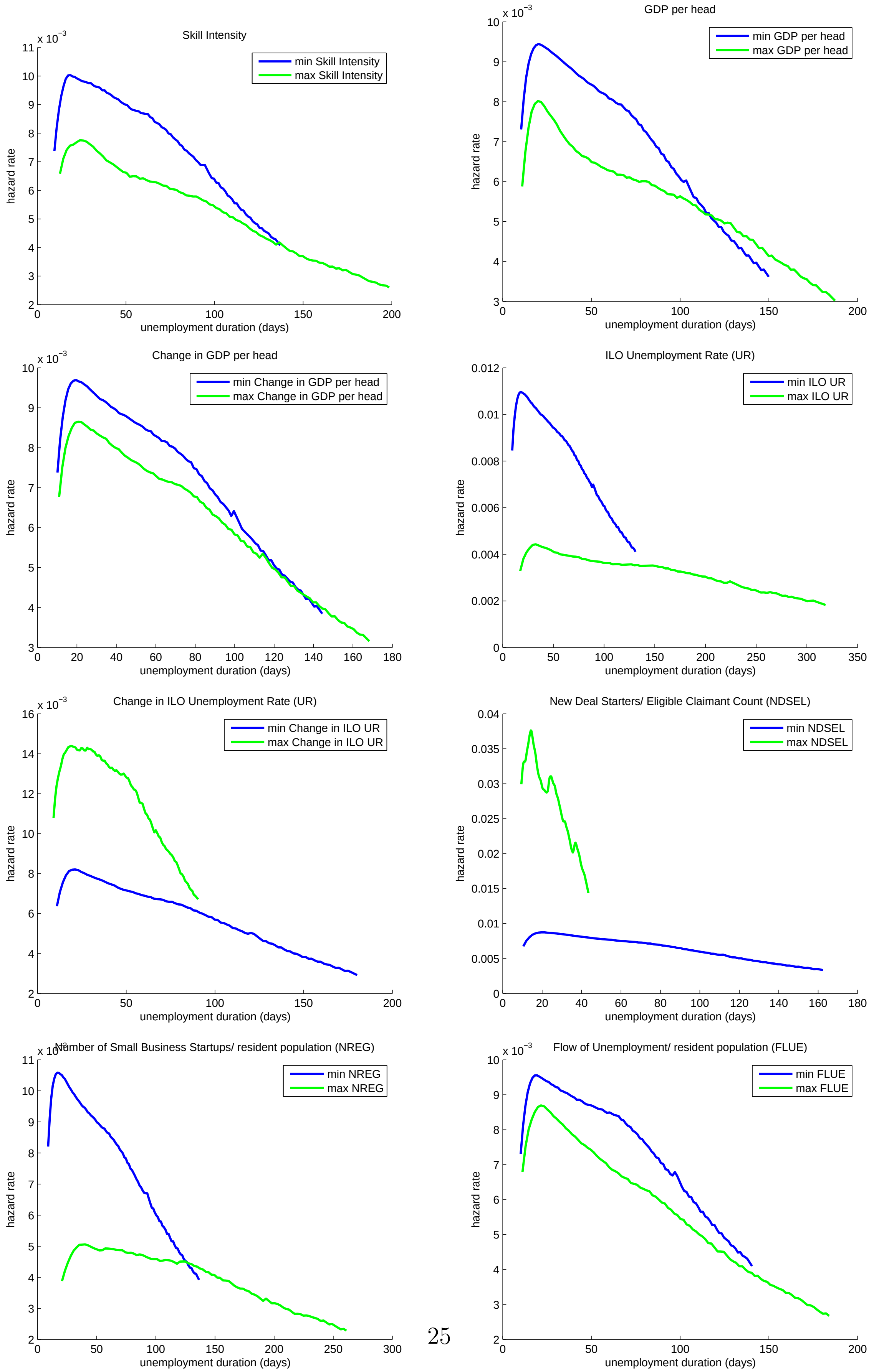
Figure 2: Estimated conditional hazard rates: change from 0 (blue line) to 1 (green line) in one individual level variable; sample means of all other variables.



While several quantile regression coefficient change their sign over the quantiles, we observe a significant change of the sign for only two variables (y2001, q3). These cases imply an immediate violation of the proportionality assumption. Moreover, hazard rates may be disproportionate in absence of sign changes of quantile regression coefficients and it is therefore inevitable to look directly at the estimated conditional hazard rates to obtain a clearer picture.

Figures 2 and 3 present a selection of the estimated conditional hazard rates, where we do not display results for the calender time and if the effect of a regressor is very small. The support of the estimated hazards is limited to a certain interval as we have only estimated the quantile regression model for $\theta \in [0.05, 0.7]$. The Figures suggest that the covariate effect is mainly limited to shorter durations of up to about 150 days. Moreover, they provide evidence that conditional hazard rates often appear disproportionate (see for example aged < 26). Unfortunately, since higher moments of the hazard rate estimator are unknown, we cannot test for this type of shape regularity. However, given the very large number of observations we believe that it is likely that

Figure 3: Estimated conditional hazard rates: changes from sample min (blue line) to sample max (green line) in one regional variable ; sample means of all other variables.



some of the non-proportionalities cannot be rejected. Due to the more restrictive nature of the Cox model, we mainly base the following discussion of estimation results on the quantile regression estimates. In general the results in Table 4 suggest that if a variable has an economically and statistically significant effect, then this will be implied by the two models. If there is a change of sign in the quantile regression model, then the Cox estimator is more likely to produce the effect at higher quantiles, thus the conflicting effect appears more likely for shorter durations.

Individual variables and calender time The following attributes have a considerably prolonging influence on the length of JSA claim periods: aged <26; past participation in Active Labour Market Programmes (ALMP); and past experience of long term unemployment. Moreover, we observe a clear time pattern with longer durations in later years. Being aged <26 and being aged >56 display a reverse trend across the quantiles, relative to being prime-aged. Relative to the least skill-intensive occupations, being employed in all other occupational groups significantly shortens claimants' spells. Furthermore, the aforementioned time pattern is reversed relative to the base line category. As some individual variables have a stronger association with the dependent variable, relative to the other controls, the direct implication is that the individual-level seems to be more important than the regional-level of aggregation. Estimated effects of the individual level coefficients are generally similar to the estimates of Wilke (2009) and for this reason we omit here a more detailed discussion.

Regional variables Although regional labour market conditions generally possess a significant association with the length of claim periods, the size of these effects is often considerably smaller than for the individual level variables. This pattern is not unique to Britain as the same observation was made with data from Germany after controlling for institutional factors (Arntz and Wilke, 2009). This contrasts the findings of Collier (2005) which suggest that regional labour market conditions are more important. For the set of regional variables, we do not observe any change of sign of the quantile regression coefficients over the quantiles.

Better accessibility of a region increases the length of JSA claim periods, in particular for very short durations. This is roughly compatible with the findings of Arntz and Wilke (2009) for Germany, who observe that a longer driving time to a higher level city increases the job finding probability for singles. However, the sign of this effect is inconsistent with our previous hypothesis. Urban regions are associated with relatively shorter claimant periods, a result which concords with our expectations from economic theory and is consistent with Arntz and Wilke (2009). Although consistent with out priors, this result is insignificant at conventional levels. Given that urban regions are also accessible, the accessibility indicator is capturing the effect of being accessible conditional on being urban. As in the case of Germany, over the time period of observation, the presence of a university lengthens JSA claim periods whilst the relationship is not significant for shorter durations. The presence of more skill-intensive jobs increases the length of claim periods. This suggests that a better social environment is related with poorer employment prospects. This finding contrasts the results of Arntz and Wilke (2009) and the interpretation of the effect is unclear and may be affected by endogeneity as the individual level variable suggests the contrary.

Higher local GDP per head has a positive association with the length of claim periods although this effect is small in economic terms. Although surprising, this result pattern is also compatible with the observation of Arntz and Wilke (2009) for Germany. Similar to the results of Petrongolo (2001), our analysis suggests that a higher local unemployment rate is related with longer claim periods. The effect increases over the quantiles and it is one of the most important regional variables. In addition to the unemployment rate, Arntz and Wilke (2009) also control for the share of long term unemployed and their results suggest that this indicator significantly increases spell lengths. For the reasons mentioned earlier, we have not included the share of long term unemployed in our final model. However, other model specifications suggest that the indicator has a comparable effect when used instead of the unemployment rate. It is not confirmed by our results that emerging regions (in terms of increase in GDP per head and decrease in unemployment rate) improve JSA claimants' job finding probabilities. However, the estimated effects are small in terms of economic significance, notably in the lowest quantiles. Granted, the impact of higher unemployment rates may suggest a 'stigma' effect of living in high unemployment regions that increases with the duration of unemployment.

The share of New Deal programme starters amongst the eligible claimant count (18 - 24) has a negative association with the length of JSA claimant durations. Although this unlikely to be causal, it suggests that assignment activity in local jobcentres may be related to regional labour market outcomes and not fully random.

In line with our expectations we find that the rate of unemployment flows - which proxies Seasonal Unemployment - has a positive impact on JSA claimant durations. This effect increases across the quantiles.

Arntz and Wilke (2009) find that the rate of new business startups in an area - which proxies local 'business activity' - has a positive, and significant, impact on the prospects of low-wage earners being re-employed in the local area. The effect on high-wage earners is found to be insignificant, which may be due to higher levels of job mobility. Due to lack of earnings information, we are unable to make this distinction. However, this indicator is one of the most economically significant amongst the regional variables. The estimates suggest that higher levels of 'business activity', relative to the resident population, have a lengthening effect on claimant spells which is most notable in the bottom quantile. This estimate is difficult to interpret, given our priors. Although most of the estimated effects of the regional variables are rather small in magnitude, being accessible, the local unemployment rate and 'business activity' in a region turn out to be the most important among them. There are strong shifts in the estimated conditional hazards for changes from the sample minimum to the sample maximum in these continuous regional variables. This suggests that extreme regional labour market conditions do have strong effects, although Table 4 suggests that sample effects on the conditional quantiles are mainly limited as they are in response to a shift by one standard deviation.

Table 5: Top and worse performing regions in Great Britain. Results from a Cox regression with individual variables, calender time variables and 128 NUTS3 region dummies.

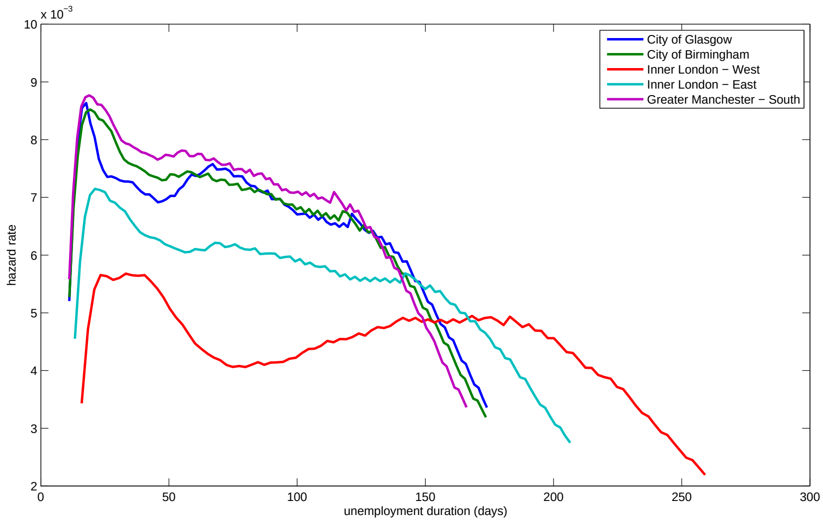
Rank	Top Performer	Worst Performer
1	Eilean Siar (Western Isles)	Inner London - East
2	Caithness/Sutherland/Ross/Cromarty	Inner London - West
3	Shetland Islands	Outer London - West and North West
4	East Cumbria	Birmingham
5	Aberdeen/Aberdeenshire/North East Moray	Berkshire

Comparison of regions As a next step we directly compare selected regions. Table 5 reports a ranking of region dummies obtained by the Cox model with individual variables and region dummies only (e.g. omitting the other regional variables). The region dummies capture both the observable and unobservable region specific effects and thus provide us a simple performance ranking in terms of the length of claim periods by controlling for individual specific characteristics and calendar time. We refer to the Cox estimates as it was technically impossible to obtain censored quantile regression estimates when regional dummies were included in the model.

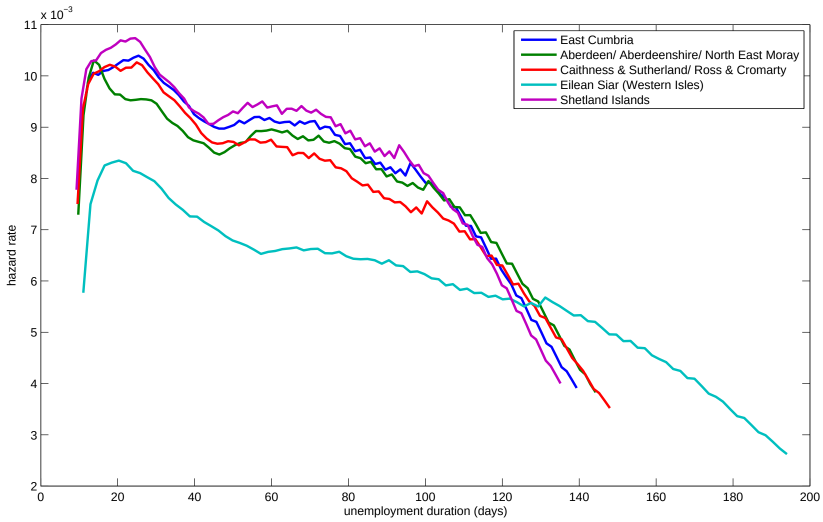
The table suggests that large cities such as London and Birmingham and the London commuter belt have the strongest positive association with the length of claim periods. In contrast, remote regions such as the Western Isles and Shetland Islands are among the regions with the shortest conditional claim duration. It is remarkable that four out of the five top performing regions are located in Northern Scotland. Since the results in the table are based on the simple Cox model's regional dummies, it is unclear whether the ranking is due to the observable regional labour market environment or due to unobservable regional characteristics. It is, however, of interest to explore this in more detail. For this reason we also compute the resampling based conditional hazard rates for the duration of benefit claim periods conditional on being a sample average individual and residing in the region specific labour market environment (sample average for each region). This is done for five NUTS3 regions which overlap with the four major cities in Britain: Inner London (East and West), Birmingham, Manchester and Glasgow.

Figure 4: Estimated conditional hazard rates for several regions.

four major agglomerations (top)



top five performer (bottom)



Apart from Glasgow and Manchester these cities belong to the poorest performers according to the Cox estimates. Moreover, we compute the hazard rates conditional to the top five performers according to Table 5. The resulting estimates are presented in Figure 4. It is apparent that the observed regional labour market conditions in Greater Manchester (South), Glasgow and Birmingham result in rather similar conditional hazards. The observed characteristics for London and in particular for West-London point to a considerably worse observed labour market environment. Hazard rates for the top performers are considerably higher than for the large cities if we ignore the Western Isles. Moreover, the figure suggests that simultaneous changes in several regional variables can lead to considerable shifts in conditional hazard rates. The shape of the hazard rates looks rather disproportionate and differences are mainly relevant in the interval up to 100 days. These results therefore provide evidence that the region specific environment matters much less for longer durations. Our findings therefore suggest that regional policies may fail to improve employment prospects of the long term unemployed.

6 Summary and Conclusion

We create a comprehensive British data set by merging individual claim periods of unemployment benefits with a large set of regional indicators. In our empirical analysis we use this data to investigate the relevance of individual characteristics and local labour market conditions on the length of JSA claim periods. We employ censored quantile regression and apply a resampling method to estimate nonparametric conditional hazard rates.

We find evidence that both individual level variables and the local labour market environment shape the distribution of re-employment times. Although individual level variables turn out to be more important, in particular the local labour demand and supply conditions and structural indicators of a region are also important determinants of the length of claim periods. Our results therefore contrast the results of Collier (2005) who observes regional variables to be more important, while they are often similar to the results of Arntz and Wilke (2009) for Germany. This includes the relative relevance and the sign of the estimated effects. Moreover, we observe that covariate effects are mainly limited to a duration of up to 150 days while they are generally negligible for longer duration. Our results therefore suggests that regional policies are likely to be ineffective for improving employment prospects of long term unemployed. This is an interesting observation which could not be made by employing a proportional hazard model.

From a policy point of view we draw the conclusion that regional labour market conditions and therefore regional policies can affect individual labour market outcomes. Our results, however, do not suggest that these are the principal driving forces for re-employment times. Therefore, regional policies seem more to have a supportive role and they cannot substitute for a lack of individual qualities in the job search process. Surprisingly, we observe that large cities such as London and Birmingham provide worse local labour market conditions than rural and even remote regions such as Northern Scotland. This finding is important as many people likely believe the reverse, although the Government is already targeting problematic neighborhoods in these cities.

Our research also leaves some scope for extensions in some respects. From a methodological point of view, the use of censored quantile regression extends standard econometric techniques in several dimensions. However, it also limits our econometric model in several aspects. First, it cannot deal with time varying covariates and thus we only take into account the information at the start of claim periods. Moreover, we cannot take into account multiple spells in our analysis as this is also still to be developed for censored quantile regression.

From a data point of view, we are unable to fully map individuals' employment biographies (their movements in and out of the labour market, wage changes, etc.) due to the lack of merged administrative individual data. The availability of individual data from additional registers would enable us to perform an extension of our analysis. We do not directly address the issue of commuting as the resident population may not be contributing to the productivity of a region. Job density can be used to indirectly control for commuting. We constructed a job density indicator from the regional data, however this variable turned out to be highly correlated with other indicators used in the analysis and for this reason it was not included in our final model. More comprehensive data with information about the workplace would enable us to directly analyse commuting and even intra regional migration.

References

Adda, J., Costa Dias, M., Meghir, C., and Sianesi, B. (2007). Labour market programmes and labour market outcomes: a study of the Swedish active labour market interventions. Working Paper Series 2007:27, IFAU - Institute for Labour Market Policy Evaluation.

Angrist, J. D. and Krueger, A. B. (1999). Empirical strategies in labor economics. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3 of *Handbook of Labor Economics*, chapter 23, pages 1277–1366. Elsevier.

Arntz, M. and Wilke, R. (2007). An application of cartographic area interpolation to German administrative data. *AStA Advances in Statistical Analysis*, 91(2): 159–180.

Arntz, M. and Wilke, R. (2009). Unemployment Duration in Germany: Individual and Regional Determinants of Local Job Finding, Migration and Subsidized Employment. *Regional Studies*, 43(1): 43–61.

Arulampalam, W. (2001). Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wage. *The Economic Journal*, 111(475): 585–606. Royal Economic Society.

Atkinson, A. and Micklewright, J. (1991). Unemployment Compensation and Labor-Market Transitions - A Critical Review. *Journal of Economic Literature*, 29(4): 1679–1727.

Ball, P. (2009). Construction of a linked postcode district to regional-level dataset for Great Britain. *Nottingham University Working Paper No. 09/09. Available at: http://www.nottingham.ac.uk/economics/research/dp/school_dp/dp.09.09.pdf.*

Becker, G. (1964). *Human Capital.* New York: National Bureau of Economic Research, 1st edition.

Beerten, R., Rainford, L., and Jones, A. (2001). Changing to standard occupational classification (soc) 2000 - dual coding on the labour force survey. *Labour Market Trends. Available at: http://findarticles.com/p/articles/mi_qa3999/is_/ai_n8955599.*

Bentley, R. (2005). Publication of jobcentre plus vacancy statistics. *Labour Market Trends.* ONS.

Bilias, Y., Chen, S., and Ying, Z. (2000). Simple Resampling Methods for Censored Regression Quantiles. *Journal of Econometrics*, 99(2): 373–386.

Brown, S. and Sessions, J. (1997). A profile of UK unemployment: Regional versus demographic influences. *Regional Studies*, 31(4): 351–366.

Cahuc, P. and Zylberberg, A. (2004). *Labor Economics.* Cambridge : The MIT Press.

Collier, W. (2005). Unemployment duration and individual heterogeneity: a regional study. *Applied Economics*, 37(2): 133–153.

Cox, D. (1972). Regression models and life tables. *Journal of the Royal Statistical Society*, 34: 187–220.

DEFRA (2007). Defra classification of Local Authority Districts and Unitary Authorities in England: An introductory guide. *Available at: http://www.defra.gov.uk/rural/ruralstats/rural-defn/LAClassifications_introguide.pdf.*

Dolton, P. and O'Neill, D. (1996). Unemployment duration and the RESTART effect: Some experimental evidence. *Economic Journal*, 106(435): 387–400.

Fitzenberger, B. and Wilke, R. (2006). Using quantile regression for duration analysis. *Allgemeines Statistisches Archiv*, 90(1): 105–120.

Folmer, H. and van Dijk, J. (1988). Differences in unemployment duration - a regional or personal problem? *Applied Economics*, 20(9): 1233–1251.

Jobcentre (2008). Jobseeker's Allowance - Help while you look for work. *Available at: http://www.jobcentreplus.gov.uk/JCP/stellent/groups/jcp/documents/websitecontent/dev_015482.pdf.*

Kalwij, S. (2004). Unemployment experiences of young men: on the road to stable employment? *Oxford Bulletin of Economics and Statistics*, 66: 205–237.

Koenker, R. (2008). Censored quantile regression redux. *Journal of Statistical Software*, 27(6): 1–24.

Koenker, R. and Bassett, G., J. (1978). Regression quantiles. *Econometrica*, 46(1): 33–50.

Koenker, R. and Geling, O. (2001). Reappraising medfly longevity: A quantile regression survival analysis. *Journal of the American Statistical Association*, 96: 458–468.

Lancaster, T. (1979). Econometric Methods for the Duration of Unemployment. *Econometrica*, 47(4): 939–956.

Lüdemann, E., Wilke, R., and Zhang, X. (2006). Censored quantile regressions and the length of unemployment periods in West Germany. *Empirical Economics*, 31: 1003–1024.

Machado, J., Portugal, P., and Guimaraes, J. (2006). U.S. Unemployment Duration: Has Long Become Longer or Short Become Shorter? *IZA Discussion Papers No. 2174, Available at: http://ideas.repec.org/p/iza/izadps/dp2174.html.*

Martin (1997). Regional unemployment disparities and their dynamics. *Regional Studies*, 31(3): 237–252.

Meyer, B. (1990). Unemployment-Insurance and Unemployment Spells. *Econometrica*, 58(4): 757–782.

Mortensen, D. T. (1970). Job search, the duration of unemployment, and the phillips curve. *American Economic Review*, 60(5): 847–862.

NOMIS (2009). VAT Registrations & Deregistrations By Industry. *Available at: https://www.nomisweb.co.uk/Default.asp.*

NSPD (2007). National Statistics Postcode Directory 2007 User Guide. *Available at: http://borders.edina.ac.uk/html/pcluts_download/pcluts_2007aug.html.*

Öberg, S. and Oscarsson, G. (1979). Regional policy and interregional migration: Matching jobs and individuals on local labour markets. *Regional Studies*, 13(1): 1–14.

Office for National Statistics (2004). Counties, Non-metropolitan Districts and Unitary Authorities. *Available at: http://www.statistics.gov.uk/geography/counties_nonmet_ua.asp.*

Office for National Statistics (2007). Regional GVA: highest in London & South East. *Available at: http://www.statistics.gov.uk/CCI/nugget.asp?ID=420.*

Petrongolo, B. (2001). Re-employment probabilities and returns to matching. *Journal of Labour Economics*, 19: 716–741.

Powell, J. (1984). Least absolute deviations estimation for the censored regression model. *Journal of Econometrics*, 25(3): 303–325.

Powell, J. (1986). Censored regression quantiles. *Journal of Econometrics*, 32(1): 143–155.

Silles, M. (2008). The returns to education for the united kingdom. *Journal of Applied Economics*, 10(2): 391–413.

Spence, M. (1973). Job market signalling. *Quarterly Journal of Economics*, 87(3): 355–374.

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(423): 432–443.

van Ours, J. (2004). The locking-in effect of subsidized jobs. *Journal of Comparative Economics*, 32(1): 37–55.

Vigneau, E. and Qannari, E. (2003). Clustering of Variables Around Latent Components. *Communications in Statistics - Simulation and Computation*, 32(4): 1131–1150.

Ward, H. and Bird, D. (1995). The JUVOS Cohort: A longitudinal database of the claimant unemployed. *Employment Gazette, September*, pages 345–350.

Wilke, R. (2009). Unemployment Duration in the United Kingdom: An Incomplete Data Approach. *Nottingham University Working Paper, mimeo. Available at: http://www.nottingham.ac.uk/economics/research/dp/school_dp/dp.09.02.pdf.*

Appendix

A1: Summary Statistics

Table 6: Description & summary statistics of individual-level covariates, 1999-2005

Variable	Mean	SD
Calendar time		
Years (ref. 1999):		
y2000	.158	.365
y2001	.156	.363
y2002	.137	.344
y2003	.139	.346
y2004	.125	.331
y2005	.121	.326
Quarters (ref.q1):		
q2	.231	.422
q3	.244	.429
q4	.263	.440
Socio Demographics:		
Age<25	.493	.500
Age>55	.012	.110
Female	.223	.416
Occupation (ref. unkown):		
Elementary	.374	.484
Manufacturing	.071	.257
Trade & Services	.364	.481
Technical	.048	.213
Senior & Professional	.057	.231
Work History variables:		
Active Labour Market Participation	.146	.354
Long-Term Unemployment	.357	.479
Incapacity Benefits	.059	.237
Income Support	.014	.117
Number of obs = 187,032, 39.1% censored.		
Min/Median/Max duration in days: 1/60/2,899.		

Source: Joint Unemployment and Vacancies Operating System (JUVOS) 5% cohort.

A2: Linking the Individual & Regional Levels

This appendix briefly describes how the link between the individual and regional data was established. For more details and a full description of the regional data see Ball (2009).

Overview of Process

Main data sources included the JUVOS, National Statistics Postcode Directory (NSPD), NOMIS and the Local Area Quarterly Labour Force Survey (available from the UK Data Archive). The linked data set matching the individual- and regional-level data to the UK geography is conditioned on the start of claimant spells. In order to match the continuous individual-level data to the regional information, individual spells were matched to the regional information pertaining to the month in which they started. Merging the two data sources was a non trivial exercise which involved several technical difficulties due a lack of a one-to-one link between regional entities. Due to censorship of the full postcode information in the individual-level JUVOS data[19], this introduced an *overlapping regions problem*, removing the one-to-one link between the individual- and regional-levels. In addition a one-to-one match between local authorities and NUTS3 regions does not exist for Scotland.

In an attempt to overcome this problem, postcode districts were matched to full postcode information using the National Statistics Postcode Directory (NSPD). Merging schemes were defined in order to create a one-to-one link between the different regional classifications. Although more complicated methods are available, e.g. map-based area interpolation (see Arntz and Wilke, 2007), a simple average weighting method was employed that assigns a postcode district to the local authority in which it most falls based on the full postcode information. This link was established for all regional definitions of interest resulting in a one-to-one link between the postcode district, local authority and NUTS3 levels of aggregation.

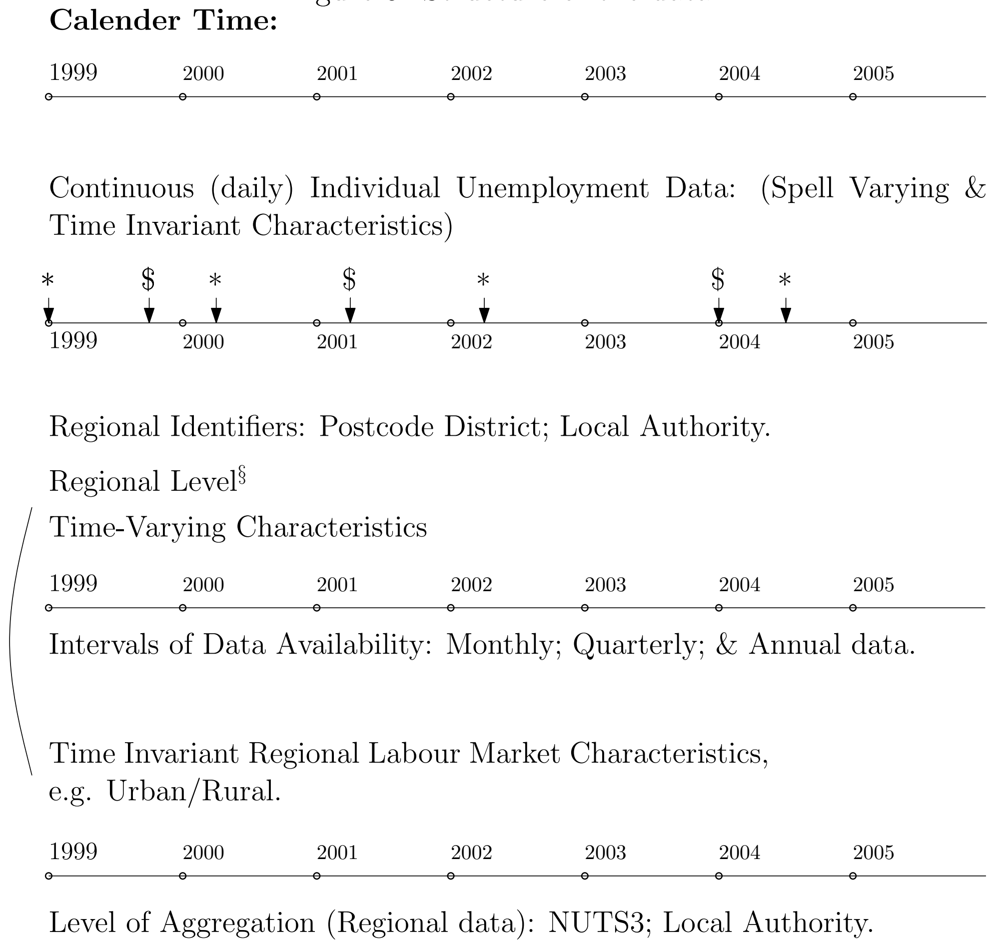
Regional Identifier

Our aim is to exploit the regional variation in the JUVOS in order to create a link between individual-level and the economic and institutional environment in which claimants reside. To this end, we identified the following geographical information in the JUVOS:

- Self-reported residential postcode data (censored to the postcode district level).
- Unemployment Benefit Office (UBO) codes.

[19] Only postcode district information is available in the JUVOS

Figure 5: Structure of the data:

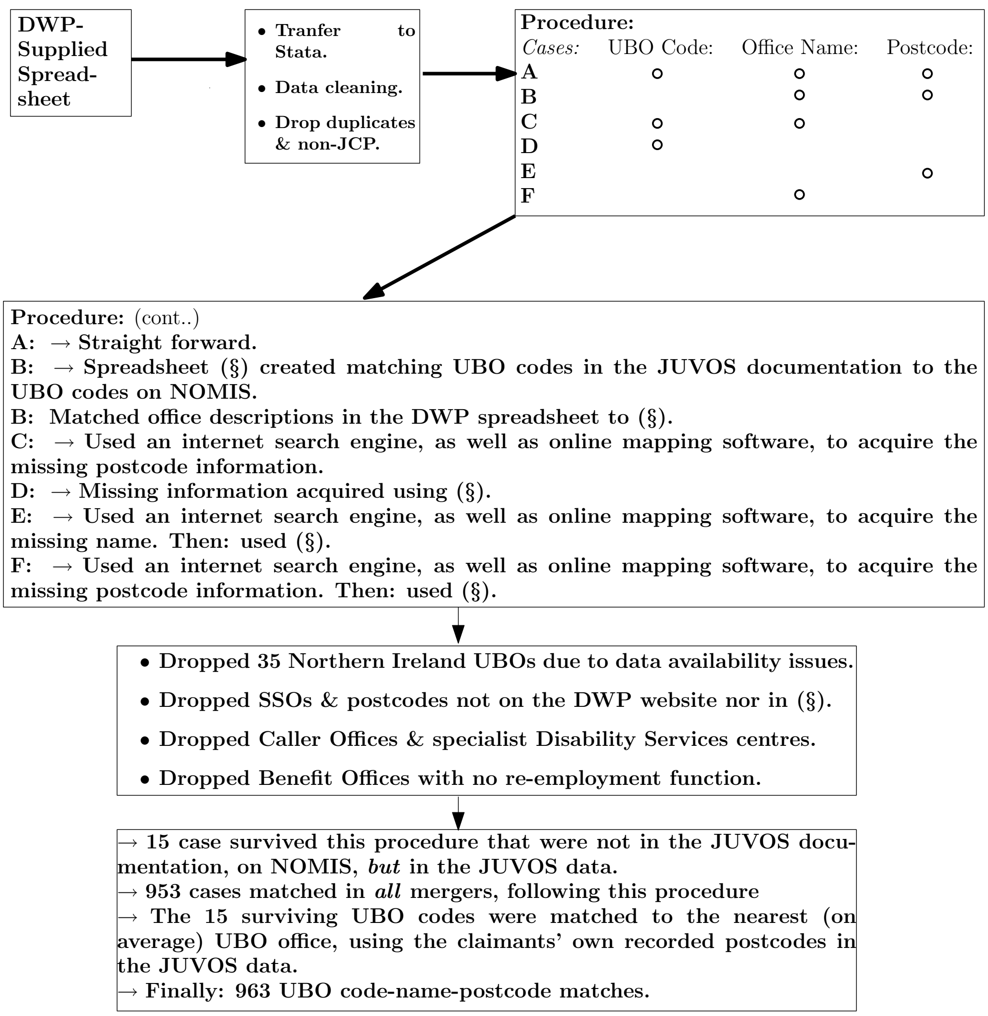


* = Individual characteristics collected at the start of each unemployment spell (assuming constant during spell).
$ = Exit destinations (restricted to Employment vs. Non-Employment).
§= Regional data merged to individual data by *month* of claimant spell commencement.

Given the self-reported nature of the first option, we were faced with data quality issues were present with postcode information missing or wrongly imputed at times. In order to improve the quality of this indicator, we used the following imputation strategy: *Replace the current postcode with the self-reported postcode during the relevant claimant's previous unemployment spell (assuming that the individual did not move during the intervening period).* This strategy was implemented in 2.8% of the cases. In order to maintain some regional variation we only replaced this self-reported variable with the UBO postcode district when this variable was missing and no information could be obtained from previous spells (implemented in 2% of cases). Each observation in the JUVOS is reported by Unemployment Benefit Office, which allows us to complete this assignment in a relatively straightforward manner. The UBO postcodes were derived using the following procedure.

A spreadsheet containing detailed (but incomplete) information about benefit office locations was sourced from the Department of Work & Pensions (DWP). In order to prepare this information for use we followed the procedure outlined in Figure 2.

Figure 6: Procedure for construction of Unemployment Benefit Office indicator:



· SSO = Social Security Office
· (§) = Spreadsheet matching UBO codes provided in the JUVOS documentation to the codes published on NOMIS.
· The JUVOS dataset is conditioned to include post 01/01/1996 data only.

The first step was to clean the data in the supplied spreadsheet, dropping certain entries and duplicates as well as checking whether the supplied information matched that available from the Department of Work & Pensions (DWP) online system[20]. Cases with missing UBO codes were

[20]Available at: http://www.jobcentreplus.gov.uk/JCP/Aboutus/Ouroffices/Search/LocalOfficeSearch.aspx

noted, and where necessary postcode information was ammended using internet search engines, Job Centre web pages, as well as the aforementioned DWP online search system. In some cases, all that was missing was the relevant postcode. However this problem was easily overcome by following the above procedure.

Jobcentres and jobcentre plus with the same postcode were assigned to the same UBO code, i.e. The Jobcentre was dropped. The spreadsheet provided by the DWP contained Social Security Office (SSO) locations. Since these offices are exclusively for the receipt of benefits and have no job related function, we decide to drop this information from the data. Specialist Disability Services centres were also dropped. After conditioning on post 01/01/1996 data, there were 963 Unemployment Benefit Office code-location matches in the data.