

# A Picture of Long-Term Unemployment in England and Wales

Melvyn Coles (University of Essex and I.A.E. (ICREA))

Paul Jones (University of Essex)

Eric Smith (University of Essex)

December 2002.

## Abstract

This paper develops a new methodology for measuring long-term unemployment. The principal drawback of the standard measure of long-term unemployment (the proportion unemployed with current spell exceeding 12 months) is that it does not distinguish between workers who have currently short unemployment durations and workers who expect to be unemployed into the long-term. The approach outlined here instead identifies the matching parameters of interest - the incidence and average matching rates of the long-term unemployed. Estimates for England and Wales (1986 to 2000) find that the expected duration of long-term unemployment peaked in the middle of the 1990-92 recession (at 15 months), and fell to 9 months by Jan 1999. The incidence of long-term unemployment also peaked during the recession. Cross section estimates identify a 'North-South' divide and a large city effect - the long-term unemployed in the largest cities, and in the 'North', experience longer spells of unemployment.

Corresponding Author: M.G. Coles, Dept of Economics, University of Essex, Colchester. CO4 3SQ. [mcole@essex.ac.uk]

Keywords: Long-term Unemployment, Stock-Flow Matching.

## 1. Introduction

This paper develops a new methodology for understanding the incidence and expected duration of long-term unemployment. In Britain, long-term unemployment has been a significant and persistent problem. For example, Layard, Nickell and Jackman (1991, p. 228) report that for workers in the unemployment *stock* (in 1989), the average duration of an uncompleted spell of unemployment was twenty-one months. In contrast, for *flow entrants* into the pool of unemployment, the average duration of a completed spell of unemployment was a much lower seven months. Machin and Manning (1999) further report that (in 1995) 44% of those unemployed in the U.K. had been unemployed for more than twelve months (61% for more than six months).<sup>1</sup>

Although the search framework provides a useful benchmark for describing the re-employment rates of job seekers by unemployment duration (e.g., Machin and Manning, 1999), it is perhaps not ideally suited for studying long-term unemployment. Given average uncompleted jobless spells typically exceed a year, it appears unlikely that such spells arise from matching frictions alone. This paper instead adopts the stock-flow matching approach to analyse long-term unemployment.<sup>2</sup> The main advantage of this approach is that the notion of long-term unemployment is directly related to a well known economic concept - the long-term unemployed are defined as job seekers who are on the long side of their particular occupations and so must wait for suitable new vacancies to enter the market. Based on this definition, the paper shows how to estimate measures of the incidence and expected duration of long-term unemployment. In contrast, the concept of long-term unemployment in the empirical search literature is vague, often measured as the proportion of unemployed workers whose current jobless spell exceeds some arbitrary threshold, usually six or twelve months. As will be shown, this latter measure of long-term unemployment has significant drawbacks.

Stock-flow matching essentially assumes workers are fully informed about the current stock of vacancies on the market. For example, help-wanted ads in newspapers and employment agencies (both private and public), along with networks of friends and relatives, provide details of available job opportunities. This in-

---

<sup>1</sup>They as well as Nickell (1997) also demonstrate that long-term unemployment is not unique to the U.K. The phenomenon is pervasive across many OECD countries with exceptions including Japan, Canada and the U.S.

<sup>2</sup>See Taylor (1995), Coles and Smith (1998), Coles and Muthoo (1998), Lagos (2000) for applications.

formation structure is the same as in the directed search approach (Montgomery, 1991; Acemoglu and Shimer, 1999). The difference between the two approaches is that directed search assumes job seekers are identical and each can only apply to one job per period, whereas stock-flow matching assumes workers are heterogeneous and each can send off multiple job applications.<sup>3</sup> In particular, stock-flow matching generates a simple sampling effect. A job seeker keeps sampling from the vacancy stock until either employment is found or there are no further vacancies to be sampled. Once a job seeker has failed to find a match with existing vacancies, the worker becomes long-term unemployed - this worker must now wait for suitable vacancies to come onto the market.

Coles and Smith (1998) obtain compelling evidence in favour of this sampling effect (also see Gregg and Petrongolo, 1999; Andrews and Bradley, 2001; Coles and Petrongolo, 2002). Using aggregate labour market information for England and Wales, Coles/Smith find that the re-employment probabilities of those workers unemployed for more than one month are highly correlated with the inflow of new vacancies, and uncorrelated with the stock. Their explanation is that such workers have had the time to fully sample the stock of current vacancies on the market, did not find a suitable match, and so wait to match with new vacancies. Coles/Smith also find that recently unemployed workers (those with durations less than one month) experience re-employment rates which are significantly correlated with the stock of vacancies, but that the correlation disappears as the duration of the worker's unemployment spell increases.

The stock-flow matching structure therefore implies that the stock of long-term unemployed workers matches with the inflow of new vacancies. This paper uses that framework to decompose information on matching and unemployment spells to provide measures of

- (a) the probability of becoming long-term unemployed (conditional on entering unemployment) and
- (b) the expected duration of unemployment (conditional on being long-term unemployed).

Based on a monthly panel of local labour market data in England and Wales (from 1986 to 2000), the results find that the incidence and expected duration of long-term unemployment is highest in the recession. Averaging across local markets,

---

<sup>3</sup>Albrecht, Gautier and Vroman (forthcoming) consider the implications of multiple applications without the sampling effect.

the estimated duration of long-term unemployment peaked around January 1992 (the middle of the recession) at approximately fifteen months. Following this peak, the average duration of long-term unemployment fell gradually over time to around the nine month mark by 1999.

These findings differ markedly from the time series properties of the conventional measure of long-term unemployment - the proportion of those unemployed with spells exceeding six or twelve months. That measure instead falls at the onset of the recession which, naively interpreted, suggests that long-term unemployment is less of a problem in recessions.<sup>4</sup> Our estimates show why this outcome is misleading. The recession is in fact characterised by an increased inflow of workers into the pool of long-term unemployment. As such workers initially have short unemployment spells, the proportion of workers with durations exceeding 12 months falls. This latter measure of long-term unemployment is flawed as it does not distinguish between workers who have short durations and workers who expect to be unemployed into the long-term.

The paper also describes the variation in long-term unemployment across local labour markets, based on Travel-To-Work-Areas (TTWAs). The results find there is much more variation in the expected duration of long-term unemployment across markets than in incidence. Estimates identify not only a large city effect - the expected duration of long-term unemployment is significantly higher in the big cities (London, Manchester, Birmingham) - but also a North-South divide. The expected duration of long-term unemployment is lowest in the South East and significantly higher in the North.

The paper consists of two main parts. The first part demonstrates how to decompose information on matching and spells of unemployment into measures of the incidence and expected duration of long-term unemployment. As the data are recorded as a monthly time series while matching is considered as a continuous time process, the paper controls econometrically for temporal aggregation of the data. The second part implements this procedure for each TTWA in England and Wales. To provide an overall picture of what happened in England and Wales, these figures are first aggregated across local markets in each period. The final section then considers cross-section variation in long-term unemployment across the TTWAs.

---

<sup>4</sup>see figure 9 in the conclusion.

## 2. Framework

At any point in time  $t$ , stock-flow matching is described by a pair  $(p(t), \lambda(t))$ .  $p(t)$  is the proportion of entrants into unemployment at date  $t$  who find a suitable job opportunity in the current stock of vacancies and so quickly match. Workers who do not find a suitable vacancy currently on the market become long-term unemployed (LTU from now on) and need to wait for suitable new vacancies to come onto the market.  $1 - p(t)$  therefore describes the incidence of LTU (conditional on becoming unemployed) and  $\lambda(t)$  then describes the average matching rate of LTU workers. Note, the random matching, the directed search and the efficiency wage approaches all assume  $p = 0$  and  $\lambda$  then describes the average matching rate of all job seekers.

In what follows, we do not test the stock-flow matching structure on the data but use it as an identifying assumption. It is worth, however, briefly documenting the extent to which stock-flow matching is consistent with the data. In its favour, stock-flow matching is consistent with the observation that roughly a quarter of new vacancies posted in Job Centres are filled on the first day. The explanation being that there is a large stock of LTU job seekers who frequently check the vacancy boards and quickly snap up good new vacancies. It is therefore not surprising that aggregate matching rates are highly correlated with the inflow of new vacancies and that formal tests against the random matching function approach using aggregate data find evidence in favour of stock-flow matching (Coles and Smith, 1998; Gregg and Petrongolo, 1999; Coles and Petrongolo, 2002).

On the other hand, there is little evidence of workers matching on the first day of unemployment. Indeed, the results of Coles and Smith (1998) suggest that workers who match with the initial stock of vacancies experience unemployment spells of up to a month before starting new employment. One potential explanation for this observation is that once filled, vacancies are quickly withdrawn to deter further job applicants, but the starting date for employment may be set at some mutually convenient future time. Alternatively, search frictions may bind for workers on the short side of the market who might take, say, two or three weeks to decide on their preferred employment opportunity before starting work. Such delays smooth out the distribution of completed unemployment spells at short durations. The focus here, however, is on the matching experience of the LTU who are on the long side of the market and frequently realise unemployment spells exceeding a year. For econometric tractability, this paper abstracts from short side frictions by assuming that entrants who find a suitable vacancy in the

current stock match arbitrarily quickly. This is clearly a strong assumption. We shall return to it in the Conclusion and argue that this simplification biases the estimated expected duration of long-term unemployment upwards by a few weeks. But given the mean expected duration of long-term unemployment exceeds a year, this bias seems small.

The specification that all LTU workers match at the same rate  $\lambda(t)$  is also inconsistent with evidence that average hazard rates typically decline with duration. This phenomenon is usually explained by unobserved heterogeneity, where perhaps different job seekers choose different search efforts (e.g. Lancaster and Nickell, 1980; Heckman and Singer, 1984; Meyer, 1990; Portugal and Addison, 2000). But this observation is also consistent with heterogeneous search efforts in a stock-flow matching economy - different LTU workers may check vacancy boards with different frequencies. Section 2.3 below returns to this issue in detail.

Imposing stock-flow matching as an identifying assumption, the next two subsections demonstrate how to infer values of  $(p, \lambda)$  directly from the data. These estimates are then used to describe the incidence  $(1 - p)$  and expected duration of long-term unemployment, taking into account that these matching parameters vary over the business cycle.

## 2.1. Matches Per Period

Let  $u(t)$  denote the inflow of newly unemployed workers and  $U(t)$  denote the stock of unemployed workers at time  $t$ . Assuming time is continuous, then stock-flow matching implies the match flow,  $M(t)$ , is given by

$$M(t) = p(t)u(t) + \lambda(t)U(t) \tag{1}$$

where proportion  $p(t)$  of the unemployment inflow  $u(t)$  match immediately (or at least very quickly), while the stock of unemployed workers  $U(t)$  match at rate  $\lambda(t)$ . Given data on matching ( $M$ ), unemployment stocks ( $U$ ) and inflows ( $u$ ), this equation provides the first identifying equation for  $p(t)$  and  $\lambda(t)$ .

An unfortunate complication, however, is that the identifying relationships are set in continuous time, but the available data are recorded monthly. We deal with this time aggregation problem using the arguments of Gregg and Petrongolo (1999). Let  $t \in [n, n + 1)$  denote calendar time in each month  $n \in \mathbb{N}$  and assume that the unemployment inflow and matching parameters are constant within the month; i.e. for each month  $n$ , assume

$$u(t) = u_n, \lambda(t) = \lambda_n, p(t) = p_n \text{ for all } t \in [n, n + 1).$$

At time  $t$ , the unemployment stock evolves according to the differential equation

$$\frac{dU}{dt} = u_n [1 - p_n] - \lambda_n U(t)$$

where  $u_n[1 - p_n]$  describes the net inflow of entrants into unemployment, while  $\lambda_n U(t)$  describes the outflow. Integrating implies

$$U(t) = \left[ U_n - \frac{u_n}{\lambda_n} (1 - p_n) \right] e^{-\lambda_n(t-n)} + \frac{u_n}{\lambda_n} (1 - p_n)$$

where  $U_n = U(n)$  denotes the level of unemployment at the start of month  $n$ . As the total number of matches during the month,  $M_n$  satisfies

$$U_{n+1} - U_n = u_n - M_n$$

Using  $U_{n+1} = U(n+1)$ , the previous expression yields

$$M_n = U_n(1 - e^{-\lambda_n}) + p_n u_n + u_n(1 - p_n) \left[ 1 - \frac{1 - e^{-\lambda_n}}{\lambda_n} \right]. \quad (2)$$

Equation (2) says that the number of matches within the month includes those who were unemployed at the start of the month and found work, those who entered the market and immediately found work, and those entrants who became LTU but were fortunate enough to find a job before the end of the month. As we have data on  $(M_n, U_n, u_n)$ , equation (2) describes our first piece of information for  $(p_n, \lambda_n)$ .

## 2.2. Average Durations of Unemployment Spells.

The second piece of information used to estimate  $(p, \lambda)$  compares the average duration of uncompleted spells of unemployment against the average duration of completed spells. In continuous time and conditional on a filled vacancy, stock-flow matching implies the average completed spell of unemployment,  $X^c(t)$ , is

$$X^c(t) = \frac{\lambda(t)U(t)}{\lambda(t)U(t) + p(t)u(t)} X^u(t), \quad (3)$$

where a LTU worker who matches has average uncompleted spell  $X^u(t)$ , while a newly unemployed worker on the short side who matches has an (arbitrarily) short unemployment spell. Given information on the ratio of average unemployment

spells,  $X^c(t)/X^u(t)$ , (3) is a second identifying equation for  $p(t)$  and  $\lambda(t)$ . Note that if  $p(t) \equiv 0$ , so that all unemployed workers match at the same rate  $\lambda(t)$  (as in the standard random matching model), a matched worker is then a random draw from the current stock of unemployed workers and so the average duration of a completed spell equals the average uncompleted spell. On the other hand, if  $p(t)$  is strictly positive, a significant fraction of those finding employment are those who have been recently laid off and have quickly found work. A higher  $p(t)$  implies a smaller ratio  $X^c(t)/X^u(t)$ .

As the data only record average completed spells over the month then, as in the previous section, estimates have to correct for temporal aggregation bias. Let  $X_n^C$  denote the mean completed unemployment spell in month  $n$ . By definition,  $X_n^C$  is the sum of spell lengths of all workers who match during the month,  $\Sigma_n$ , divided by the number of matches,  $M_n$ ; i.e.

$$X_n^C = \Sigma_n/M_n$$

Assuming newly unemployed workers on the short side of the market match (arbitrarily) quickly,  $\Sigma_n$  is determined (almost) entirely by completed spells of unemployment by LTU workers.

These completed spells come from one of two instances. Either

- (a) a newly unemployed worker initially fails to match (with probability  $1 - p_n$ ), becomes LTU but is still fortunate enough to match before the end of the month. The sum of spell lengths from such workers is denoted  $\Sigma_n^F$ ;

or

- (b) a worker is already in long-term unemployment at the start of the month and finds work before the end of the month. The sum of spell lengths from such workers is denoted  $\Sigma_n^S$ .

Total spell lengths  $\Sigma_n$  is simply the sum of these two objects

$$\Sigma_n = \Sigma_n^F + \Sigma_n^S$$

- (a) Calculating  $\Sigma_n^F$ : Consider the inflow  $u_n(1 - p_n)$  of workers into the stock of long-term unemployment. For each entry date  $t \in [n, n + 1)$ , these workers match within the rest of the month according to a Poisson process with parameter  $\lambda_n$ . Over an arbitrarily small time interval  $ds > 0$ , proportion  $[\lambda_n \exp(-\lambda_n(s - t))] ds$

of these workers find work at date  $s \in (t, n + 1)$  with a corresponding completed unemployment spell  $(s - t)$ . Adding up all those spells implies

$$\Sigma_n^F = \int_{t=n}^{n+1} \left\{ u_n(1 - p_n) dt \left[ \int_{s=t}^{n+1} (s - t) [\lambda_n e^{-\lambda_n(s-t)}] ds \right] \right\}$$

Integration establishes

$$\Sigma_n^F = \frac{u_n(1 - p_n)}{\lambda_n} \left[ (1 + e^{-\lambda_n}) - 2 \left( \frac{1 - e^{-\lambda_n}}{\lambda_n} \right) \right].$$

(b) Calculating  $\Sigma_n^S$ : Recall that  $U_n$  describes the stock of unemployed workers at the start of the month, and let  $X_{n-1}^U$  denote the mean uncompleted spell of unemployment carried over from the previous period. Since LTU workers match according to a Poisson process with parameter  $\lambda_n$ , it follows that

$$\Sigma_n^S = U_n \int_n^{n+1} \lambda_n e^{-\lambda_n(t-n)} [X_{n-1}^U + t - n] dt$$

The  $\lambda_n \exp(-\lambda_n(t - n)) dt$  term describes the proportion of LTU workers from last period's stock who find work (over small time period  $dt > 0$ ) at date  $t \in [n, n + 1)$ . Their average duration of unemployment is the average uncompleted spell inherited from the previous period,  $X_{n-1}^U$ , plus the additional time  $(t - n)$  spent in unemployment this period. Integration this time gives

$$\Sigma_n^S = U_n \left[ (X_{n-1}^U + \frac{1}{\lambda_n})(1 - e^{-\lambda_n}) - e^{-\lambda_n} \right]$$

Using  $X_n^C = [\Sigma_n^F + \Sigma_n^S]/M_n$  implies that the average duration of completed spells satisfies:

$$\begin{aligned} X_n^C &= \frac{u_n}{M_n} \left( \frac{1 - p_n}{\lambda_n} \right) \left[ (1 + e^{-\lambda_n}) - 2 \left( \frac{1 - e^{-\lambda_n}}{\lambda_n} \right) \right] \\ &\quad + \frac{U_n}{M_n} \left[ (X_{n-1}^U + \frac{1}{\lambda_n})(1 - e^{-\lambda_n}) - e^{-\lambda_n} \right] \end{aligned} \quad (4)$$

Given data on  $(X_n^C, X_n^U, U_n, u_n, M_n)$ , equations (2) and (4) identify the underlying matching parameters  $(p_n, \lambda_n)$ . Although the equations are non-linear, it is straightforward to compute solutions using standard numerical techniques.

### 2.3. Search Heterogeneity

A criticism of the above approach is that it assumes all LTU job seekers match at the same rate. This subsection briefly considers introducing search heterogeneity across the LTU. To keep things simple, suppose all LTU workers can match with each new vacancy (e.g; an efficiency wage story as in Shapiro and Stiglitz, 1984). In addition, assume that the first job seeker on the long side who contacts a new vacancy gets the job.

Stock-flow matching implies that the LTU chase new vacancies as they come onto the market. The frequency with which a LTU worker checks the vacancy boards then affects that worker's probability of winning the next vacancy chase. It does not, however, affect the aggregate outcome - a vacancy on the short side is always quickly filled by someone.<sup>5</sup>

More precisely, suppose a LTU worker  $j$  checks Job Centre vacancies with frequency  $k_j$ , and let  $K = \sum_j k_j$  denote aggregate search effort by the long-term unemployed,  $j = 1, \dots, U$  where  $U$  denotes the number of LTU workers. Job rationing implies individual  $j$ 's re-employment rate is

$$\lambda_j = \frac{k_j}{K}v,$$

where  $v$  describes the inflow of new vacancies. A worker who checks the ads twice as often is twice as likely to win the next vacancy chase. Note, however, that if all job seekers double their search efforts, no change in matching probabilities occurs. There are only pure displacement effects - the worker who wins the next vacancy chase does so at the cost of the other workers. Moreover, the aggregate matching rate of all LTU workers,  $\sum_j \lambda_j$ , is given by  $\sum_j \lambda_j = v$ , which is independent of  $\{k_j\}_{j=1}^U$ . The number of LTU workers who match depends entirely on the inflow of new vacancies.

Should some LTU workers check the vacancy boards more frequently than others, the stock-flow matching framework is consistent with negative duration dependence in the hazard function. But most importantly, dispersion in  $k_j$  does not affect the aggregate outcome. A new vacancy on the short side is quickly filled by a LTU worker. The underlying point is that identifying matching behaviour using micro level data is problematic given individual  $k_j$  are unobserved (and

---

<sup>5</sup>Given the data is recorded as a monthly time series, it is immaterial whether a short-side vacancy is filled the same day or within a few days of being posted.

potentially time varying). In contrast, the representative LTU worker who chooses average search effort  $\bar{k} = K/U$  matches at rate  $\lambda = v/U$ . The average matching rate is then independent of the unobserved  $\{k_j\}$ , and so can be identified using aggregate data.

Nevertheless, dispersion in  $\{k_j\}$  does generate interesting composition effects. Figure 7 below reveals that the expected duration of long-term unemployment (computed using average matching rates) peaks at fifteen months in January 1992. At first sight this figure appears to be inconsistent with the Layard et al (1991) statistic that the mean uncompleted jobless spell is twenty-one months (in 1989). Heterogeneity in  $\{k_j\}$  can account for this discrepancy. In a steady state, the average expected duration of unemployment for all LTU workers is

$$\overline{ED} = \frac{1}{U} \sum_{j=1}^U \frac{1}{\lambda_j} = \frac{1}{v} \sum_{j=1}^U \frac{\bar{k}}{k_j}.$$

If some choose zero search effort,  $k_j = 0$ , then  $\overline{ED}$  must be arbitrarily large and so the mean uncompleted spell of unemployment in the stock of unemployed workers can also be arbitrarily large. More interestingly, note that  $\overline{ED}$  is a convex function of  $k_j$ . A mean preserving spread in  $k_j$  implies an increase in  $\overline{ED}$ . The average duration of unemployment for all LTU workers is therefore minimised when all choose the same search effort; i.e.  $k_j = \bar{k}$  in which case  $\lambda_j = \lambda = v/U$ .

By identifying the average matching rate of the LTU, the above suggests that the estimates below of the expected duration of unemployment are a *lower bound* for the average duration of uncompleted unemployment spells. Of course, variations in  $\{k_j\}$  do not have any implications for aggregate matching, as matching by the LTU is driven by the arrival of suitable new vacancies. Variations in  $k_j$ , however, affect the distribution of unemployment spells across the LTU.

### 3. Estimation

The previous section describes how to identify the incidence of LTU,  $1 - p_n^i$ , and the average re-employment rate of LTU workers,  $\lambda_n^i$ , for labour market  $i$  in month  $n$ . (The notation here is extended to allow for spatially distinct labour markets.) Using data on local labour market matching, this section now identifies how LTU has varied over time and across regions in England and Wales from March 1986

until December 2000.<sup>6</sup>

### 3.1. Data

In the UK, the Travel-To-Work-Area (TTWA) is the standard measure of a self-contained labour market. Based on census figures, these are geographical regions that have:

- a minimum of 3500 residents;
- at least 75% of the people working in the area live in it;
- at least 75% of those living in the area also work there.

There are 254 such areas in England and Wales.

The working population ( $POP^i$  in market  $i$ ) is concentrated in a few, very large markets. From the 1991 census figures, the largest TTWA is London which has a working age population of 3.8 million workers whereas the mean working age population across all TTWAs is a mere 120,000. The median population is an even smaller 64,000 while the standard deviation is 268,000. To characterise the representative experience of a LTU worker in the England and Wales, in what follows observations of each TTWA  $i$  are weighted by the proportion of total working age adults who live in that market.

For each TTWA  $i$  in each month  $n$ , the Department of Employment reports:

- unemployment inflow ( $u_n^i$ ) during the month;
- unemployment outflow ( $M_n^i$ ) disaggregated by time spent unemployed;
- unemployed stock ( $U_{n+1}^i$ ) at the end of month  $n$ , broken down by the duration of the current spell of unemployment.

In the data, unemployment is defined as the number of people claiming unemployment-related benefits. All workers claiming state benefit are required to ‘sign on’ at the

---

<sup>6</sup>All data are available from the National On-line Manpower Information service (NOMIS) located at the University of Durham and homepage <http://www.nomisweb.co.uk/>

unemployment benefit office on becoming unemployed and ‘sign off’ on leaving unemployment.<sup>7</sup> The flow of individuals signing off unemployment benefits is used to capture matched outflow  $M_n^i$ . This outflow is assumed to be into employment thereby ignoring any movements into states of non-participation. Figures relating to time spent unemployed are available for sixteen duration categories: 0-1 week, 1-2 weeks, 2-4 weeks, 4-6 weeks, 6-8 weeks, 8-13 weeks, 13-26 weeks, 26-39 weeks, 39-52 weeks, 52-65 weeks, 65 to 78 weeks, 78 to 104 weeks, 104 to 156 weeks, 156 to 208 weeks, 208 to 260 weeks, and greater than 260 weeks.

Given this definition of unemployment, Figure 1 plots aggregate unemployment in England and Wales from March 1986 until December 2000. Defining a recession as two consecutive quarters of rising unemployment, it follows that the aggregate economy was in recession from March 1990 to January 1993 (as marked by dashed vertical lines). During this recession, total unemployment more than doubled, rising from 1.2 million to 2.6 million.

Figure 2 plots aggregate unemployment inflows and outflows over this period (seasonally detrended and standardized to a four week month).<sup>8</sup> The main feature of the recession is that both the inflow and outflow from unemployment are initially low, and both increase over time during the recession. The difference between these two series is that the inflow increases more rapidly at the beginning of the recession. Taken together, these features generate two important implications for long-term unemployment. First, low outflow rates imply that it takes longer for each worker to find suitable work, which implies a relatively high expected duration of LTU. Second, the number of LTU workers increases over the recession. With more LTU workers chasing the same new vacancies, the expected duration of LTU increases still further.

Using the data described above, it is straightforward to compute the average duration of completed and uncompleted spells of unemployment. Figure 3 graphs these series for England and Wales. Note the large gap between the two series. This implies that the newly unemployed match much more quickly than the stock of long-term unemployed workers. These data also exhibit a curious property - the average uncompleted and completed spell of unemployment is lowest in the

---

<sup>7</sup>Unemployment-related benefits are currently: the Jobseekers Allowance(JSA) introduced from October 1996, and National Insurance (NI) credits. A claimant must declare that they are out of work, capable of, available for and actively seeking work during the week in which their claim is made. This count differs very slightly from the alternative count of people who register at the Job Centre seeking work.

<sup>8</sup>The estimates for  $p_n^i, \lambda_n^i$  below are based on the raw, non-seasonally adjusted data.

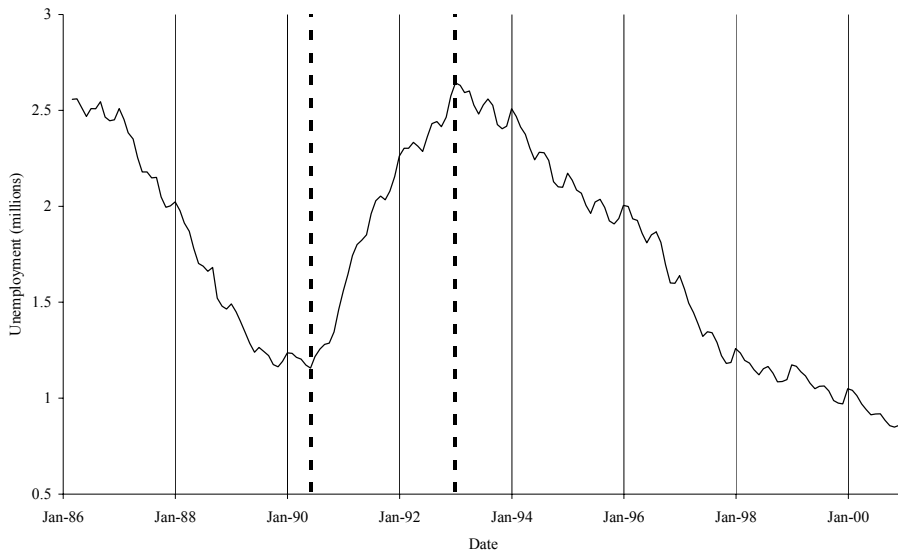


Figure 1: Aggregate Unemployment in England and Wales

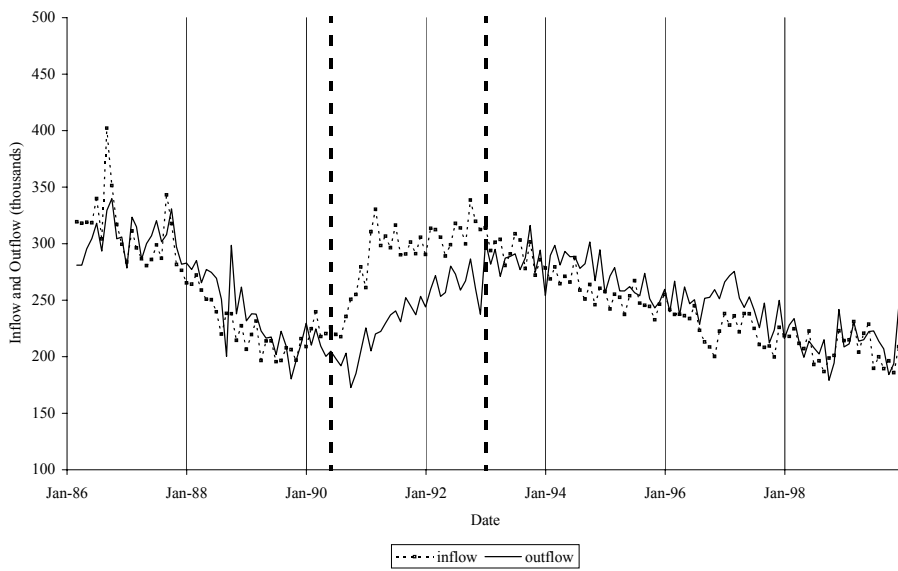


Figure 2: Aggregate Unemployment Inflow and Outflow

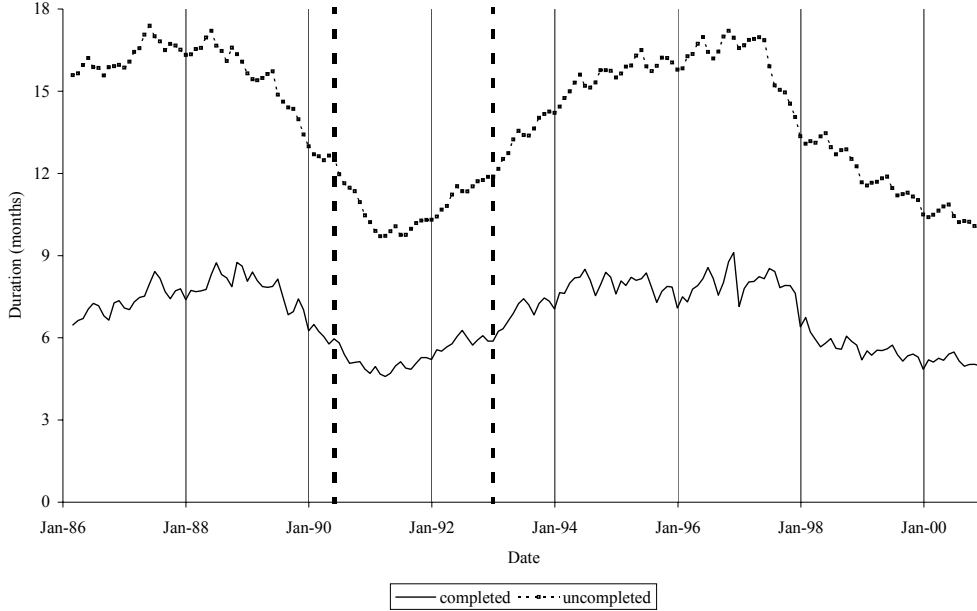


Figure 3: Average Duration of Completed and Uncompleted Spells

recession. The standard measure of LTU (the proportion unemployed with current unemployment spells exceeding a year) also exhibits this behaviour- it too falls at the onset of the recession (see Figure 9 in the Conclusion). Based on the estimates for  $(p(t), \lambda(t))$  obtained below, the time series features of these data will be fully explained later.

For each TTWA  $i$  and each month  $n$ , the data described above are used to compute the average duration of completed  $(X_n^C)^i$  and uncompleted spells  $(X_n^U)^i$ , thereby creating the panel of data  $(X_n^C, X_n^U, U_n, u_n, M_n)^i$ . In addition, for each TTWA  $i$ , let the average unemployment rate be  $\bar{u}^i = \bar{U}^i / POP^i$  where  $\bar{U}^i = \sum_n U_n^i / N$ . Let the average re-employment rate be  $\bar{M}^i / \bar{U}^i$  where  $\bar{M}^i = \sum_n M_n^i / N$ . Figure 4 presents a weighted scatter plot of these mean statistics across TTWA, where the size of a ‘bubble’ reflects the population weight of the TTWA.

Not surprisingly, re-employment rates are negatively correlated with unemployment rates. The largest ‘bubble’ in Figure 4 is London which has a high average unemployment rate and a low average re-employment rate. Aside from London, there is no clear correlation between city size and unemployment - the

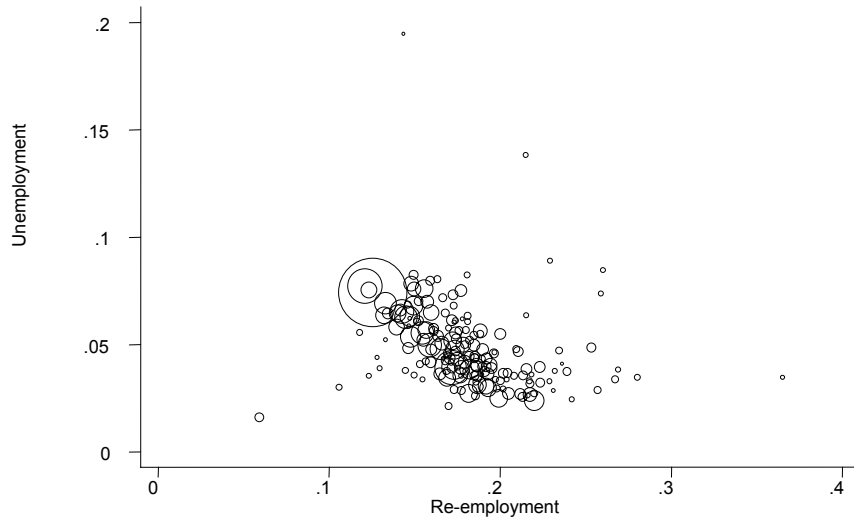


Figure 4: Re-employment Rate vs. Unemployment Rate by TTWA

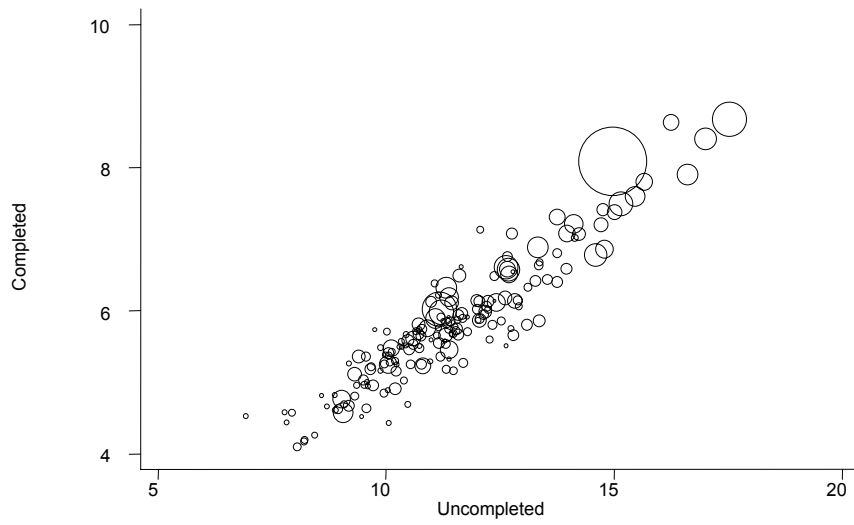


Figure 5: Completed vs. Uncompleted Spells by TTWA

larger bubbles seem evenly distributed across the scatter plot.

Figure 5 is the (weighted) scatter plot of the average completed and uncompleted spells of unemployment across TTWA. Two stylised facts are evident. First the scatter plot is approximately linear in that across all TTWA, the average completed spell of unemployment is roughly half the average uncompleted spell. Figure 5 also demonstrates a strong city size effect - the larger bubbles are concentrated at high completed and uncompleted spells of unemployment.

Given the panel of data  $(X_n^C, X_n^U, U_n, u_n, M_n)^i$ , it is straightforward to solve the matching equations (2) and (4) for a panel of estimated matching probabilities  $(p_n^i, \lambda_n^i)$ . The next subsection describes how those estimates vary over the business cycle. The subsection after that considers how those estimates vary across TTWA.

### 3.2. Long-Term Unemployment and the Business Cycle

Recall that  $1 - p_n^i$  is defined as the incidence of LTU in market  $i$ , in month  $n$ . Aggregating over the TTWA, define the average incidence of LTU in England and Wales in month  $n$  as<sup>9</sup>

$$1 - \bar{p}_n = \sum_i \omega^i (1 - p_n^i)$$

where  $\omega^i$  is the working age population weight of TTWA  $i$ . Similarly, define the average matching rate of the LTU in England and Wales as

$$\bar{\lambda}_n = \sum_i \omega^i \lambda_n^i.$$

As these statistics are highly seasonal, Figure 6 plots these statistics seasonally detrended.<sup>10</sup> The top frame of Figure 6 plots average incidence,  $1 - \bar{p}_n$ , the bottom plots the average matching rate,  $\bar{\lambda}_n$ . To illustrate TTWA dispersion over time, Figure 6 also plots the 10th and 90th percentiles of  $1 - p_n^i$  and  $\lambda_n^i$  in month  $n$  (using population weights  $\omega^i$  and seasonally detrended).

Figure 6 establishes that on average, the estimated incidence of LTU is around one half; i.e. half of those who become unemployed quickly find work, the others are stuck waiting for something suitable to come onto the market. Consistent

---

<sup>9</sup>An alternative measure can be found by first aggregating the data economy-wide and then estimating. Doing so yields nearly equivalent estimates - there are only very minor differences between the two series.

<sup>10</sup>For example, the incidence of LTU is highest in January and lowest in October.

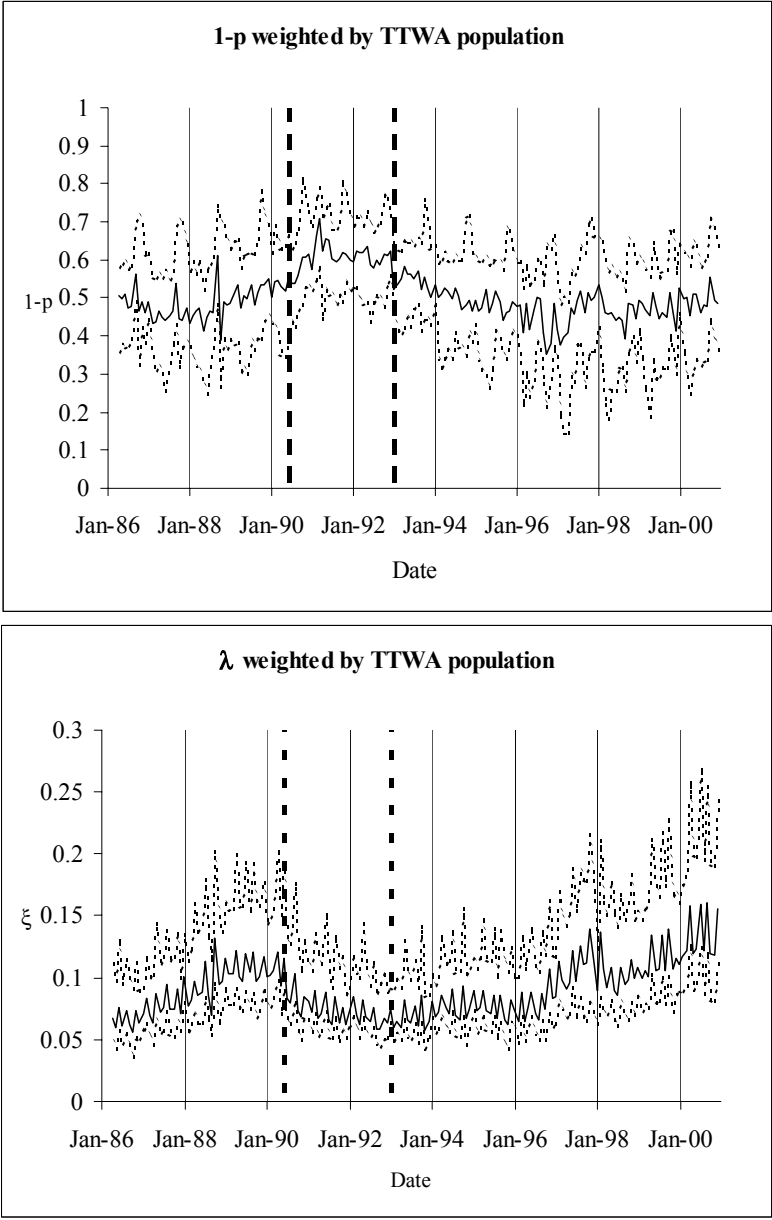


Figure 6: LTU Incidence and Matching Rates

with intuition, LTU incidence is highest in the recession - peaking at around 0.65 - and lowest in the boom with a value around 0.4 by mid-1996.

The matching rates of the LTU are also strongly countercyclical. Immediately prior to the recession, LTU monthly matching rates  $\bar{\lambda}_n$  were around 0.1 suggesting an expected long-term unemployment spell  $1/\bar{\lambda}_n$  of around 10 months. The onset of the recession resulted in a collapse of this matching rate to  $\bar{\lambda}_n \approx 0.06$  suggesting an expected unemployment spell  $1/\bar{\lambda}_n$  of 16 months. Those rates did not begin to recover until long after the end of the recession (July 1996 rather than January 1993).

A test of these results is to ask whether they are consistent with the time series properties of the unemployment spells data  $X_n^c, X_n^u$  as described in Figure 3. To see why this is a relevant test, note that in the absence of the temporal aggregation problem, (1) and (3) imply the estimates for  $(p(t), \lambda(t))$  depend only on  $X^c, X^u$  through their ratio  $X^c/X^u$ . Taking temporal aggregation bias into account implies a more complicated estimator. Nevertheless there is no reason to expect that the estimated values of  $1/\bar{\lambda}_n$  will necessarily be consistent with the observed variation in  $X_n^c, X_n^u$ .

First consider the average uncompleted spell of unemployment data. Note that prior to the recession, the average uncompleted spell of unemployment fell from 18 months to around 12 months. The estimated matching rates are  $\bar{\lambda}_n \approx 0.06$  at the start of this phase (suggesting an expected unemployment spell  $1/\bar{\lambda}_n = 16$  months), rising to  $\bar{\lambda}_n \approx 0.1$  ( $1/\bar{\lambda}_n = 10$  months) by the end. With turnover, where newly unemployed workers replace LTU workers who match, these estimates of  $1/\lambda$  are not only roughly consistent with the mean uncompleted spell averages, they clearly explain the direction of change.<sup>11</sup>

Somewhat surprisingly, the recession implies a further fall in the average uncompleted spell of unemployment, reaching a low value of 10 months. To see why this occurs, note that the recession generates a large increase in the inflow of newly unemployed workers (Figure 2) and an increase in the estimated incidence of long-term unemployment (Figure 6). Together these imply a large increase in newly unemployed workers into the stock of LTU, which lowers the average uncompleted spell of unemployment.

Following this low point, the average uncompleted spell of unemployment grows inexorably over time, reaching its highest value of around 18 months by July, 1996. During this phase, the estimated LTU matching rates collapse to

---

<sup>11</sup>Recall that in a steady state, heterogeneity in  $k_j$  suggests that the  $1/\bar{\lambda}$  estimates are a 'lower bound' for the average uncompleted spell data.

$\bar{\lambda}_n \approx 0.06$  ( $1/\bar{\lambda}_n \approx 16$  months). This not only explains the observed increase in the average uncompleted spell during this phase given the starting average of 10 months, but note that the matching rate  $\bar{\lambda}_n$  recovers after July 1997 (Figure 6) which coincides with the subsequent decline in the average uncompleted spell of unemployment.

Of course  $1/\bar{\lambda}_n$  is only a rough approximation of the expected duration of unemployment. Figure 7 below provides better measures of this statistic. Before constructing those estimates however, consider briefly the average completed spell of unemployment data. Figure 3 implies this data series has the same time profile as the uncompleted spell data series, but with roughly half its value (also see Figure 5). Given that, (3) implies  $p(t)u(t) \approx \lambda(t)U(t)$  on average over the cycle. The stock-flow matching interpretation is that roughly half of all vacancies filled are on the short side of their markets (and so are filled by the long-term unemployed), the other half are on the long side and wait to be filled by newly unemployed workers. This result is reassuringly consistent with the above estimates of LTU incidence.  $1 - \bar{p}_n \approx 0.5$  also implies that approximately half of newly unemployed workers are on the long side of their markets (and become LTU), the others are on the short side.<sup>12</sup>

Based on the estimated values of  $\bar{\lambda}_n$ , Figure 7 below plots estimates of the expected duration of long-term unemployment, denoted  $\overline{ED}_n$ . This statistic is computed by assuming an LTU worker matches at Poisson rate  $\{\bar{\lambda}_s\}_{s=n}^{\infty}$  in each month  $s$ .  $\overline{ED}_n$  is then defined recursively by

$$\overline{ED}_n = \int_n^{n+1} [t - n] e^{-\bar{\lambda}_n(t-n)} \bar{\lambda}_n dt + e^{-\bar{\lambda}_n} [1 + \overline{ED}_{n+1}].$$

Given an LTU worker at date  $n$ , with probability  $e^{-\bar{\lambda}_n(t-n)} \bar{\lambda}_n dt$  that worker gets a job at date  $t < n + 1$ , and so experiences a further  $(t - n)$  spell of unemployment. With probability  $e^{-\bar{\lambda}_n}$  the worker remains unemployed at the end of the month  $[n, n + 1)$  and so has expected duration  $[1 + \overline{ED}_{n+1}]$  in that event. Integration implies

$$\overline{ED}_n = \frac{1 - e^{-\bar{\lambda}_n}}{\bar{\lambda}_n} + e^{-\bar{\lambda}_n} \overline{ED}_{n+1}.$$

As the estimates for  $\bar{\lambda}_n$  terminate at December 2000, constructing the series  $\{\overline{ED}_n\}_{n=1}^N$  requires a terminal value for  $\overline{ED}_{N+1}$  at January 2001. Robustness

---

<sup>12</sup>Note, a 50-50 split of entrants between the short and long sides is not an implication of (long run) steady state; e.g. Coles (1999).

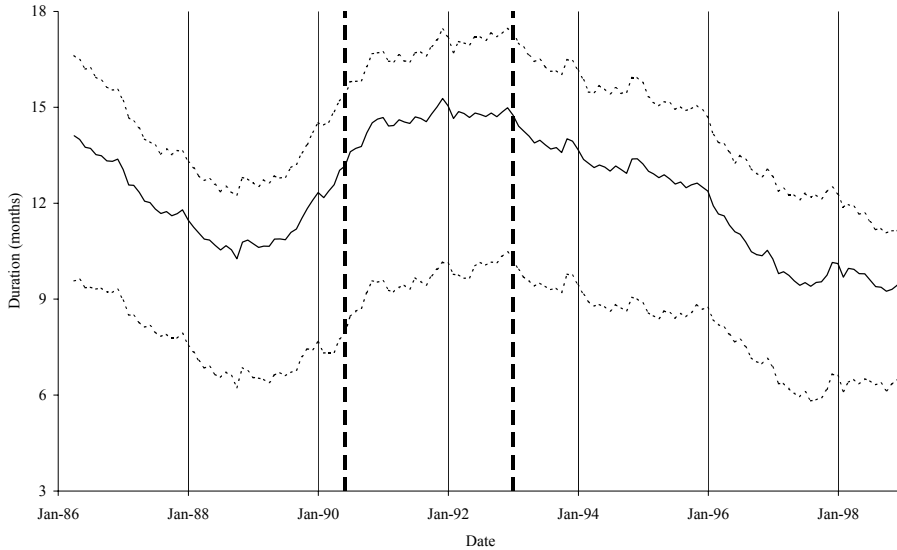


Figure 7: Expected Duration of Long-Term Unemployment

checks find that the estimates for  $\overline{ED}_n$  prior to January 1999 are largely unaffected by any reasonable choice of  $\overline{ED}_{N+1}$ . This occurs as the probability a LTU worker in January 1999 remains unemployed by January 2001 is very small and hence the choice of  $\overline{ED}_{N+1}$  has only a very small impact on  $\overline{ED}_n$  prior to January 1999.<sup>13</sup> The series for  $\overline{ED}_n$  plotted in Figure 7 in fact uses boundary value  $\overline{ED}_{N+1} = 1/\overline{\lambda}_N$ , but reflecting the above robustness issue, we only plot the estimated series  $\{\overline{ED}_n\}$  as far as January 1999. To provide some idea of dispersion across TTWAs over time, Figure 7 also displays the 90th and the 10th percentiles of  $ED_n^i$  (individually constructed using estimated values  $\lambda_n^i$ ), again weighted by  $\omega^i$ .<sup>14</sup>

As  $\overline{ED}_n$  is constructed using realised values of  $\{\overline{\lambda}_s\}_{s=n}^N$ , it is not the ‘expected duration of unemployment given period  $n$  information’. Instead conditional on how the economy evolved over time,  $\overline{ED}_n$  is the prediction on how long it took an LTU worker at date  $n$  to find employment. Consistent with intuition, the expected duration of LTU peaks in the middle of the recession. Not only were matching

<sup>13</sup>Put differently, the difference equation for  $\overline{ED}_n$  is stable when iterating backwards.

<sup>14</sup>These estimates are not seasonally adjusted - integration largely washes out the seasonal effects.

rates lowest there,  $\bar{\lambda}_n \approx 0.06$ , they remained at that low point for several more years. The stock-flow interpretation is that these workers were on the long side of their markets and were stuck waiting for suitable new vacancies to come onto the market. Further, Figure 2 suggests that the low LTU matching rate occurred not because the inflow of new vacancies was particularly small - the number of matches rose steeply during the recession - but because there was a large increase in the number of LTU workers chasing those new vacancies. By the end of the recession, the number unemployed had more than doubled to 2.6 million, and displacement effects lead to a collapse in individual LTU re-employment rates.

After the recession, Figure 7 establishes that the expected duration of LTU gradually falls over time, falling to a low of 9 months by 1999. Note, however, that this decline cannot be attributed to a large increase in the outflow of unemployed workers. Figure 2 shows the outflow gradually decreases over this time period. Instead the rise in  $\bar{\lambda}_n$  observed after July 1996 is entirely due to the decrease in the number unemployed. This results in fewer LTU workers chasing each new vacancy, which raises individual re-employment rates.

### 3.3. Cross Section Dispersion

To see the way in which LTU varies across TTWAs, define the mean incidence for TTWA  $i$  as

$$1 - \bar{p}^i = \frac{\sum_n (1 - p_n^i)}{N}$$

where  $N = 166$  is the total number of months in the data sample, and the mean expected duration as

$$\overline{ED}^i = \frac{\sum_n ED_n^i}{N}.$$

Table 1 provides the descriptive statistics for these measures.<sup>15</sup>

---

<sup>15</sup>All figures are calculated up to December 1998 since expected duration becomes unreliable after this time.

Table 1  
Descriptive Statistics

	$1 - \bar{p}^i$	$\bar{\lambda}^i$	$\overline{ED}^i$
Mean	0.510	0.084	12.35
Standard Deviation	0.033	0.019	2.56
Coefficient of Variation	0.065	0.223	0.21
10th Percentile	0.464	0.069	8.53
90th Percentile	0.553	0.120	14.25
Number of Observations	254		
All figures weighted by population			

The coefficient of variation implies there is much more variation in LTU matching rates  $\bar{\lambda}^i$  (and expected durations) across local labour markets, than in incidence  $1 - \bar{p}^i$ . For example, the average duration of LTU was 12.3 months over this sample period, but 10% of the population mass (recall, cities are weighted by population) had average LTU durations exceeding 14.2 months, while another 10% had LTU durations less than 8.5 months. In contrast, the incidence of LTU ranged only from 0.46 to 0.55. This result is consistent with the scatter plot of the unemployment spells data,  $X^c, X^u$ , in Figure 3. The scatter is tightly distributed around the line  $X^c = 0.5X^u$ , suggesting little variation in the incidence of LTU across TTWA, while the larger cities clearly have longer uncompleted unemployment spells.

Figure 8 presents a (weighted) scatter plot of the incidence of unemployment against expected duration. Two main features are evident; (a) the larger cities are correlated with high expected durations of LTU, and (b) cities with longer expected durations of unemployment have lower LTU incidence (the raw correlation being  $-0.278$ ). Using a log linear regression equation and weighted least squares, Table 2 describes how these variables are correlated with city size (population and geographical area) and region across England and Wales.

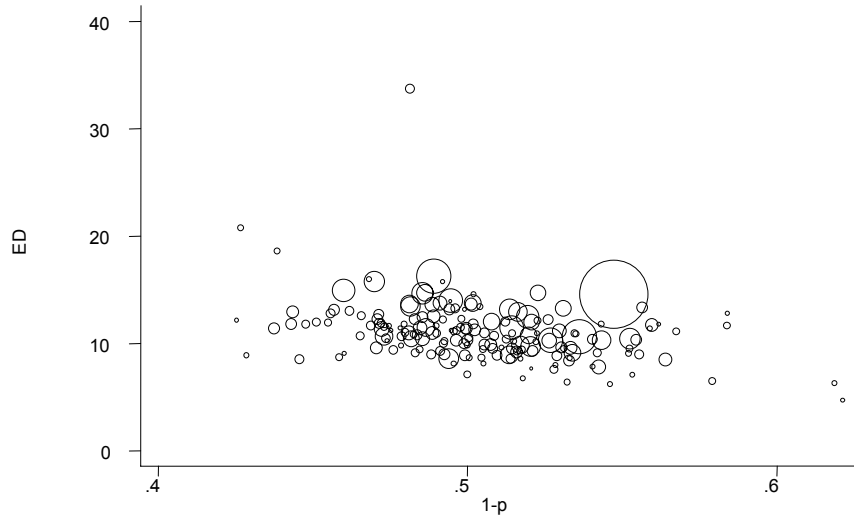


Figure 8: Expected Duration vs. Incidence by TTWA

Table 2  
Weighted Least Squares Estimates

Variable	$\log(1 - \bar{p}^i)$	$\log(\overline{ED}^i)$
Constant	-0.487* (.078)	1.143* (.228)
$\log(\text{Population})$	-0.011* (.004)	0.127* (.013)
$\log(\text{Area})$	0.003 (.007)	-0.036 (.020)
REGION		
East Anglia	-0.148 (.022)	0.194* (.066)
East Midlands	-0.113* (.019)	0.244* (.055)
North	-0.188* (.019)	0.325* (.056)
North West	-0.107* (.015)	0.158* (.047)
South East	-0.083* (.015)	0.026 (.044)
South West	-0.118* (.019)	0.188* (.056)
Wales	-0.124* (.021)	0.287* (.062)
West Midlands	-0.111* (.016)	0.240* (.046)
Yorkshire & H/side	-0.148* (.017)	0.262* (.050)
$\bar{r}^2$	0.40	0.47
$N$	254	

As suggested by the scatter plots, there are significant city size effects. The largest cities tend to have a slightly smaller incidence of LTU, but each LTU worker expects a much longer unemployment spells. A doubling of the city size leads to a 1% fall in incidence, and a 13% increase in the expected durations of unemployment. The overall effect implies that LTU is much more of a problem in the largest cities.<sup>16</sup>

There is also strong evidence of a North-South divide on LTU. The omitted region variable is the ‘London Area’. The regressions find that the incidence of LTU in the ‘North’ is lower but the expected durations of LTU are much higher. Evaluated at the means, expected durations on average rise by more than 25%, or approximately three months, in the eight regions outside southeastern England. In contrast, the average incidence of becoming LTU in these regions is 12% lower. This implies that in the ‘North’, a smaller minority of workers experience much longer spells of long-term unemployment.

## 4. Conclusion

This paper has shown how to decompose matching and unemployment data into measures of the incidence ( $1 - p$ ) and average matching rate ( $\lambda$ ) of the long-term unemployed. It has also identified estimates of the expected duration of long-term unemployment. The results are not only consistent with intuition - the incidence and expected duration of LTU is highest in the recession and lowest in the boom - they are also consistent with how the average completed and uncompleted spells of unemployment change over the cycle. The cross section results also identify

- (i) LTU is a bigger problem in the larger cities (higher expected durations) and
- (ii) a “North-South” regional divide, where the ‘North’ is characterised by a smaller minority of workers experiencing much longer spells of long-term unemployment

It is worth comparing these results against the standard statistics for LTU - the proportion of workers unemployed more than 6 or 12 months. Figure 9 plots those statistics for England and Wales.

Somewhat paradoxically, those measures of LTU take their lowest value in the middle of the recession (January, 1991). In fact those measures have the same

---

<sup>16</sup>Other regressions not reported here find that demographic, industrial and occupational characteristics of the TTWAs from the 1991 census do not correlate strongly with either risk or duration.

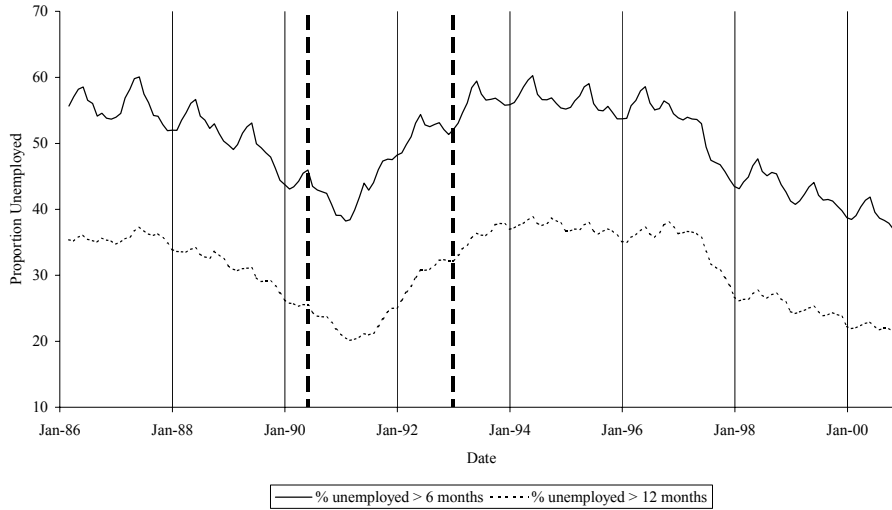


Figure 9: Standard Measures of Long-Term Unemployment

time series profile as the average completed and uncompleted unemployment spell statistics (Figure 3). As explained in the text, these measures fall at the onset of the recession as there is a large increase in newly unemployed workers (also see Machin and Manning, 1999). But such a time series profile does not capture the underlying LTU situation - Figure 6 demonstrates not only that the incidence of LTU peaks in the middle of the recession, but also  $\lambda$  collapses to its low value of 0.06 (and the average expected duration of LTU peaks at 15 months). The principal drawback of the standard measure of LTU is that it does not distinguish between workers who have currently short unemployment durations and workers who expect to be unemployed into the long-term. The approach outlined here instead identifies the matching parameters of interest - the incidence and average matching rates of the long-term unemployed.

An important direction for future research is to relax the assumption that workers on the short side of the market match arbitrarily quickly. A natural extension would be to assume search frictions bind for these workers - it takes time to locate and take up the most preferred job opportunity in the current vacancy stock. The results of Coles and Smith (1998) suggest this search process may take up to a month in time. The results presented here find that roughly half

of all newly unemployed workers are on the short side of their skill market. If instead such workers match, on average, in say two or three weeks (rather than immediately), then given our matching parameters fit the observed outflow rate  $M$ , our results overstate the average duration of long-term unemployment by a similar two or three weeks. This bias is small relative to the estimated durations of long-term unemployment.<sup>17</sup>

Such an extension would nevertheless provide a natural bridge between stock-flow matching and the standard search literature. It would also provide an alternative interpretation of the empirical search literature which assumes two types of job seekers - those that match quickly and those that match slowly (e.g. Lancaster and Nickell, 1980). The interpretation instead is that the former are on the short side of their skills market but matching frictions delay their immediate re-employment, whereas the latter are LTU and have to wait for something suitable to come onto the market.

## References

- [1] Acemoglu, D. and R. Shimer (1999), "Efficient Unemployment Insurance", *Journal of Political Economy*, 107:893-929.
- [2] Albrecht, J., P. Gautier and S. Vroman (2002), "Matching with Multiple Applications," *Economics Letters*, forthcoming.
- [3] Andrews M., S Bradley and R. Upward (2001) "Estimating the probability of a match using microeconomic data for the youth labour market", *Labour Economics*, 8:335-357
- [4] Budd A., P. Levine and P. Smith (1988), "Unemployment, vacancies and the long-term unemployed", *Economic Journal*, 98:1071-1091.
- [5] Burgess, S and S. Profit (2001), "Externalities in the matching of workers and firms in Britain", *Labour Economics*, 8:313-333.
- [6] Coles, M.G. (1999), "Turnover externalities with marketplace trading", *International Economics Review*.

---

<sup>17</sup>Simulations not reported here where, conditional on a match, the starting date for employment is exponentially distributed with mean equal to two weeks, suggest the estimates of  $p$  are slightly downward biased, while there is little effect on the estimates of  $\lambda$ .

- [7] Coles, M.G. and B. Petrongolo (2001), "A test between unemployment theories using matching data", mimeo, University of Essex.
- [8] Coles, M.G. and E. Smith (1998) "Marketplaces and matching", *International Economic Review*, 39:239-255.
- [9] Coles, M.G. and A. Muthoo (1998) "Strategic bargaining and competitive bidding in a dynamic market equilibrium" *Review of Economic Studies*, 65:235-260
- [10] Gregg, P. and B. Petrongolo (1997), "Random or non-random matching? Implications for the use of the UV curve as a measure of matching effectiveness", Institute for Economics and Statistics (Oxford) Discussion Paper No. 13.
- [11] Gregg, P. and J. Wadsworth (1996), "How effective are state employment agencies?", *Oxford Bulletin of Economics and Statistics*, 58:443-467.
- [12] 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.
- [13] Jackman, R., R. Layard and C. Pissarides (1989), "On vacancies", *Oxford Bulletin of Economics and Statistics*, 51:377-394.
- [14] Lancaster, T. and S. Nickell (1980), "The analysis of re-employment probabilities for the unemployed", *Journal of the Royal Statistical Society A*, 143, part 2: 141-152.
- [15] Layard R., S. Nickell and R. Jackman (1991), *Unemployment*, Oxford University Press, Oxford.
- [16] Lagos, R. (2000), "An alternative approach to search frictions", *Journal of Political Economy*, 108:851-73
- [17] Machin and Manning (1999), "The causes and consequences of long-term unemployment in Europe", in *Handbook of Labor Economics*, O. Ashenfelter and D. Card, eds, North- Holland, Amsterdam.
- [18] Meyer, B. (1990), "Unemployment insurance and unemployment spells", *Econometrica*, 58:757-782.

- [19] Montgomery J. (1991), "Equilibrium wage dispersion and interindustry wage differentials", *Quarterly Journal of Economics* 106:163-179
- [20] Petrongolo, B. and C. Pissarides (2001), "Looking into the black box: a survey of the matching function", *Journal of Economic Literature*, 39:390-431.
- [21] Portugal, P. and J. Addison (2000), "Short- and long-term unemployment", *Economic Letters*, 66:107-112.
- [22] Shapiro, C. and J. Stiglitz (1984), "Equilibrium unemployment as a worker discipline device", *American Economic Review*, 74:433-444
- [23] Taylor, C. (1995), "The long side of the market and the short end of the stick: bargaining power and price formation in buyers', sellers' and balanced markets," *Quarterly Journal of Economics*, 110:837-855.