MANCHESTER
1824

The University of Manchester

Economics Discussion Paper

EDP-0635

Why do job-seeker and vacancy hazards slope downwards? Estimating a two-sided search model of the labour market

By Martyn Andrews, Steve Bradley, Dave Stott and Richard Upward

November 2006

Economics
School of Social Sciences
The University of Manchester
Oxford Road
Manchester M13 9PL
United Kingdom

Why do job-seeker and vacancy hazards slope downwards? Estimating a two-sided search model of the labour market*

Martyn Andrews
University of Manchester

Steve Bradley
Lancaster University

Dave Stott
Lancaster University

Richard Upward
University of Nottingham

February 2006

Version submitted to WPEG Annual Conference, Kent 2006

*The authors thank The Economic and Social Research Council (under grant R000239142) for financial assistance. We are grateful to Barbara Petrongolo for helpful clarifying conversations. The comments of participants at various presentations are gratefully acknowledged. These include the 2003 Annual Conference of the European Association of Labour Economists (Seville), the RES (2003), and the Department of Economics at Manchester. The usual disclaimer applies. The data were kindly supplied by Lancashire Careers Service.

Abstract

In this paper we provide the first microeconometric estimates of the hazards to matching on *both* sides of a labour market, decomposed into their constituent parts. Namely, the rate at which job-seekers and vacancies contact each other, and the probability that these contacts result in a match. This allows us to determine whether it becomes harder for agents to match as time passes because they receive fewer contacts or because contacts are less likely to be successful.

In the raw data the decline in the matching rate is driven by a decline in the contact rate, and not by any fall in the probability of a match conditional on a contact. We estimate a two-sided matching model to determine whether this result is caused by omitted observed or unobserved heterogeneity in job-seekers and vacancies. It also allows us to estimate the parameters of the individual components of the matching function. We find that the same result applies as in the raw data: the decline in the matching rate on both sides of the market is driven by the decline in the contact rate.

Keywords: two-sided search, random matching, hazards
New JEL Classification: C41, E24, J41, J63, J64

7032 words

Address for Correspondence:

Dr. M.J. Andrews
School of Social Sciences
University of Manchester
Manchester, M13 9PL

Email: `martyn.andrews@manchester.ac.uk`

Filename: `h-lambda2-WPEG.tex`
Processed at: 09:16 Tuesday 28th February, 2006

1 Introduction

There is a large empirical literature which has estimated and tested various components of the search and matching framework. But as noted by Petrongolo & Pissarides (2005), "...aggregate matching functions and individual hazard rates conceal more than one structural dimension. They are both a composite of the mechanics of the meeting technology and the willingness of employers and job-seekers to accept the other side's offer." In this paper we directly estimate these components of the matching function. Uniquely, we use micro-level data to decompose both vacancy and job-seeker matching rates into the matching probability and the arrival rate of applicants from the other side of the same well-defined labour market.

Why is this decomposition of interest? The search and matching framework is becoming the dominant paradigm in explaining both micro- and macro- labour-market phenomena: see Mortensen & Pissarides' recent (1998, 1999) surveys. Estimating the individual components of the matching function allows us to shed light on and test various elements of this framework.

First, we can examine whether changes in the per-period matching probability (or hazard rate) for vacancy and job-seeker spells is determined by changes in the arrival rate of suitable partners, or changes in the matching probability. In the raw data both unemployment spells and vacancies tend to have declining hazard rates: is this because suitable partners become harder to find as time passes, or because the matching probability declines?

Second, Eckstein & van den Berg (2006) note that parameter estimates from a reduced-form hazard model "cannot separate between the relative magnitudes of the job offer arrival rate and the acceptance probability, or estimate the relative magnitudes of the effects of the x variables on them." Our estimates allow us to calculate precisely these magnitudes.

Third, the decomposition of the matching function allows us to test one of the crucial restrictions underlying the search and matching framework: that of *random matching*.[1] In the context of labour markets, when a job-seeker contacts a vacancy it is assumed that the vacancy is a random draw from the set of all vacancies.

An alternative assumption about the meeting technology is that job-seekers and vacancies do not meet randomly over time, but are able to use a marketplace to search the other side of the market. This assumption, referred to as *non-random*

[1] See, for example, Burdett & Coles (1999).

matching, leads to the *stock-flow matching* model.[2] To test this model formally, we require micro-level data which contains the identity of each pair who contact and match.

This paper is the first we are aware of to estimate hazards from both sides of the *same market* using microeconomic data, and which decomposes both hazards into their constituent parts.

In the next section, we present a stylised version of the two-sided random matching model. In Section 3, we describe the institutional background to the youth labour market in the UK in the late 1980s and then describe fully the information we observe in our dataset of contacts and matches. In Section 4 we set the two-sided search model of Section 2 in a stochastic environment, from which we develop the econometric methodology. In Section 5 we discuss our results. Section 6 concludes.

2 Theoretical framework

The model we outline here is a stylised version of the random matching model. There are stocks of vacancies V and job-seekers U (all of whom are assumed unemployed) attempting to meet and eventually form matched pairs. The rate at which they randomly contact each other per period is $\lambda(U, V)$, where $\lambda()$ has the same properties as a production function (concave and increasing in both arguments). If $\lambda(U, V)$ also exhibits constant returns to scale, the average number of contacts per vacancy is

$$\lambda^e(\theta) = \lambda/V = \lambda(U/V, 1)$$

and is decreasing in labour-market tightness $\theta \equiv V/U$. Similarly, the average number of contacts per job-seeker is

$$\lambda^w(\theta) = \lambda/U = \lambda(1, V/U)$$

and is increasing in θ. The corresponding hazards are:

$$h^e(\theta) = \lambda^e(\theta)\mu(\theta) \qquad h^w(\theta) = \lambda^w(\theta)\mu(\theta), \tag{1}$$

where μ is joint probability that a job-seeker finds an employer acceptable and an employer finds a job-seeker acceptable. In some two-sided search models $\mu(\theta)$ is

[2] See Coles & Smith (1998) and Coles & Petrongolo (2003).

an increasing function in slack markets and then becomes a decreasing function in tighter markets.

The aggregate matching (or hiring) function can be obtained by aggregating either hazard over the corresponding stock of market participants:

$$\delta(U,V) = Vh^e(\theta) = V\lambda^e(\theta)\mu(\theta) \tag{2}$$

$$= Uh^w(\theta) = U\lambda^w(\theta)\mu(\theta) = \lambda(U,V)\mu(\theta). \tag{3}$$

This shows how the matching function δ is decomposed into the contact function and the matching probability. It will exhibit constant returns to scale if $\lambda(\theta)$ does the same.

There is a large microeconometric literature that has estimated the hazard out of unemployment using unemployment duration data,[3] but there is far less evidence for vacancies.[4] Search in a stationary environment predicts that the hazard is constant, although most estimates show declining hazards. This is thought to be due to either some form of negative duration dependence or unmodelled unobserved heterogeneity. Assuming the latter can be controlled for using appropriate econometric techniques, negative duration dependence can arise either because the arrival rate of suitable offers falls or the matching probability falls or both, as seen in decomposing the hazard in (1) above.[5] Other microeconometric studies do not estimate either hazard directly. Some have estimated the hiring function $\delta(U,V)$ directly[6] or the matching probability[7] or better still, have decomposed the hiring function into λ and μ (see equation 3).[8] However, the great majority of empirical work on the hiring function has used aggregate time-series data.[9]

3 The data

The data we use are the computerised records of the Lancashire Careers Service (LCS) over the period March 1988 to June 1992. The Careers Service was a

[3]See van den Berg (1999, Footnote 1) for a list of contributions and surveys.

[4]See, for example, van Ours & Ridder (1991, 1992, 1993), Barron, Berger & Black (1997), Burdett & Cunningham (1998), and Russo & van Ommeren (1998), Andrews, Bradley & Upward (2003).

[5]See van Ours (1990) for vacancies and van den Berg (1990) for unemployment.

[6]See Lindeboom, van Ours & Renes (1994), Anderson & Burgess (2000), and Broersma & van Ours (1999).

[7]See Teyssière (1996) and Andrews, Bradley & Upward (2001).

[8]See van Ours & Lindeboom (1996).

[9]See Petrongolo & Pissarides (2001) for a comprehensive survey.

Government-funded network which provided vocational guidance for school-leavers and which operated a free matching service for employers and youths.

The data comprise a longitudinal record of all youths in Lancashire aged 15–18, including those in education, employment, training and unemployment. For each individual we observe the start and end dates of every labour market spell over the sample period. The data also include a record of all vacancies notified to the Careers Service over the sample period. Approximately 30% of all job spells observed in the data resulted from a match with a vacancy posted with the Careers Service. Vacancies for which the Careers Service were not the method of search are not included in the data. However, in the youth labour market (in contrast to the adult labour market) vacancies posted with the Careers Service were generally representative of all vacancies available for this age group.

Job-seekers can come from one of four labour market states: unemployment, employment, government-sponsored training or education. Each vacancy is filled by one of these types of job-seeker or it is withdrawn from the market, or it is censored. Job vacancies can either be filled via the Careers Service, or filled by some other means. Each job-seeker finds one of these types of vacancy, or she leaves the labour market and stops actively searching, or she is censored. A job-seeker who stops searching and leaves the labour market is the analogue of a vacancy which is withdrawn from the market.

We analyse contacts and matches between job vacancies and unemployed job-seekers. Matches involving school-leavers and those on training programmes are less relevant for the purpose of estimating models of labour-market matching. However, we do need to consider other types of job-seeker when specifying the arguments of the matching function, because it might be the case that the stock of those engaged in on-the-job search affects the probability of a match between unemployed job-seekers and vacancies because they are competing for the same vacancies. We therefore use two definitions of job-seekers. The first, *narrow* definition refers only to unemployed job-seekers. The second, *wide* definition includes those who are on training programmes and those who are in jobs, and who are registered as actively searching with the Careers Service. The narrow definition corresponds more closely to the existing literature.

Figure 1 illustrates the information we observe on the timing of spells, contacts and matches. The data are observed roughly every calendar month, and the duration of each interval is known to the day. Hereafter, we refer to each interval as a 'month', denoted τ. We emphasise that τ refers to elapsed duration rather than calendar

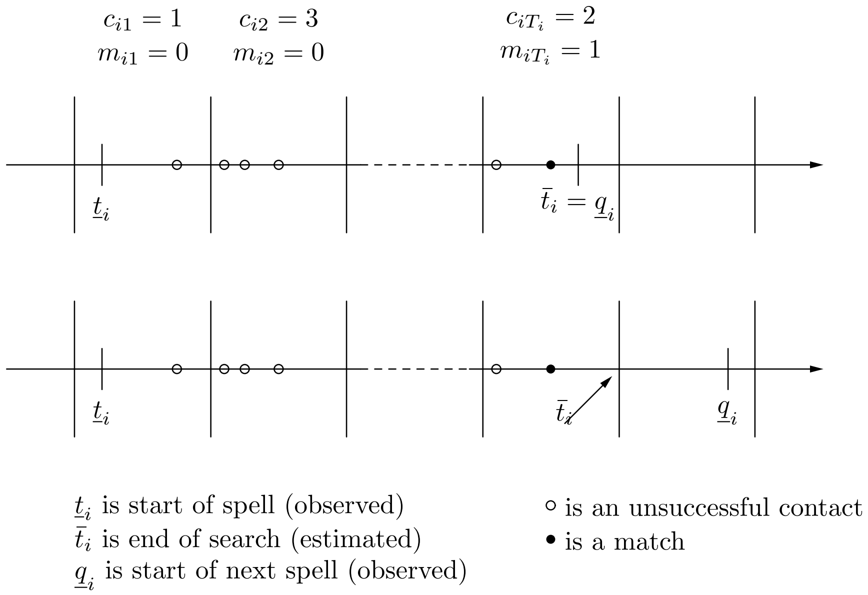


Figure 1: Contacts and matches: two representative job-seekers

time. We observe the day on which the job-seeker i became unemployed, $\underline{t}_i$ We also observe the number of contacts received by each job-seeker i for each month i was unemployed, $c_{i\tau}$, $\tau = 1, \ldots, T_i$, and the associated number of matches, $m_{i\tau}$. By definition, $m_{i\tau} = 0$ except for the last month, when $m_{iT_i} = 1$, and only if a match occurs (the spell can be censored by the end of the sample or the agents can exit to matches not analysed here).

We do not observe exactly when each contact took place (including the final successful contact), only the month in which it occurred. In our empirical work we therefore estimate hazards and matching probabilities as functions of elapsed duration measured to the nearest month, as is standard in discrete-time duration models. However, we do observe the start and end date of each spell dated to the nearest day, and we use this information to calculate more precise measures of time at risk within a month. The only inaccuracy here comes about because there is likely to be a gap between the date of the successful contact and the start date of the resulting job. The end date is taken to be either the last day of the month in which the contact takes place, $\bar{t}_i$, or the start date of the resulting job spell, $\underline{q}_i$, whichever occurs first. (In the figure, we illustrate both possibilities). This will always be an overestimate of time at risk within a month.

Total search duration for job-seeker i is given by $t_i = \sum_{\tau=1}^{T_i} t_{i\tau}$, where $t_{i\tau}$ is the time spent unemployed each month (roughly 30 days except for the first and last months). T_i is the integer number of months, wholly or partly, spent unemployed. Similarly, the total number of contacts is given by $c_i = \sum_\tau c_{i\tau}$ and the total number of matches is given by $m_i = \sum_\tau m_{i\tau}$ (again, $m_i = 1$ unless the spell is censored). Therefore we have a monthly unbalanced panel of observations for each job-seeker; the two variables being modelled are the number of contacts $c_{i\tau}$, $i = 1, \ldots, N^w$, $\tau = 1, \ldots, T_i$, and whether or not a given contact results in a match. Denote the total number of observations as $\bar{N}^w \equiv N^w \sum_i T_i$.

Exactly the same considerations apply to vacancies, except that all of T_j, $c_{j\tau}$, $m_{j\tau}$, c_j, m_j, and t_j are now indexed j for vacancy and s for elapsed duration. Consequently, we have a second monthly unbalanced panel of observations for each vacancy, now with $\bar{N}^e$ observations.

For both datasets, there are $25,267$ such contacts, resulting in $2,761$ matches. For each contact/match, we observe the following:

- the wide or narrow stock of job-seekers, U, and the stock of vacancies, V. Both vary by month through the duration of the job-seeker's spell of unemployment or employer's vacancy spell;
- a vector of job-seeker characteristics $\mathbf{x}^w$, and a vector of employer characteristics $\mathbf{x}^e$. Both are observed at the beginning of the spell;
- the wage ω on offer by the employer (not observed for all contactes/matches).

In fact, the stocks of unemployed job-seekers and vacancies do not vary by i or j, but by the labour market in which the job-seeker and employer are located. The data cover the whole of Lancashire, a county in the United Kingdom that comprises 14 towns/cities (in fact, local authority districts). When constructing the covariates, in fact we group Lancashire into just three labour markets (West, Central and East), recognising that job-seekers can travel between certain towns when looking for work. 96% of all matches take place between a job-seeker and vacancy from the same labour market. This number drops to 75% when Lancashire is treated as 14 towns/cities. It follows that there is very little cross-section variation in the data. Identification is achieved through the so-called recruitment cycle, which generates a lot of time-series variation in the data. There are very large peaks in unemployed stocks, arising from young people leaving school between May and August each year, which, of course is when employers post their vacancies. There is a similar annual variation in the data for vacancy stocks, but less pronounced.

4 Econometric issues

4.1 A basic statistical model

In this subsection, we set the two-sided search model of Section 2 in a stochastic environment. Recall that we have stocks of unmatched vacancies, denoted V, and unmatched job-seekers, denoted U, attempting to contact each other in a particular market. The standard assumption in this literature is that pairs are drawn randomly from U and V and meet each other according to a Poisson process. This is the contact or encounter function. Seen from the point of view of a single job-seeker, the number of vacancies he encounters every period, denoted C, is also Poisson distributed:

$$C \sim \text{Poisson}(\lambda^w t).$$

λ^w is the Poisson parameter, and denotes the average number of contacts per job-seeker each period, with exposure t. From a theoretical perspective, the length of the period tends to zero, and, in the limit, the probability of two or more contacts occurring also tends to zero. In practice, the length of the period is a day, with the observation interval t being roughly a month (the length of which varies from month to month) and the number of contacts is any non-negative integer.

Every time a vacancy is encountered, the pair either consummates the contact by matching with each other, or they do not. The probability that a given contact results in a match is denoted μ. The agents from unsuccessful contacts return to the stocks of U and V. Then the number of matches that result from C contacts is distributed as a Binomial:

$$M|C \sim \text{Binomial}(\mu, C).$$

Once a job-seeker has matched successfully, then there are no more matches. It follows that the marginal distribution for the number of matches per period has a Bernoulli distribution (see the Appendix):

$$M \sim \text{Bernoulli}(\lambda^w \mu t).$$

This means that $\Pr(m = 1) = \lambda^w \mu t$ in a month with exposure t. Intuitively, controlling for the length of the month, if λ^w vacancies are encountered on average each day and a proportion μ of them match on average, then the average daily matching rate of job-seekers is $\lambda^w \mu \equiv h^w$. Although C is Poisson distributed, M

has a binary outcome. Looking at marginal distribution for M is useful, because it shows that, if one had data on matches only, the average number of matches per day, with exposure t, is $h^w \equiv \mu\lambda^w$. Clearly, without data on contacts, one cannot separately identify μ and λ^w.

With a sample of $\bar{N}^w$ job-seeker observations on pairs of contacts and matches ($m_{i\tau}$,$c_{i\tau}$), using Equation (A.1), the log-likelihood of observing the sample is:

$$\ell_{m,c} = \sum_{i\tau}[c_{i\tau}\log(\lambda^w_{i\tau}t_{i\tau}) - \lambda^w_{i\tau}t_{i\tau}] + \text{const} \\ + \sum_{i\tau|c_{i\tau}>0}[m_{i\tau}\log\mu_{i\tau} + (1-m_{i\tau})\log(1-\mu_{i\tau})], \quad (4)$$

where

$$t_{i\tau}\lambda^w_{i\tau} = a^w_i \exp(\mathbf{x}_{i\tau}\boldsymbol{\beta}^{cw}) \quad \text{and} \quad (5)$$

$$\mu_{i\tau} = \epsilon^w_i[1-\exp(-\exp(\mathbf{x}_{i\tau}\boldsymbol{\beta}^w))]. \quad (6)$$

a^w_i is a time-invariant unobserved heterogeneity term. This way of writing Equation (5) means that exposure can be entered into the model easily as an extra covariate $\log t_{i\tau}$ with its parameter constrained at unity.

In passing, note that constraining $\lambda^w_{i\tau}t_{i\tau} = \lambda^w_i t_i$ and adding up over $\tau = 1,\ldots,T_i$ observations for each job-seeker, gives the same log-likelihood as Lancaster (1990, Eqn (2.14)). In his example, all that is observed is continuously measured duration/exposure for a given job-seeker, and the number of unacceptable contacts.

The log-likelihood given in Equation (4) factors in two separate terms. The first summation is a standard Poisson likelihood that uses all the data to estimate the parameters of $\lambda^w_{i\tau}$. The second summation is a standard Binary Choice likelihood that uses only job-seeker/months where there are one for more contacts to estimate the parameters of $\mu_{i\tau}$. In this second term, for each month the data on $c_{i\tau}$ and $m_{i\tau}$ can be expanded into a sequence of ones for $c_{i\tau}$ and a sequence of zeros ending with a final one for $m_{i\tau}$. Thus the data on $c_{i\tau}$ can be discarded, and the model estimated as a Binary Choice model with m as the dependent variable, using each contact as a separate observation.

Thus we estimate the parameters of the contact process separately from the parameters of the matching process. For the first, we have a panel dataset whose unit of observation is a month/job-seeker i,τ. For the second, we have a job-seeker panel whose unit of observation is a contact c. This also means that one could change the specification of the contact process to something more general, like a Negative Binomial, and still estimate the two models separately.

Suppose now that the sample comprises $\bar{N}^w$ job-seeker observations on matches $m_{i\tau}$ only. From the Marginal Distribution of M given in the Appendix (see Equation A.2), the log-likelihood of observing the sample is

$$\ell_m = \sum_{i\tau} [m_{i\tau} \log(h^w_{i\tau} t_{i\tau}) + (1 - m_{i\tau}) \log(h^w_{i\tau} t_{i\tau})], \tag{7}$$

where

$$t_{i\tau} h^w_{i\tau} = u^w_i [1 - \exp(-\exp(\mathbf{x}_{i\tau} \boldsymbol{\beta}^{mw}))]. \tag{8}$$

This is very similar to a discrete-time duration model, where one organises the data into sequential binary response form, using months as the unit of observation. For each job-seeker, every month $m_{i\tau} = 0$ except the last, where $m_{i\tau}$ can be unity. However, it is only contacts, successful or otherwise, that are observed monthly. Because the duration of the spell is observed to the nearest day, we can use this information in the above Binary Choice formulation, but not the sequential binary response form, where daily duration enters the likelihood via exposure $t_{i\tau}$. Also, it seems more natural to model matches as a binary outcome when contacts are observed as a count.

It is slightly unusual to use the complimentary log-log link for the matching model in Equation (8) and the model for μ in Equation (5). (The Canonical link for a binary outcome is the logistic link). There are two reasons. First, it means that exposure can be entered into both models easily. Second, because $h^w \equiv \mu \lambda^w$, it follows that

$$\boldsymbol{\beta}^{mw} \approx \boldsymbol{\beta}^{w} + \boldsymbol{\beta}^{cw}. \tag{9}$$

This means we can decompose the effect of any covariate that affects the exit hazard to see whether its effect is via the matching probability, or the arrival rate of agents from the other side of the market. For example, consider the effect of the stock of unemployment:

$$\frac{\partial \log h^w}{\partial \log U} \approx \frac{\partial \log \mu}{\partial \log U} + \frac{\partial \log \lambda^w}{\partial \log U}. \tag{10}$$

Do job-seekers searching in a labour market with a high stock of unemployed job-seekers have a lower unemployment hazard because they contact fewer vacancies or because there is a lower probability of success, once contacted? Similarly, thinking about elapsed unemployment duration, is the standard finding that unemployment hazards fall due to falling arrival rate of applicants, falling matching probabilities, or both?

All of the above is repeated for employers, where we seek to estimate the parameter vectors $\boldsymbol{\beta}^{me}$, $\boldsymbol{\beta}^{e}$ and $\boldsymbol{\beta}^{ce}$ using a panel dataset of match/contact pairs $(m_{j\tau}, c_{j\tau})$. The panel of contacts is the same dataset as for job-seekers, and so there we use three datasets in what follows.

4.2 Non-parametric estimates of matching and contact hazards

For each contact, we observe the current duration of unemployment for the job-seeker, τ^w, and the time the vacancy has been open, τ^e, where τ is measured in 'months'. Because we have no *a priori* view about the shape of the hazard of arrival rate of job-seekers λ^w and employers λ^e, or matches h^w and h^e, or the matching probability μ, we model their shapes non-parametrically. Specifically, we define 12 dummy variables for each 'month' the spell has lasted:

$$1(\tau = 1), \ldots, 1(\tau = 10),\ 1(\tau = 11, 12), \text{ and } 1(\tau \geq 13)$$

for both sides of the market.

Writing the raw hazards as $\bar{\lambda}^w_\tau$, $\bar{h}^w_\tau$ and $\bar{\mu}^w_\tau$, and writing Equations (5) and (8) without covariates, it is easy to show that ML estimates of these raw hazards are given by:

$$\widehat{\bar{\lambda}}^w_\tau = \frac{\sum_i c_{i\tau}}{\sum_i t_{i\tau}} \qquad \widehat{\bar{\mu}}^w_\tau = \frac{\sum_i m_{i\tau}}{\sum_i c_{i\tau}} \qquad \widehat{\bar{h}}^w_\tau = \frac{\sum_i m_{i\tau}}{\sum_i t_{i\tau}}, \tag{11}$$

where $\widehat{\bar{h}}^w_\tau = \widehat{\bar{\mu}}^w_\tau \widehat{\bar{\lambda}}^w_\tau$. Exactly the same expressions occur for employers:

$$\widehat{\bar{\lambda}}^e_\tau = \frac{\sum_j c_{j\tau}}{\sum_j t_{j\tau}} \qquad \widehat{\bar{\mu}}^e_\tau = \frac{\sum_j m_{j\tau}}{\sum_j c_{j\tau}} \qquad \widehat{\bar{h}}^e_\tau = \frac{\sum_j m_{j\tau}}{\sum_j t_{j\tau}},$$

where $\widehat{\bar{h}}^e_\tau = \widehat{\bar{\mu}}^e_\tau \widehat{\bar{\lambda}}^e_\tau$. Notice that the expression for $\widehat{\bar{\mu}}^w_\tau$ is not the same as for $\widehat{\bar{\mu}}^e_\tau$ for two reasons. The first is because the rate at which job-seekers arrive at employers is not the same as the rate at which employers arrive at job-seekers. Second, the matching probabilities will differ even though a job-seeker might have been unemployed the same length of time that an employer has been looking for a job-seeker.

4.3 Covariates

In Section 3, we note that we have access to a vector of job-seeker characteristics $\mathbf{x}^w$, and a vector of employer characteristics $\mathbf{x}^e$, and the wage ω on offer by the employer. The five models we estimate are written as follows.

$$\begin{aligned}
\lambda^w_{i\tau} &= \bar{\lambda}^w_\tau a^w_i \lambda^w(U_{i\tau}, V_{i\tau}, \mathbf{x}^w_i) \\
h^w_{i\tau} &= \bar{h}^w_\tau u^w_i h^w(U_{i\tau}, V_{i\tau}, \mathbf{x}^w_i) \\
\mu_c &= \bar{\mu}^w_{\tau(c)} \bar{\mu}^e_{s(c)} \epsilon^w_{i(c)} \epsilon^e_{j(c)} \mu(U_c, V_c, \mathbf{x}^w_{i(c)}, \mathbf{x}^e_{j(c)}) \\
\lambda^e_{js} &= \bar{\lambda}^e_s a^e_j \lambda^e(U_{js}, V_{js}, \mathbf{x}^e_j) \\
h^e_{js} &= \bar{h}^e_s u^e_j h^e(U_{js}, V_{js}, \mathbf{x}^e_j)
\end{aligned}$$

It is standard to parameterise the contact function, the matching probability and the matching function and as Cobb-Douglas, which means that U and V enter all five functions in logs as follows:

$$\log \lambda^w = (\alpha_1 - 1)\log U_{i\tau} + \beta_1 \log V_{i\tau} + \mathbf{x}^w_i \boldsymbol{\beta}^{cw} + \log t_{i\tau} \tag{12}$$

$$\log(-\log(1 - h^w)) = (\alpha_2 - 1)\log U_{i\tau} + \beta_2 \log V_{i\tau} + \mathbf{x}^w_i \boldsymbol{\beta}^{mw} + \log t_{i\tau} \tag{13}$$

$$\log(-\log(1 - \mu)) = \alpha_3 \log U_c + \beta_3 \log V_c + \mathbf{x}^w_{i(c)} \boldsymbol{\beta}^w + \mathbf{x}^e_{j(c)} \boldsymbol{\beta}^e \tag{14}$$

$$\log \lambda^e = \alpha_1 \log U_{js} + (\beta_1 - 1)\log V_{js} + \mathbf{x}^e_j \boldsymbol{\beta}^{ce} + \log t_{js} \tag{15}$$

$$\log(-\log(1 - h^e)) = \alpha_2 \log U_{js} + (\beta_2 - 1)\log V_{js} + \mathbf{x}^e_j \boldsymbol{\beta}^{me} + \log t_{js} \tag{16}$$

Equations (12) and (13) replace Equations (5) and (8) for job-seekers, estimated using the job-seeker panel $(c_{i\tau}, m_{i\tau})$. Equations (15) and (16) are equivalent equations for vacancies, estimated using the job-seeker panel (c_{js}, m_{js}). Equation (14) replaces Equation (6), and refers to both job-seekers and vacancies, and is estimated using the panel of contacts m_c. Notice that this model for μ involves two sets of duration dummies and two random effects, leading to a multi-level model. Also, given the discussion at the end of Section 4.1, we expect $\alpha_2 \approx \alpha_1 + \alpha_3$, $\beta_2 \approx \beta_1 + \beta_3$; more importantly, we would hope that the estimates of α_1 and β_1 from (12) are similar to those from (15), and that the estimates of α_2 and β_2 from (13) are similar to those from (16).

4.4 Econometric techniques

If one ignores the unobserved heterogeneity terms a_i^w, ϵ_i^w, u_i^w, a_j^e, ϵ_j^e, and u_j^e, then we refer to the models outlined above as *Pooled Models*. They comprise two Poisson models for contacts, two Bernoulli models for matches, and one more Bernoulli model for whether or not a given contact results in a match. When we control for unobserved heterogeneity, we assume that all six random effects are Normally distributed with variances $\sigma^2_{a,w}$, $\sigma^2_{\epsilon,w}$, $\sigma^2_{u,w}$, $\sigma^2_{a,e}$, $\sigma^2_{\epsilon,e}$, and $\sigma^2_{u,e}$ respectively. The Bernoulli model for whether or not a given contact results in a match has two random effects, ϵ_i^w and ϵ_j^e. These panel-data techniques are well-established and we therefore do not provide further details (but see, in particular, Cameron & Trivedi (1998, chapter 9)). We refer to these as *Random Effects Models*.

5 Results

5.1 Raw data, hazards and matching probabilities

Table 1 describes all the raw data needed to estimate h, λ and μ for both sides of the market. As explained in Section 3, the data are effectively a monthly panel, with monthly duration index τ, shown in column (1).

Consider first the totals over all durations in both panels (a) and (b). There are $25,267$ contacts (column 3) of which $2,761$ (column 2) result in matches. The overall matching probability is given by $\widehat{\mu} = \sum_{i\tau} m_{i\tau} / \sum_{i\tau} c_{i\tau}$, and is therefore 0.109 (column 7). These are the same matches and contacts seen from both sides of the market, and therefore the same numbers appear in the *Total* row in both panel (a) and panel (b). Similarly, using $\widehat{h}^w = \sum_{i\tau} m_{i\tau} / \sum_{i\tau} t_{i\tau}$, dividing $2,761$ matches by $3,235,810$ days at risk for job-seekers (column 4), gives an average hazard of 0.00085. Note that this is a daily hazard rate, because we are dividing by the number of days at risk. The corresponding hazard for employers is 0.00301, nearly 4 times higher, because the total days at risk are correspondingly lower. The ratio of the days at risk is an estimate of labour-market tightness, θ, therefore. Finally, using $\widehat{\lambda}^w = \sum_{i\tau} c_{i\tau} / \sum_{i\tau} t_{i\tau}$, the average contact rates for job-seekers is $\widehat{\lambda}^w = 0.00781$. For vacancies, we have $\widehat{\lambda}^e = 0.02759$. These are again in the same 1:4 ratio, and both are about 10 times higher than the corresponding hazard rates, because $\widehat{\mu} = 0.109$.

A similar analysis applies to each row of the table, except that $\widehat{\mu}_\tau^w \neq \widehat{\mu}_\tau^e$ for reasons given above. For example, column (3) of panel (a) shows that job-seekers receive

Table 1: Raw data, hazards and matching probabilities

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				(a) Job-Seekers				
τ	$\sum_i m_{i\tau}$	$\sum_i c_{i\tau}$	$\sum_i t_{i\tau}$	$\hat{\hat{h}}^w_\tau$	$\hat{\hat{\lambda}}^w_\tau$	$\hat{\hat{\mu}}^w_\tau$	r_τ	$\frac{\sum_i t_{i\tau}}{r_\tau}$
1	684	10783	613051	0.00112	0.01759	0.06343	34657	17.69
2	824	5540	759901	0.00108	0.00729	0.14874	29176	26.05
3	484	3337	544566	0.00089	0.00613	0.14504	19617	27.76
4	355	2145	394597	0.00090	0.00544	0.16550	13342	29.58
5	155	1160	225311	0.00069	0.00515	0.13362	7944	28.36
6	97	709	164508	0.00059	0.00431	0.13681	5715	28.79
7	51	438	124784	0.00041	0.00351	0.11644	4141	30.13
8	44	317	96664	0.00046	0.00328	0.13880	3107	31.11
9	23	210	70104	0.00033	0.00300	0.10952	2363	29.67
10	11	168	52715	0.00021	0.00319	0.06548	1813	29.08
11-12	16	212	76362	0.00021	0.00278	0.07547	2424	31.50
13+	17	248	113247	0.00015	0.00219	0.06855	3688	30.71
Total	2761	25267	3235810	0.00085	0.00781	0.10927	127987	25.28
				(b) Vacancies				
s	$\sum_j m_{js}$	$\sum_j c_{js}$	$\sum_j t_{js}$	$\hat{\hat{h}}^w_s$	$\hat{\hat{\lambda}}^w_s$	$\hat{\hat{\mu}}^w_s$	r_s	$\frac{\sum_j t_{js}}{\sum_j r_s}$
1	1895	16334	191109	0.00992	0.08547	0.11602	14154	13.50
2	470	4855	190954	0.00246	0.02542	0.09681	7584	25.18
3	108	1306	117603	0.00092	0.01111	0.08270	4179	28.14
4	70	993	90781	0.00077	0.01094	0.07049	3056	29.71
5	63	487	62467	0.00101	0.00780	0.12936	2159	28.93
6	35	315	46396	0.00075	0.00679	0.11111	1697	27.34
7	23	295	39594	0.00058	0.00745	0.07797	1316	30.09
8	27	201	34673	0.00078	0.00580	0.13433	1070	32.40
9	25	138	27561	0.00091	0.00501	0.18116	878	31.39
10	10	73	18868	0.00053	0.00387	0.13699	645	29.25
11-12	17	85	30602	0.00056	0.00278	0.20000	988	30.97
13+	18	185	65170	0.00028	0.00284	0.09730	2036	32.01
Total	2761	25267	915778	0.00301	0.02759	0.10927	39762	23.03

[a] r_τ is the number of job-seekers at risk after τ months searching. Mathematically, $r_\tau = \sum_i m_{i\tau} + \sum_i (1 - m_{i\tau}) = \sum_i 1_{i\tau}$.

$10,783$ contacts in the first month of their search duration, and column (2) shows that 684 of those contacts result in a match. Column (4) shows that total days at risk in this first month are $613,051$.[10] Our estimate of the job-seeker hazard to a successful contact is given in column (5), and comes from Equation (11).

$$\hat{h}_1^w = \frac{\sum_i m_{i1}}{\sum_i t_{i1}} = 0.00112$$

Similarly, the contact function reported in column (6) is estimated from

$$\hat{\lambda}_1^w = \frac{\sum_i c_{i1}}{\sum_i t_{i1}} = 0.01759,$$

and the matching probability in column (7) is

$$\hat{\mu}_1^w = \frac{\sum_i m_{i1}}{\sum_i c_{i1}} = 0.06343.$$

It is clear from the above that one can also calculate $\hat{h}_\tau^w$ as the product of $\hat{\lambda}_\tau^w$ and $\hat{\mu}_\tau^w$.

Column (8) reports the total number of job-seekers and vacancies "at risk" in each duration category. There are $34,657$ job-seekers and $14,154$ vacancies in total, all of whom are at risk in month 1. Finally, column (9) is an estimate of the total number of days at risk per job-seeker or vacancy in that duration category. The $34,657$ job-seekers at risk in month 1 were at risk for 17.69 days. As noted already, this is an over-estimate because we overestimate column (4).

In Figure 2 we plot the estimates of h, λ and μ for both sides of the market from Table 1. Panels (a) and (b) show that the hazard to a successful contact is declining with duration for both job-seekers and vacancies. This is the standard result in most of the microeconometric literature, although these are the first results we are aware of which show both sides of the *same* market. The most obvious difference between panels (a) and (b) is the extent to which the hazard falls after one month, and the overall level of the hazard. The data come from a slack labour market ($\theta < 1$) and therefore vacancies exit faster than job-seekers. In addition, the exit hazard for vacancies collapses after the first month, whereas that for jobs-seekers declines only slowly.

Panels (c)–(f) decompose the job-seeker and vacancy hazard into the contact hazard and the matching probability. It is immediately clear, that on both sides of the

[10] Note that this is an overestimate, as explained in Section 3.

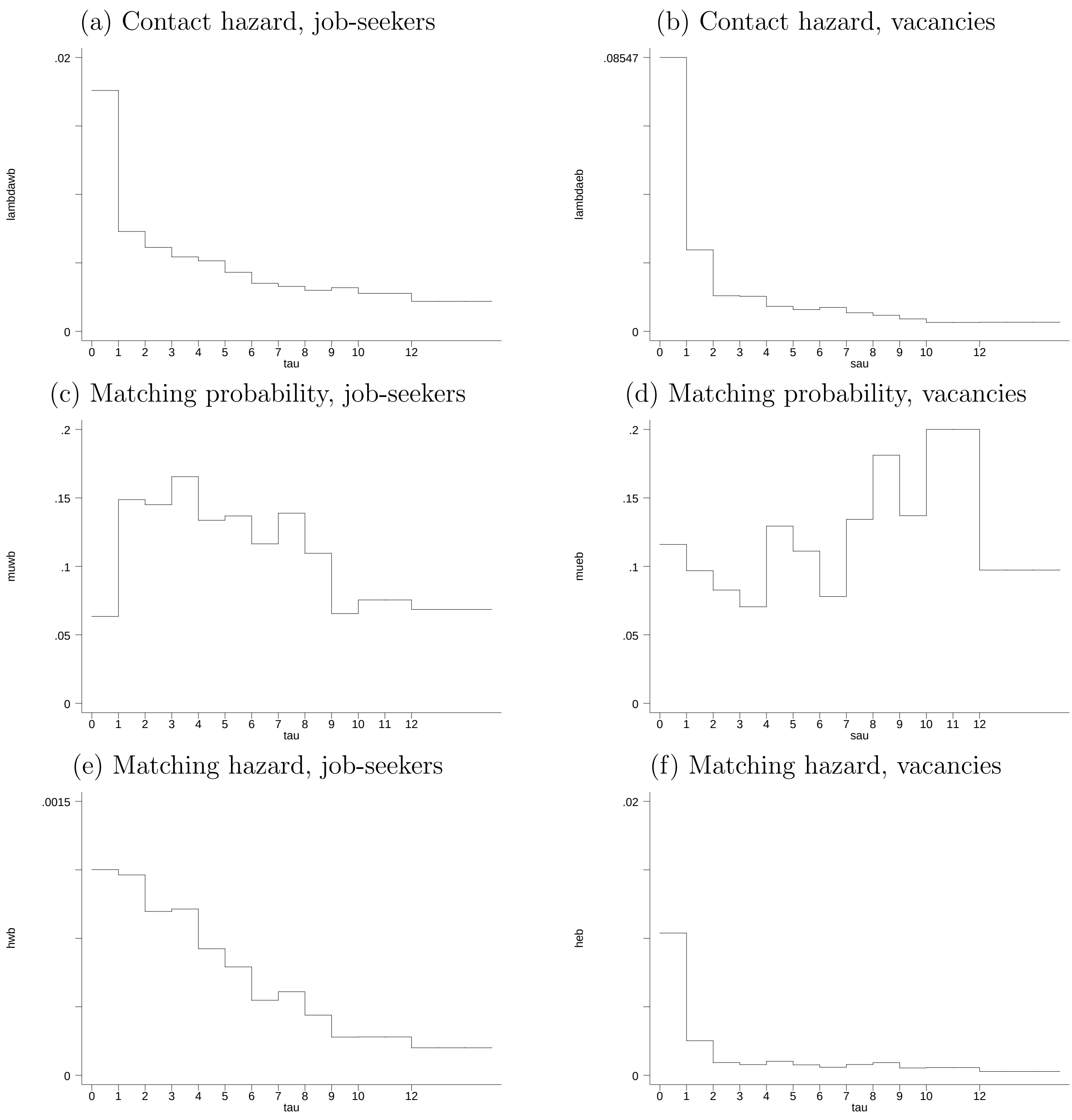


Figure 2: Raw hazards and matching probabilities

market, the declining matching hazard is due *entirely* to a decline in the number of contacts. Indeed, the matching probability actually increases with duration both for job-seekers and vacancies over a certain range. This result is consistent with the idea that a fall in either the job-seeker's or the employer's reservation utility offsets the decline in the contact rate.

We should also note that the large fall in the contact rate observed on both sides of the market is consistent with the non-random or stock-flow matching model. If there is a marketplace in which job-seekers and vacancies can contact each other quickly, we would expect very high initial contact rates. But once this initial period is over, agents only contact new entrants on the other side of the market, and the contact rate therefore falls sharply.

Of course, these contact hazards are also consistent with the random matching model if one assumes either (a) that there is considerable heterogeneity in job-seekers and vacancies, leading to spurious duration dependence or (b) if there is genuine duration dependence as a result of the "quality" of job-seekers and vacancies declining with duration. The latter seems unlikely, however, over such a short period, and in particular we do not think that vacancies' characteristics change with duration. However, the former is a real possibility, because we have not controlled for any observed or unobserved differences of the job-seekers and vacancies. To deal with this issue we now estimate the matching hazard, the contact hazard and the matching probability using the econometric methods outlined in Section 4.4.

5.2 Estimates of random matching models

Table 2 reports estimates of Equations (12) to (16) across four different specifications. Panel (a) is the Pooled Model, which is reported for comparison with estimates of the matching function from aggregated data. The estimates in panels (b) and (c) also allow for Normally distributed unobserved heterogeneity; in (b) we use the narrow definition of the stocks whereas in (c) we use the wide definition. Under constant returns in the matching function we would expect $\alpha + \beta = 1$ in the matching and contact functions, and $\alpha + \beta = 0$ in the matching probability.

Table 2: Estimated hazards and matching probabilities for job-seekers and vacancies, random matching models[a]

	Job-seekers				*Contacts*[c]		*Vacancies*				unlogged
	λ^w		h^w		μ		λ^e		h^e		mean
(a) Pooled Model, Narrow Stocks											
$\log U$	0.012	(0.044)	−0.147	(0.026)	−0.126	(0.044)	0.427	(0.015)	0.423	(0.057)	7xx
$\log V$	0.461	(0.072)	0.413	(0.064)	−0.115	(0.073)	−0.064	(0.051)	−0.221	(0.053)	2xx
$\alpha + \beta$	1.473	(0.029)	1.266	(0.044)	−0.241	(0.034)	1.363	(0.037)	1.202	(0.024)	
Log likelihood	−74055.2		−14306.0		−8410.4		−56323.8		−11059.7		
(b) Random Effects Model, Narrow Stocks[b]											
$\log U$	0.025	(0.009)	−0.151	(0.021)	−0.046	(0.030)	0.257	(0.013)	0.584	(0.041)	7xx
$\log V$	0.353	(0.014)	0.427	(0.029)	−0.161	(0.040)	0.144	(0.022)	−0.120	(0.063)	2xx
$\alpha + \beta$	1.378	(0.013)	1.276	(0.028)	−0.207	(0.040)	1.401	(0.020)	1.465	(0.061)	
Unobs het (σ)	1.329	(0.015)	0.783	(0.073)	1.299	(0.047)	1.286	(0.015)	3.102	(0.064)	
Log likelihood	−66710.0		−14286.8		−7970.3		−41350.6		−9893.1		
(c) Random Effects Model, Wide Stocks											
$\log U$	0.072	(0.009)	−0.126	(0.020)	−0.099	(0.029)	0.221	(0.013)	0.453	(0.042)	7xx
$\log V$	0.309	(0.015)	0.432	(0.031)	−0.119	(0.042)	0.149	(0.023)	−0.080	(0.065)	2xx
$\alpha + \beta$	1.381	(0.013)	1.306	(0.027)	−0.219	(0.038)	1.370	(0.020)	1.373	(0.060)	
Unobs het (σ)	1.331	(0.015)	0.787	(0.073)	1.289	(0.047)	1.299	(0.015)	3.135	(0.064)	
Log likelihood	−66683.3		−14293.5		−7965.7		−41414.9		−9943.9		
Observations	127987		127987		25267		39762		39762		3*T

[a] Estimates based on 25267 contacts and 2761 matches between 34657 unemployed job-seeker spells (26113 job-seekers) and 14154 Careers Service job vacancies. All standard errors adjusted for clustering on 3 labour-markets (see text), except for RE models.

[b] Normally distributed random effects. Adaptive quadrature used throughout, with numbers of quadrature points as follows: $\lambda^w = 24$, $h^w = 12$, $\mu = 16$, $\lambda^e = 32$, $h^e = 36$.

[c] This is *not* double RE, ie includes $\epsilon^e_{j(c)}$, but drops $\epsilon^w_{i(c)}$.

Our estimates of α and β generally show a slight but significant degree of increasing returns to scale ($\alpha + \beta > 1$) in the matching function. The estimates in the h column are approximately equal to the sum of the estimates in the λ and μ columns — see Equation (9), and so the novelty of our results is that we can decompose the effect of unemployment, for example, across the contact function and the matching probability (Equation (10)). So, for example, in panel (a) the elasticity of h^w with respect to unemployment is -0.147, and so this effect is not being driven by the elasticity of U with respect to contacts (0.012), but rather by the probability of a match (-0.126). This suggests that the negative impact of high unemployment on unemployment durations does *not* arise because of congestion or competition between job-seekers in terms of contact rates. Rather, it arises because, once a contact is made, the probability of a match is significantly lower. Since the matching probability is the joint probability that the both the job-seeker and the employer find that other side's offer acceptable, this is strong evidence that it is the employer effect which dominates. That is, as unemployment increases employers become more selective. However, we can do not see the same effects on the other side of the same market. The elasticity of h^e with respect to U is 0.423 and this is be driven again by the increase in the number of contacts (0.427) rather than the effect on the matching probability. We expect μ to be homogeneous of degree zero, so that $\alpha + \beta = 0$, but this is strongly rejected.

A potential inconsistency in these estimates arises because of unobserved heterogeneity. Job-seekers and vacancies differ in their "quality", which affects the rate of which they contact potential matches a_i^w, a_j^e and the probability of matching ϵ_i^w, ϵ_i^w, in ways that are unobserved in the data. This leads to overdispersion in the data. It should be stressed that this will only lead to inconsistent estimates of α and β if these heterogeneity terms are correlated with the other right hand side variables. In this case the right hand side variables are the stocks of job-seekers and vacancies in each district month, and it is not obvious that these will be correlated with unobserved characteristics of individual job-seekers and vacancies. Of more importance is the fact that we wish to estimate the relationship between h, λ, μ and elapsed duration. As is well known, failure to control for heterogeneity when estimating baseline hazards may lead to spurious duration dependence.

The panel nature of our data allows us to control for unobserved heterogeneity, reported in panel (b) of Table 2. It turns out that estimates of α and β on the job-seeker's side of the market are quite robust to unobserved heterogeneity at the individual level. This is consistent with the idea that U and V are essentially exogenous to the characteristics of individual job-seekers. On the vacancy side of the market, however, models with unobserved heterogeneity produce different estimates of α and β. For example, the elasticity of λ^e with respect to the stock of vacancies moves from -0.064 (pooled model) to 0.144 (random effects) and for h^e from -0.221 to -0.120. Similarly, the estimates with respect to the stock of unemployed move from 0.427 to 0.257 for λ^e and from 0.423 to 0.584 for λ^e This implies that the unobserved quality of the vacancies in our data are correlated with the stocks of V in each district-month. Also notice that the variance of the heterogeneity is much bigger for the vacancy hazard 3.102 than for the job-seeker hazard 0.783.

Panel (c) re-estimates (b), but using wide rather than narrow stocks. The estimates hardly move, and so, in what follows, we use the latter, since they are the more plausible measures of stocks for this particular (youth) labour market.

What effect does controlling for U, V and unobserved heterogeneity have on the estimates of the baseline hazards? Figure 3 plots the estimated baseline hazard from panel (c) in Table 2.[11] The baseline hazards remain very similar to those plotted from the raw data (Figure 2), although the exit rate for vacancies is considerably higher. The basic story remains very clear: declining exit rates for both job-seekers and vacancies is a result solely of declining contact rates. If we believe that we have controlled successfully for the heterogeneity of job-seekers and vacancies, this is a genuine effect of duration dependence.

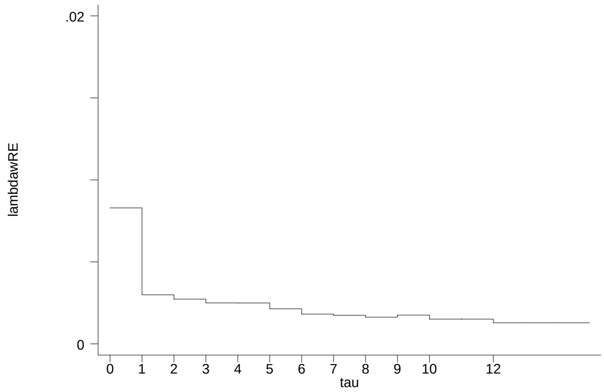
5.3 Estimates of non-random matching models

In Andrews, Bradley, Stott & Upward (2003) we develop a statistical test for the hypothesis that the data are generated by the stock-flow matching model, rather than the random matching model. Here, we give only the intuition behind the method.[12] We first separate the data between job-seekers and vacancies which are "old" and those which are "new". By "old" we mean having been in the market for more than one month. This notion of "old" and "new" corresponds quite closely to the theoretical notion of the "stock" and the "flow".

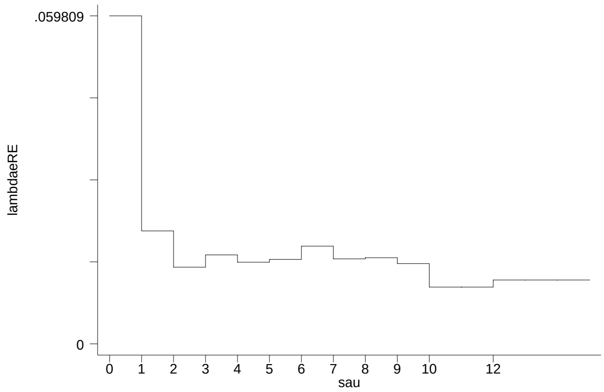
The argument is then quite simple. New vacancies and job-seekers can match with *any* agent from the other side of the market, because by definition they will not previously have been sampled. In contrast, once the job-seeker or vacancy becomes

[11] All hazards are plotted using the mean values of U and V.

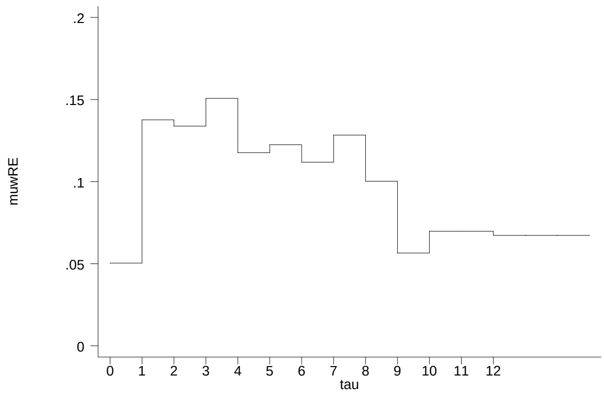
[12] For a fuller exposition, see Andrews, Bradley, Stott & Upward (2003), in particular Section 3.



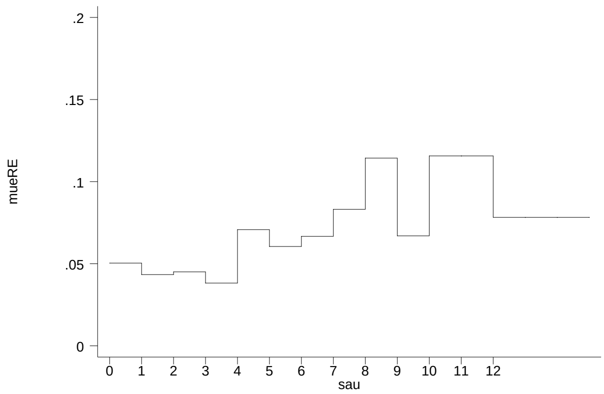
(a) Contact hazard, job-seekers



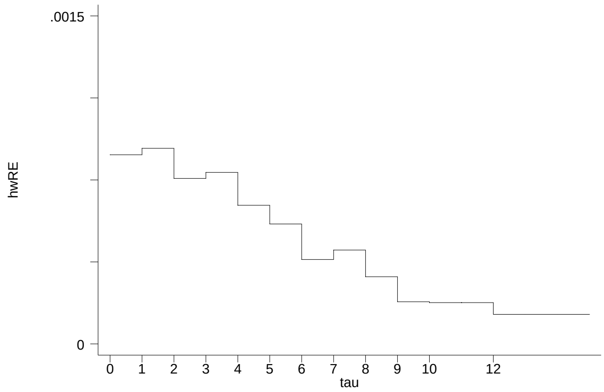
(b) Contact hazard, vacancies



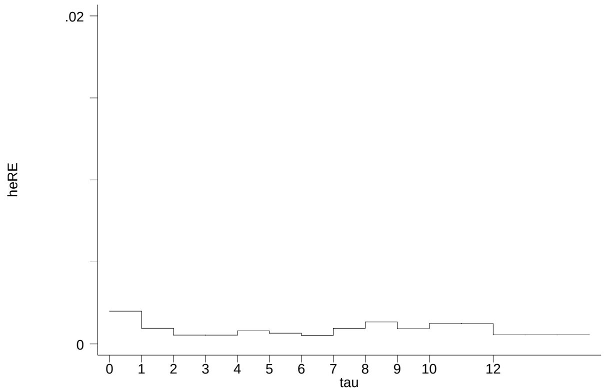
(c) Matching probability, job-seekers



(d) Matching probability, vacancies



(e) Matching hazard, job-seekers



(f) Matching hazard, vacancies

Figure 3: Hazards and matching probabilities, random matching model

old they can only match with “new” agents from the other side of the market, because all other potential matches must have been already rejected.

The empirical implication of this is threefold. First, either the contact rate or the probability of a match should reduce sharply once the initial period is over. We have already seen that this is the case for the contact function both in the raw data (Figure 2) and the econometric estimates (Figure 3). Second, an estimate of the matching function should show that for old job-seekers the stock of new vacancies v is a significant factor in determining the contact and exit rate, over and above the total stock of all vacancies, V. Similarly, the stock of new job-seekers u should be significant over and above the total stock of all job-seekers U. Third, the collapse in the matching rate should be a result of the collapse in the numbers of suitable partners on the other side of the market as the initial period of search ends.

Table 3: Hazards and matching probabilities for job-seekers and vacancies, stock-flow models, random-effects[a,b]

	Job-seekers						*Vacancies*						
	h^w		λ^w		μ^w		h^e		λ^e		μ^e		mean
(a) New (in market for one month or less)													
$\log u$	0.445	(0.070)	0.002	(0.024)	0.431	(0.094)	0.103	(0.056)	−0.040	(0.026)	0.153	(0.044)	192
$\log U$	−0.405	(0.068)	0.131	(0.022)	−0.500	(0.099)	0.366	(0.057)	0.279	(0.025)	−0.085	(0.087)	759
$\log v$	−0.450	(0.094)	−0.261	(0.029)	−0.156	(0.283)	−0.217	(0.070)	−0.025	(0.030)	0.025	(0.122)	58
$\log V$	0.771	(0.079)	0.455	(0.025)	0.194	(0.236)	−0.004	(0.063)	0.193	(0.029)	−0.260	(0.159)	216
α, β	1.040, 0.321		1.132, 0.194		−0.069, 0.038		0.469, 0.779		0.320, 1.168		0.068, −0.235		
$\alpha + \beta$	1.361	(0.060)	1.326	(0.019)	−0.031	(0.042)	1.248	(0.049)	1.488	(0.022)	−0.166	(0.009)	
(b) Old (in market for more than one month)													
$\log u$	−0.106	(0.041)	−0.216	(0.017)	0.214	(0.048)	0.193	(0.064)	0.064	(0.024)	0.321	(0.038)	192
$\log U$	−0.273	(0.056)	0.102	(0.027)	−0.475	(0.042)	0.300	(0.093)	0.090	(0.036)	−0.484	(0.137)	759
$\log v$	0.131	(0.055)	0.069	(0.022)	0.168	(0.030)	−0.339	(0.091)	−0.178	(0.031)	0.151	(0.057)	58
$\log V$	0.285	(0.050)	0.368	(0.024)	−0.314	(0.090)	0.301	(0.096)	0.391	(0.036)	−0.006	(0.069)	216
α, β	0.622, 0.416		0.886, 0.437		−0.261, −0.146		0.493, 0.963		0.155, 1.213		−0.163, 0.144		
$\alpha + \beta$	1.038	(0.047)	1.323	(0.023)	−0.407	(0.087)	1.455	(0.080)	1.368	(0.032)	−0.018	(0.119)	
Variance (σ^2)[c]	0.251	(0.099)	1.841	(0.037)			2.122	(0.136)	1.408	(0.029)			
Log likelihood	−14274.6		−66813.3		−8375.6		−10660.7		−41196.2		−8621.3		
Observations	127987		127987		25267		39762		39762		25267		

[a] Estimates based on 25267 contacts (10783 to new unemployed and 14484 to old unemployed, 16334 to new vacancies and 8933 to old vacancies) and 2761 matches (684 to new unemployed and 2077 to old unemployed, 1895 to new vacancies and 866 to old vacancies) between 34657 unemployed job-seeker spells (26113 job-seekers) and 14154 Careers Service job vacancies.

[b] Estimates of μ^w and μ^e are pooled, estimates of all other parameters are random-effects.

[c] Gamma distributed random effects. Normally distributed random effects gave very similar results.

Thus, if we estimate a matching function which includes a correctly specified measure of old and new stocks, the resulting hazard should be approximately constant.

In Table 3 we report estimates of a suitably parameterised stock-flow matching model. We use a random-effects specification for the estimates of h and λ, and a pooled specification for the estimates of μ. As before, we estimate three models on both sides of the market: the matching rate h, the contact rate λ and the matching probability μ. The top panel gives estimates for "new" job-seekers and vacancies, the bottom panel for "old" job-seekers and vacancies.

We are most interested in the relationship between "old" job-seekers and vacancies and the stock of "new" agents on the other side of the market, given in the bottom panel of Table 3. First, note that the elasticity of h^w with respect to v is positive and significant (0.133 (0.055)), over and above the estimated elasticity of h^w with respect to V. In other words, a job-seeker who has been in the market for more than a month has a significantly higher exit rate if there is a larger stock of "new" vacancies on the other side of the market. Under random matching, the coefficient on $\log v$ should be zero. Exactly the same effect is observed on the other side of the market: the elasticity of h^e with respect to u is also positive and strongly significant (0.193 (0.064)).

Does this effect come from an increase in the number of contacts (λ) or an increase in the probability of a match? It appears to be both, although the effect on μ is larger on both sides of the market. The elasticity of both λ^w and μ^w with respect to v is positive and significant (0.069 (0.022)) and 0.168(0.030) respectively). On the other side of the market, the elasticities of both λ^e and μ^e with respect to u are both positive and significant, and again the latter is larger and more significant.

Our final piece of evidence comes from the resulting hazards, plotted from the estimates of h, λ and μ in Table 3. As we noted earlier, if the decline in the overall hazard is a result of the collapse in the numbers of suitable partners on the other side of the market, then once we have controlled for this the hazard should be flat. In Figure 4 we plot the estimated hazards and matching probabilities from the model estimated in Table 3. It is noticeable that these hazards decline far less than the equivalent hazards from the estimates which included only the total stock of U and V as covariates. We can see in particular that the contact hazard is now almost flat, which is strongly suggestive that the fall observed in the raw data is due to the sudden reduction in the number of suitable partners. This is entirely consistent with the stock-flow model.

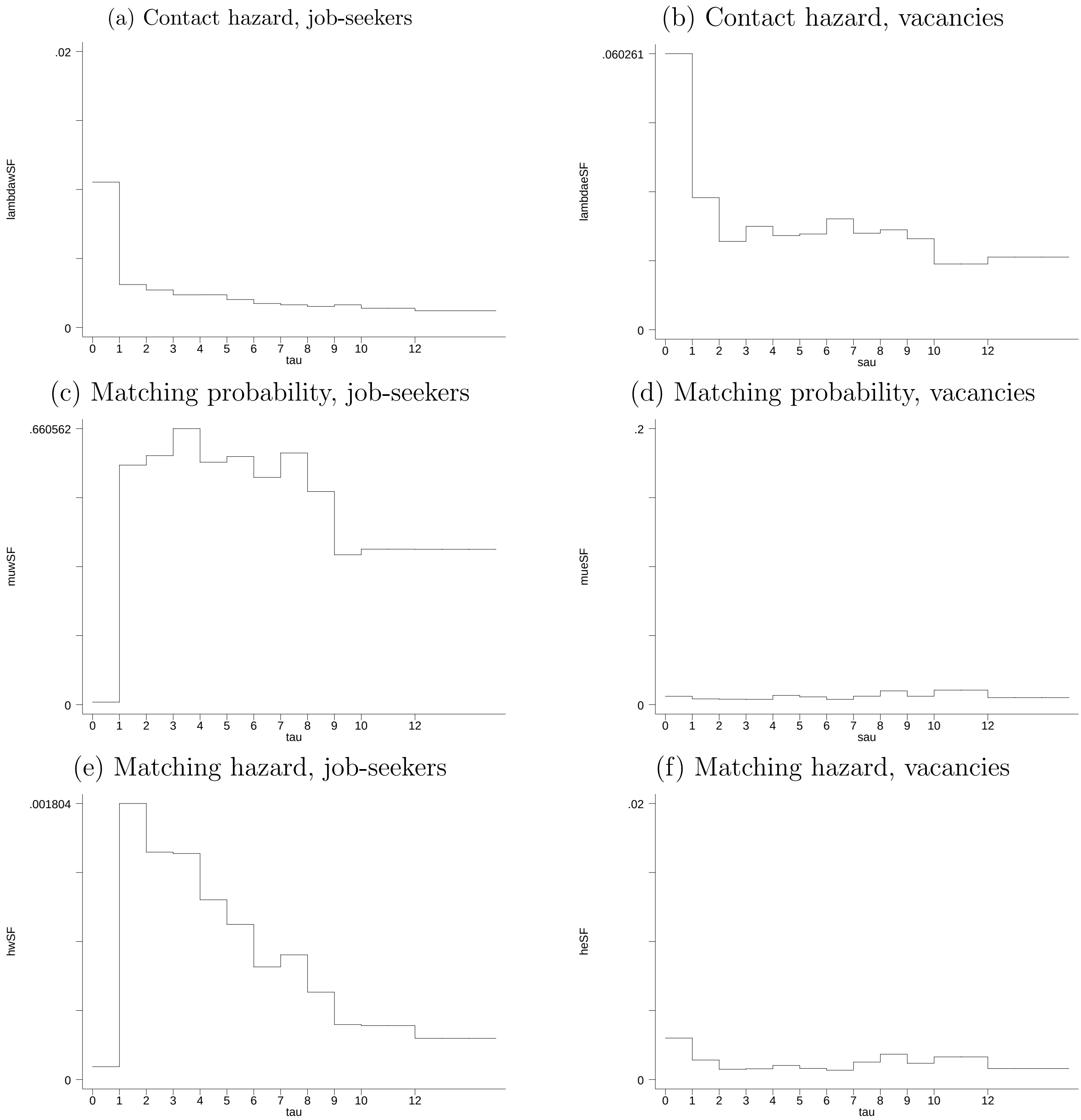


Figure 4: Hazards and matching probabilities, stock-flow model

6 Conclusions

A large empirical literature has estimated and tested various components of the search and matching framework, but very few estimates have allowed one to decompose the rate at which agents successfully match into the rate at which they contact, and the probability that a contact results in a match. Uniquely, we have been able to do this simultaneously for both sides of the same market. We have been able to do both in this paper. This allows us to shed considerable light on some of the fundamental assumptions of the search and matching framework.

To summarise, we estimate the parameters of λ^w, μ, and $h^w \equiv \lambda^w \mu$ in the rest of this paper, where the specifications for λ^w, μ, and h^w are given above. In words, we estimate the parameters of the contact function using count-data techniques using our monthly panel of contacts/matches; we estimate the parameters of the hazard function using the same data with Binary Choice techniques, and third we estimate μ by estimating binary choice models on the sample of contacts, using m as the dependent variable. However, we are able to estimate models that go beyond the basic Poisson assumption by modelling unobserved heterogeneity using random-effects techniques, by exploiting the longitudinal nature of the data, as well as model baseline hazards using non-parametric duration dummies.

We then estimate these three models for the employer side of the market, which means that we also estimate a model for the number of job-seekers arriving to a given vacancy (thereby estimating λ^e), the number of vacancy matches per period (thereby estimating $h^e \equiv \lambda^e \mu$), and a binary choice model for the probability that a contact matches (thereby estimating μ).

We find that in the raw data the decline in the matching rate for both job-seekers and vacancies is driven by a sharp decline in the contact rate, and not by any fall in the probability of a match conditional on a contact. We then estimate a two-sided matching model in order to determine whether this result is caused by omitted observed or unobserved heterogeneity in job-seekers and vacancies. It also allows us to estimate the parameters of the individual components of the matching function. We find that the same result applies as in the raw data: the decline in the matching rate on both sides of the market is driven by the decline in the contact rate.

We then estimate a more general matching model, one which nests the random matching model, in order to test whether the stock-flow matching model is consistent with the data. Our results are strongly suggestive that it is the decline in the number of suitable partners which occurs once an initial period has passed which is responsible for the sharp decline in the contact rate and hence the matching rate.

Appendix A Algebraic details

The probability density of observing a particular realisation c is

$$p(c) = \frac{e^{-\lambda^w t}(\lambda^w t)^c}{c!}, \quad c = 0, 1, 2, \ldots.$$

The probability density of observing m matches given c contacts, $c-1$ of which have already been rejected, is

$$p(m|c) = \mu^m (1-\mu)^{1-m}, \quad m = 0, 1.$$

Thus the joint density of observing m and c

$$p(m, c) = p(c)p(m|c) = \frac{e^{-\lambda^w t}(\lambda^w t)^c \mu^m (1-\mu)^{1-m}}{c!} \tag{A.1}$$

Adding up $p(m, c)$ over all values of c gives

$$\Pr(m = 1) = \mu(1 - e^{-\lambda^w t}) \approx \mu\lambda^w t, \tag{A.2}$$
$$\Pr(m = 0) = (1-\mu)(1 - e^{-\lambda^w t}) + e^{-\lambda^w t} \approx 1 - \mu\lambda^w t. \tag{A.3}$$

with the approximation being $e^x = 1 - x$ for any x. This is useful, because it shows that if one had only data on matches, the average number of matches per period, with exposure t, is $h^w \equiv \mu\lambda^w$. Clearly, without data on contacts, one cannot separately identify μ and λ^w.

References

Anderson, P. & Burgess, S. (2000), 'Empirical matching functions: estimation and interpretation using disaggregate data', *Review of Economics and Statistics* **82**, 93–102.

Andrews, M., Bradley, S., Stott, D. & Upward, R. (2003), Testing theories of labour market matching, Discussion Paper No. 0318, University of Manchester.

Andrews, M., Bradley, S. & Upward, R. (2001), 'Estimating the probability of a match using micro-economic data for the youth labour market', *Labour Economics* **8**, 335–57.

Andrews, M., Bradley, S. & Upward, R. (2003), Employer search, hard-to-fill vacancies and skill shortages, mimeo, University of Manchester.

Barron, J., Berger, M. & Black, D. (1997), Employer search, training and vacancy duration. Upjohn Institute for Employment Research, Kalamazoo: Michigan.

Broersma, L. & van Ours, J. (1999), 'Job searchers, job matches and the elasticity of matching', *Labour Economics* **6**, 77–93.

Burdett, K. & Coles, M. (1999), 'Long-term partnership formation: marriage and employment', *Economic Journal* **109**, F307–334.

Burdett, K. & Cunningham, E. (1998), 'Toward a theory of vacancies', *Journal of Labor Economics* **16**, 445–78.

Cameron, A. & Trivedi, P. (1998), *Regression Analysis of Count Data*, Cambridge University Press, Cambridge.

Coles, M. & Petrongolo, B. (2003), A test between unemployment theories using matching data. Mimeo, July.

Coles, M. & Smith, E. (1998), 'Marketplaces and matching', *International Economic Review* **39**, 239–55.

Eckstein, Z. & van den Berg, G. (2006), 'Empirical labor research: a survey', *Journal of Econometrics* p. Forthcoming.

Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Cambridge University Press, Cambridge.

Lindeboom, M., van Ours, J. & Renes, G. (1994), 'Matching employers and workers: an empirical analysis of the effectiveness of search', *Oxford Economic Papers* **46**, 45–67.

Mortensen, D. & Pissarides, C. (1998), Job reallocation, employment fluctuations and unemployment differences, *in* M. Woodford & J. Taylor, eds, 'Handbook of Macroeconomics', North-Holland, Amsterdam, chapter 18, pp. 1171–228.

Mortensen, D. & Pissarides, C. (1999), New developments in models of search in the labor market, *in* O. Ashenfelter & D. Card, eds, 'Handbook of Labor Economics', Vol. 3B, Elsevier, Amsterdam, chapter 39, pp. 2567–627.

Petrongolo, B. & Pissarides, C. (2001), 'Looking into the black box: a survey of the matching function', *Journal of Economic Literature* **39**, 390–431.

Petrongolo, B. & Pissarides, C. (2005), 'Scale effects in markets with search', *Economic Journal* p. Forthcoming.

Russo, G. & van Ommeren, J. (1998), 'Recruitment methods and vacancy duration', *Bulletin of Economic Research* **50**, 155–66.

Teyssière, G. (1996), 'Matching processes in the labour market: an econometric study', *Labour Economics* **2**, 421–35.

van den Berg, G. (1990), 'Search behaviour, transitions to non-participation and the duration of unemployment', *Economic Journal* **100**, 842–865.

van den Berg, G. (1999), 'Empirical inference with equilibrium search models of the labour market', *Economic Journal* **109**, F283–306.

van Ours, J. (1990), An empirical analysis of employers' search, *in* J. Hartog, G. Ridder & J. Theeuwes, eds, 'Panel Data and Labor Market Studies', North-Holland, Amsterdam, pp. 191–214.

van Ours, J. & Lindeboom, M. (1996), "Seek and ye shall find": an empirical analysis of the matching of job seekers and vacancies. Paper presented at Labour Market Changes and Income Dynamics conference.

van Ours, J. & Ridder, G. (1991), 'Cyclical variations in vacancy durations and vacancy flows: an empirical analysis', *European Economic Review* **35**, 1143–1155.

van Ours, J. & Ridder, G. (1992), 'Vacancies and the recruitment of new employees', *Journal of Labor Economics* **10**, 138–155.

van Ours, J. & Ridder, G. (1993), 'Vacancy durations: search or selection?', *Oxford Bulletin of Economics and Statistics* **55**, 187–198.