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 2003

Version submitted to

European Association of Labour Economists Conference 2003
and ESRC

*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 data were kindly supplied by Lancashire Careers Service. Subsequent revisions to this paper will be available at `http://les1.man.ac.uk/ses/staff/ma/`.

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.

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

Address for Correspondence:

Dr. R. Upward
School of Economics
University of Nottingham
Nottingham NG7 2RD
Email: `richard.upward@nottingham.ac.uk`

1 Introduction

Search theory is becoming the dominant paradigm in explaining micro- and macro- labour-market phenomena, see Mortensen & Pissarides' recent (1998, 1999) surveys. Examples include the "flows approach" to the study of labour markets, and how movements in the Beveridge Curve can explain whether the labour market has become more or less effective over time. See, for example, Blanchard & Diamond (1989, 1992). Other policy issues are often discussed in a search framework, see, for example, Manning's recent (2001) discussion on labour-market interventions and Marimon & Zilibotti (2000) on whether worksharing policy can reduce unemployment. Of course, search theory has also been central in modelling the impact of benefits on unemployment duration — Lancaster (1979), Meyer (1990), and many studies in between — and what causes long-term unemployment (Jackman, Layard & Pissarides 1989).

Burdett & Coles (1999) identify four key assumptions of the search and matching literature. The three that are relevant to labour markets are (i) *Poisson arrival rates*, that is, the process that generates contacts between employers and workers is a Poisson; (ii) *random matching*, that is, if an employer contacts a worker, it is assumed the worker's identity is a random draw from all possible workers; and (iii) there exists an *encounter function*. To quote Burdett & Coles (1999) directly: "deep in the heart of all search and matching models is . . . an encounter function . . . which relates the numbers of encounters per unit of time [contacts] as a function of [stocks of] unemployed workers and vacancies". In other words, the encounter function is a production function that generates a flow of contacts. If all contacts lead to matches (hires), it is also the matching function. In the real world, and also in theoretical models that have heterogeneous agents, not all contacts lead to matches and so the probability of a match is also modelled.

These three assumptions are deeply embedded in the search and matching literature. It is therefore important to establish whether they are observed in real world data. However, most evidence is based on aggregate time series data or data on individual unemployment durations, and therefore provides estimates of the matching function only. As noted by Petrongolo & Pissarides (2002), ". . . 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 firms and workers to accept the other side's offer."

In this paper we use micro-level data from a well-defined labour market to estimate directly the individual components of the matching function. We are able to decompose both the employer and worker hazard functions into a matching probability and the arrival rate of applicants from the other side of the market. In particular, we estimate the hazards to matching on both sides of the market, examine whether they are downward-sloping, and if they are, establish whether it is because job-seekers/vacancies receive fewer applicants or whether the matching probabilities decline with duration.

The decomposition of the matching function sheds light in particular on the assumption of random matching. An alternative assumption about the meeting technology is that workers 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 matching* is almost exclusively associated with Melvyn Coles and collaborators, who provide a persuasive alternative view as to how agents search and match with each other. This is the *stock-flow matching* model.[1] To test this model formally we require micro-level data which contains the identity of each pair who contact and match.

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; a companion paper discusses more fully how the stock-flow matching model of Coles and collaborators generalises the model presented here (Andrews, Bradley, Stott & Upward 2003), and thereby develops a formal test of random matching. There are stocks of vacancies V and job seekers U (all of whom

[1]The best exposition is Coles & Smith (1998), but also see Coles & Petrongolo (2003).

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 worker finds an employer acceptable and an employer finds a worker acceptable. In some two-sided search models $\mu(\theta)$ is 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,[2] but there is far less evidence for vacancies.[3] 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

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

[3] See, for example, van Ours & Ridder (1991, 1992, 1993), Barron, Berger & Black (1997), Burdett & Cunningham (1998), and Russo & van Ommeren (1998), Andrews, Bradley & Upward (2001*a*).

for using appropriate econometric techniques (see below), negative duration dependence can arise either because the arrival rate of suitable offers falls and/or the matching probability falls, as seen in decomposing the hazard in (1) above.[4] Other microeconometric studies do not estimate either hazard directly. Some have estimated the hiring function $\delta(U, V)$ directly[5] or the matching probability[6] or better still, have decomposed the hiring function into λ and μ (see equation 3).[7] However, the great majority of empirical work on the hiring function has used aggregate time-series data.[8]

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.

3 The data

The data we use are the computerised records of the Lancashire Careers Service. The Careers Service holds records on *all* youths aged between 15 and 18, including those who are seeking employment. We observe every vacancy notified by employers to the Careers Service between March 1988 and June 1992. About 30% of job vacancies are notified to the Careers Service. Job vacancies for which the Careers Service is not the method of search are not included in the data. Job vacancies require both high- and low-quality job seekers, and are representative of all entry-level jobs in the youth labour market. It follows that our data are representative of all job seekers, because we observe *all* contacts between notified job vacancies and job seekers.

Each contact, and therefore each match, in the labour market covered by the Careers Service data originates from a stock of job-seekers and a stock of vacancies. Job-seekers can come from one of four labour market states: unemployment, employment, in training, in education. Each vacancy is filled by one of these types of job-seeker, or it is "lapsed" or it is censored. A vacancy is said to have lapsed if it is withdrawn from the Careers Service database.[9]

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

[5] See Lindeboom, van Ours & Renes (1994), Anderson & Burgess (1997), and Broersma & van Ours (1999).

[6] See Teyssière (1996) and Andrews, Bradley & Upward (2001*b*).

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

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

[9] Andrews et al. (2001*a*) discuss the process of lapsing in more detail.

There are very few censored observations in these data because the sampling period is long in relation to the typical duration of a vacancy. 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.[10]

The primary unit of observation is a contact, c, and the associated binary variable m takes the value unity if a match occurs. In this paper, we examine only contacts and matches between job vacancies filled via the Careers Service and unemployed job-seekers. There are $25,267$ such contacts, resulting in $2,761$ matches. Associated with each contact is the identity of the job-seeker i and vacancy j (itself associated with an employer).

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 observe the day on which the job-seeker i became unemployed. We also observe the number of contacts received by each job-seeker i for each month i was unemployed, $c_{i\tau}$, $\tau = 1, \dots, T_i$, 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).

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. For spells which end within a month, the end date is taken to be either the last day of the month in which the contact takes place, or the start date of the resulting job spell, whichever occurs first. 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

[10] A job seeker who stops searching and leaves the labour market is the analogue of a vacancy which lapses.

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).

To summarise, when modelling search duration, we have a monthly unbalanced panel of observations for each job-seeker, $i = 1, \ldots, N$, $\tau = 1, \ldots, T_i$, where the variables being modelled are the number of contacts and the number of matches per month. 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. Consequently, we have a second monthly unbalanced panel of observations for each vacancy.

For each contact/match, we observe the following vector of information for the covariates:

$$(\tau, i, j, U, V, \tau^w, \tau^e, \mathbf{x}^w, \mathbf{x}^e, \omega),$$

where

- U is the stock of job-seekers. This varies by month through the duration of the job-seeker's spell of unemployment or employer's vacancy spell;
- τ^w is the current duration of the spell of unemployment (measured in months) and is used to construct piecewise linear duration dummies;
- a vector of characteristics $\mathbf{x}^w$, observed at the beginning of the spell;
- V is the stock of vacancies. This varies by month through the duration of the job-seeker's spell of unemployment or employer's vacancy spell;
- τ^e is the current duration of the spell of the vacancy (measured in months) and is used to construct piecewise linear duration dummies;
- a vector of characteristics $\mathbf{x}^e$, observed at the beginning of the spell;
- the wage ω on offer by the employer (if observed).

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 geographically distinct towns/cities (in fact, local authority districts). The issue here is whether the stocks should vary by these 14 districts, being distinct labour markets, or whether the same value should be used irrespective of where in Lancashire the match takes place, or something in between. For the intermediate case, we group Lancashire into just 3 labour markets (West, Central and East), recognising that job-seekers can travel between certain towns when looking for work. When we specify just three "districts" in Lancashire, 96% of all matches take place between an unemployed job seeker and job vacancy from the same district. This number drops to 75% when Lancashire is treated as 14 districts, which is convincing evidence that the 3 district specification is the best one.[11]

There are two other issues concerning the stocks of U and V. In a companion paper (Andrews et al. 2003), we test the stock-flow matching model of Coles, Smith and Petrongolo against the special case of random matching. This involves decomposing the stock of unemployed job-seekers U into stocks of new job-seekers u, ie all those who have been unemployed less than k weeks, and the rest, $\bar{U}$. The same decomposition applies to vacancies: $V = \bar{V} + v$. As noted, the decomposition of the matching hazard into the contact function and the matching probability allows us to test whether the stock-flow matching model is consistent with the behaviour of the arrival rate of contacts or the matching probability (or both, or neither).

Second, temporal aggregation bias is an important issue in this literature, and is discussed at length by Burdett, Coles & van Ours (1994), Gregg & Petrongolo (1997) and Coles & Petrongolo (2003). In the context of monthly data, the problem arises in not observing the instantaneous hiring rate, but rather flows over a discrete period (a month). The assumptions one needs to adjust the stock measures depend on how quickly agents are matching, which itself is being modelled, and so there is a simultaneity bias. Coles & Petrongolo (2003) estimate matching functions using a quite sophisticated technique that deals with this problem. In our data this may well be a problem, as our stocks are monthly, in fact observed at the beginning of the month.

We can now be more precise about the models we seek to estimate in the rest of this paper. Instead of models for h^e, h^w, λ^e, λ^w and μ based on Equation (1),

[11] Throughout Huber/White standard errors correct for within labour-market correlations between job-seekers/vacancies.

we estimate

$$\lambda^w = \lambda^w(U, V, \tau^w, \mathbf{x}^w, \omega)$$
$$\lambda^e = \lambda^e(U, V, \tau^e, \mathbf{x}^e, \omega)$$
$$\mu = \mu(U, V, \tau^w, \tau^e, \mathbf{x}^w, \mathbf{x}^e, \omega).$$

This clearly shows that the arrival rate of vacancies to job-seekers only depends on the current duration of the job-seeker τ^w and his/her characteristics $\mathbf{x}^w$, and not those of the vacancies $\tau^e, \mathbf{x}^e$. $\tau^e, \mathbf{x}^e$ only affect the hazard via the matching probability. Similar considerations apply to vacancies. The exit hazard are

$$h^w(U, V, \tau^w, \tau^e, \mathbf{x}^w, \mathbf{x}^e, \omega) = \lambda^w(U, V, \tau^w, \mathbf{x}^w, \omega)\mu(U, V, \tau^w, \tau^e, \mathbf{x}^w, \mathbf{x}^e, \omega)$$
$$h^e(U, V, \tau^w, \tau^e, \mathbf{x}^w, \mathbf{x}^e, \omega) = \lambda^e(U, V, \tau^e, \mathbf{x}^e, \omega)\mu(U, V, \tau^w, \tau^e, \mathbf{x}^w, \mathbf{x}^e, \omega).$$

giving the five functions we seek to estimate (note that the μ function is identical for both sides of the market). These hazards are potentially influenced by variables from the other side of the market—because we observe the identity of both partners in a match—which is very unusual in empirical studies of unemployment hazards.

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

$$\log \lambda^w = (\alpha_1 - 1)\log U + \beta_1 \log V \tag{4}$$
$$\log \lambda^e = \alpha_2 \log U + (\beta_2 - 1)\log V \tag{5}$$
$$\log \mu = \alpha_3 \log U + \beta_3 \log V \tag{6}$$
$$\log h^w = (\alpha_4 - 1)\log U + \beta_4 \log V \tag{7}$$
$$\log h^e = \alpha_5 \log U + (\beta_5 - 1)\log V. \tag{8}$$

Once estimated, we can decompose the effect of any covariate that affects either 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.

$$\frac{\partial \log h^w}{\partial \log U} = \frac{\partial \log \mu}{\partial \log U} + \frac{\partial \log \lambda^w}{\partial \log U}.$$

For example, 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 unemployment duration τ^w, is the standard finding that unemployment hazards fall due to falling arrival rate of applicants, falling matching probabilities, or both?

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 meet (or 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:

$$C \sim \text{Poisson}(\lambda^w). \tag{9}$$

λ^w is the Poisson parameter, and denotes the average number of contacts per job-seeker each period. The probability density of observing a particular realisation c is

$$p(c) = \frac{e^{-\lambda^w}(\lambda^w)^c}{c!}, \quad c = 0, 1, 2, \ldots. \tag{10}$$

Every time a vacancy is encountered, the pair either consummates the meeting by matching with each other, or they do not. The probability that a given meeting results in a match is denoted μ. Unsuccessful meetings return to the stocks of U and V; the resulting flow of hires per period is denoted M. Then the number of hires that result from C meetings is distributed as a Binomial:

$$M|C \sim \text{Binomial}(\mu, C) \tag{11}$$

or the probability density of observing m matches given c contacts is

$$p(m|c) = \frac{c!\mu^m(1-\mu)^{c-m}}{m!(c-m)!}, \quad m = 0, 1, 2, \ldots c. \tag{12}$$

It is easily shown that the marginal distribution for the number of hires per period is also Poisson distributed (see Section 4.2):

$$M \sim \text{Poisson}(\lambda^w \mu). \tag{13}$$

Intuitively, if λ^w vacancies are encountered on average each period and a proportion μ of them match on average, then the average matching rate of job-seekers is $\lambda^w \mu \equiv h^w$. If C is Poisson distributed, so must be M.

From the joint density of observing m and c

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

λ^w and μ can be estimated. With a sample of observations on pairs of contacts and matches (m_i,c_i), $i = 1, \ldots, n$, the log-likelihood ℓ of observing the sample is

$$\ell = -n\lambda^w + \log \lambda^w \sum c_i + \log \mu \sum m_i + \log(1-\mu) \sum (c_i - m_i) + \text{const.}$$

The first order conditions give expressions for ML estimates $\hat{\lambda}^w$ and $\hat{\mu}$:

$$\frac{\partial \ell}{\partial \lambda^w} = -n + \frac{\sum c_i}{\lambda^w} = 0 \quad \Rightarrow \quad \hat{\lambda}^w = \frac{\sum c_i}{n} \tag{14}$$

$$\frac{\partial \ell}{\partial \mu} = \frac{\sum m_i}{\mu} + \frac{\sum (m_i - c_i)}{1-\mu} = 0 \quad \Rightarrow \quad \hat{\mu} = \frac{\sum m_i}{\sum c_i} \tag{15}$$

This demonstrates that λ^w and μ can be estimated separately, from equations (10) and (12) respectively, rather than using the joint density. With a sample of data on only m_i the ML estimate of the average matching rate is given by:

$$\hat{h}^w \equiv \hat{\lambda}^w \hat{\mu} = \sum m_i / n, \tag{16}$$

which emphasises the point that only if we observe data on both hires *and* contacts can the average matching rate be decomposed into the average contact rate λ^w and the matching probability μ.[12]

To summarise, equations (9), (11) and (13) will be used to estimate parameterisations of λ^w, μ, and $h^w \equiv \lambda^w \mu$ respectively in the rest of this paper, where the precise specifications for λ^w, μ, and h^w were given at the end of Section 3. In words, we shall estimate the parameters of the contact function using count-data techniques using our monthly panel of contacts/matches; we shall estimate the parameters of the hazard function using the same data, and third we shall estimate μ by estimating binary choice models on the sample of contacts, using m as the dependent variable. However, we are able to

[12] When λ^w and μ are parameterised by regression functions, the decomposition holds exactly only when categorical data are used throughout.

estimate models that go beyond the basic Poisson assumption by modelling unobserved heterogeneity using random and fixed-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 count data model for the number of job-seekers arriving to a given vacancy (thereby estimating λ^e), the number of vacancy hires 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). In theory there is a single binary choice model for μ, which is a function of characteristics of both sides of the market, namely $(U, V, \tau^w, \tau^e, \mathbf{x}^w, \mathbf{x}^e, \omega)$. In the empirical work we actually estimate a separate model for each side of the market, namely $\mu^e(U, V, \tau^e, \mathbf{x}^e, \omega)$ and $\mu^w(U, V, \tau^w, \mathbf{x}^w, \omega)$.

4.2 Econometric techniques

Consider the result illustrated in the previous subsection, namely that if contacts are distributed Poisson (λ^w), then matches are distributed Poisson ($\lambda^w\mu$) (the same applies to $\lambda^e\mu$ on the other side of the market). This is an example of a *Poisson-stopped Binomial* (Winkelmann 1997, Cameron & Trivedi 1998). Paraphrasing Cameron & Trivedi (1998), now assume that the number of contacts is no longer Poisson distributed, but has mean λ and variance σ^2. Now define X_i, $i = 1, \ldots, n$, as a sequence of c iid Bernoulli trials, in which each X_i takes only one of two values, 0 or 1, with probabilities $1-\mu$ and μ respectively. Then M, the number of matches, is given by $M = \sum_{i=1}^{c} X_i$. Using the Law of Iterated Expectations

$$\mathrm{E}(M) = \mathrm{E}[\mathrm{E}(M|C)] = \mathrm{E}(C\mu) = \mu\mathrm{E}(C) = \lambda\mu$$
$$\mathrm{Var}(M) = \mathrm{Var}[\mathrm{E}(M|C)] + \mathrm{E}[\mathrm{Var}(M|C)] = (\sigma^2 - \lambda)\mu^2 + \lambda\mu.$$

The actual distribution of M depends on the distribution of C.

This result provides the link between the count-data technique we use for contacts and the count-data technique we use for matches, and will also allow us to decompose estimates from matches $\mathrm{E}(M) = \lambda\mu$ into estimates for contacts $\mathrm{E}(C) = \lambda$, and the probability of a match given a contact $\mathrm{E}(M|C) = \mu$.

We estimate the following three econometric models:

1. $c_{i\tau}$ and $m_{i\tau}$ are both estimated as *Pooled-Poisson count models*. Inevitably, the data will be over-dispersed, and so we introduce an individual effect (unobserved heterogeneity) term into both count models:

$$\begin{aligned} c_{i\tau} &\sim \text{Poisson}(\lambda_{i\tau}) \qquad \lambda_{i\tau} = \text{E}(m_{i\tau}|\mathbf{x}_{i\tau}, a_i^c) = \exp(\log a_i^c + \mathbf{x}_{i\tau}\boldsymbol{\beta}^c) \\ m_{i\tau} &\sim \text{Poisson}(h_{i\tau}) \qquad h_{i\tau} = \text{E}(m_{i\tau}|\mathbf{x}_{i\tau}, a_i^m) = \exp(\log a_i^m + \mathbf{x}_{i\tau}\boldsymbol{\beta}^m) \end{aligned}$$

(We suppress superscript w and e for notational clarity.) This means that

2. $c_{i\tau}$ and $m_{i\tau}$ are estimated as *Poisson fixed-effects models*, where the a_i are incidental parameters that are concentrated out of the likelihood

   or

3. $c_{i\tau}$ and $m_{i\tau}$ are estimated as *Poisson random-effects models*, with either Gamma distributed or Normal distributed individual effects. If Gamma,

$$a_i \sim \text{Gamma}(\delta, \delta)$$

so that $\text{E}(a_i) = 1$) and $\text{Var}(a_i) = 1/\delta$. If Normal,

$$\log a_i \sim \text{Normal}(0, \sigma_\delta^2)$$

In all 3 cases, $\mu_{i\tau} = \Pr(m_{i\tau} = 1)$ is estimated as a Logit:

$$m_{i\tau}|m_{i\tau} \sim \text{Binomial}(\mu_{i\tau}) \qquad \mu_{i\tau} = \text{E}(m_{i\tau}|c_{i\tau}, \mathbf{x}_{i\tau}) = \Lambda(\mathbf{x}_{i\tau}\boldsymbol{\beta})$$

These techniques are well-established and we therefore do not provide further details (but see, in particular, Cameron & Trivedi (1998, chapter 9)).

If the link function for the binary choice model were exponential rather than logistic, it follows that $\boldsymbol{\beta}^m = \boldsymbol{\beta} + \boldsymbol{\beta}^c$. Convergence is typically poor because it is the Logistic that is the Canonical link. It is easy to show that, for the Logit,

$$\boldsymbol{\beta}^m = \boldsymbol{\beta}[1 - \Lambda(\mathbf{x}_{i\tau}\boldsymbol{\beta})] + \boldsymbol{\beta}^c \tag{17}$$

where $\Lambda(\mathbf{x}_{i\tau}\boldsymbol{\beta})$ is replaced by its sample average. As the average matching probability is about 0.1, this means that the Logit estimates need scaling by 0.9.

4.3 Non-parametric 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 applicants, $\lambda^w(\tau^w)$ and $\lambda^e(\tau^e)$, or matches, $h^w(\tau^w)$ and $h^e(\tau^e)$, or the matching probability, $\mu(\tau^w, \tau^e)$, 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.

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 $\tau = 1, \ldots, 13$, 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 $\hat{\mu} = \sum_{i\tau} m_{i\tau} / \sum_{i\tau} c_{i\tau}$, 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). Using Equation (16), where n is replaced by the number of days at risk (column 4), dividing $2,761$ matches by $3,235,810$ for job-seekers 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.

Similarly, dividing $25,267$ by the total days at risk gives the average contact rates for both sides of the market. These are again in the same 1:4 ratio, and both are about 10 times higher than the corresponding hazard rates, because $\hat{\mu} = 0.109$.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				(a) Job-Seekers				
τ	$\sum_i m_{i\tau}$	$\sum_i c_{i\tau}$	$\sum_i t_{i\tau}$	$\hat{h}^w$	$\hat{\lambda}^w$	$\hat{\mu}^w$	$\sum_i 1_{i\tau}$	$\frac{\sum_i t_{i\tau}}{\sum_i 1_{i\tau}}$
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				
τ	$\sum_i m_{i\tau}$	$\sum_i c_{i\tau}$	$\sum_i t_{i\tau}$	$\hat{h}^e$	$\hat{\lambda}^e$	$\hat{\mu}^e$	$\sum_i 1_{i\tau}$	$\frac{\sum_i t_{i\tau}}{\sum_i 1_{i\tau}}$
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

Table 1: Raw data, hazards and matching probabilities

A similar analysis applies to each row of the table, except that $\hat{\mu}^w_\tau \neq \hat{\mu}^e_\tau$. For example, column (3) of panel (a) shows that job-seekers receive $10,783$ 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 $613,051$.[13] Our estimate of the worker hazard to a successful contact is given in column (5), and comes from Equation (16).

$$\hat{h}^w_1 = \frac{\sum_i m_{i1}}{\sum_i t_{i1}} = 0.00112$$

From Equation (14), the contact function reported in column (6) is estimated from

$$\hat{\lambda}^w_1 = \frac{\sum_i c_{i1}}{\sum_i t_{i1}} = 0.01759,$$

and the matching probability in column (7) is

$$\hat{\mu}^w_1 = \frac{\sum_i m_{i1}}{\sum_i c_{i1}} = 0.06343.$$

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

Column (8) reports the total number of workers 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 worker or vacancy in that duration category. The $34,657$ workers 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 1 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 workers 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.

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

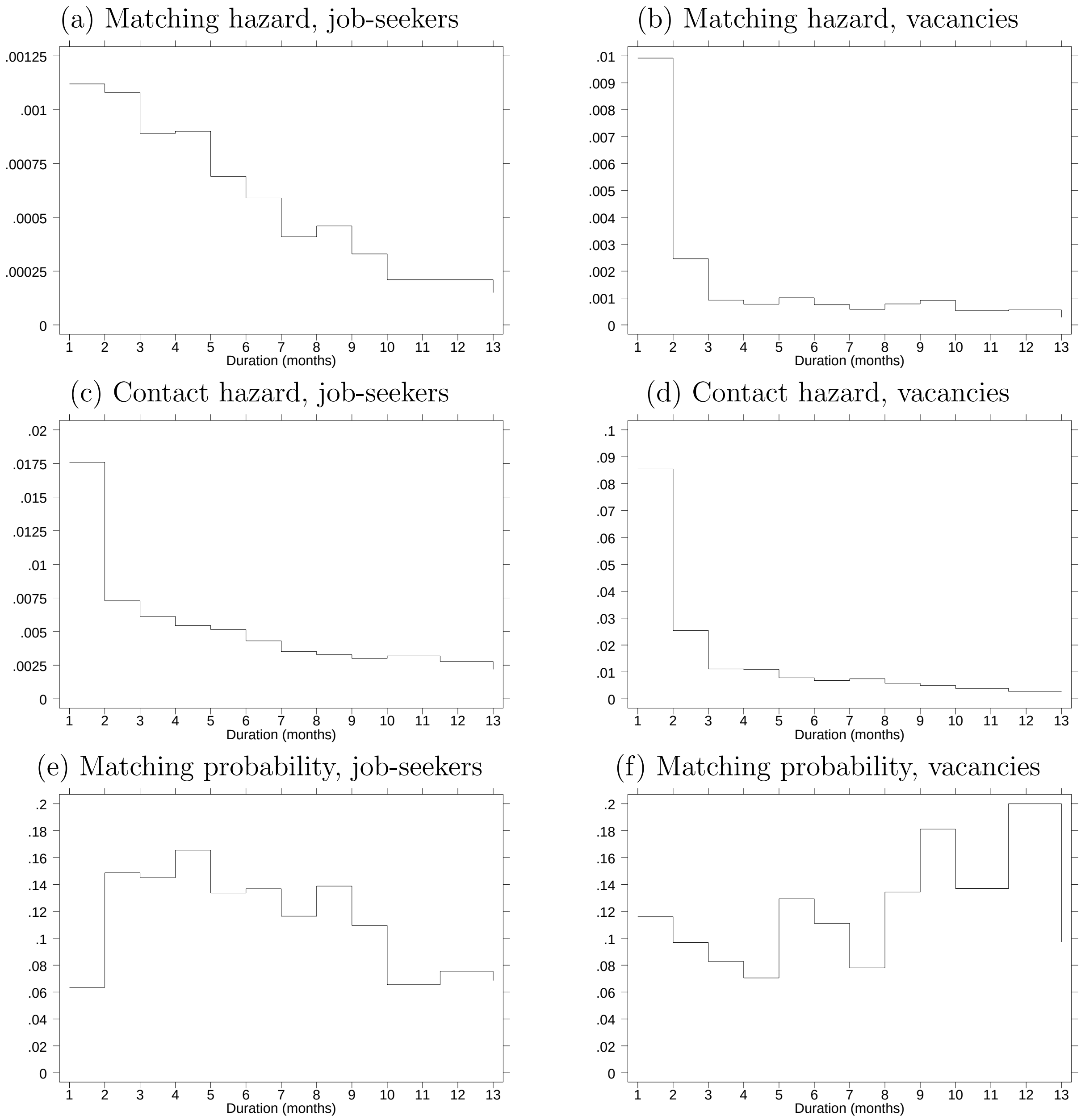

Figure 1: Raw hazards and matching probabilities

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 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 firm'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 4.2.

5.2 Estimates of random matching models

Table 2 reports estimates of Equations (4) to (8) across four different specifications. Panel (a) is the "baseline" specification, which is reported for comparison with estimates of the matching function from aggregated data. This specification imposes a constant (exponential) baseline hazard for both matches and contacts. Panels (b)–(d) allow the hazards and the matching probability to vary with τ by using the monthly piecewise-constant hazard specified in Section 4.3.

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

	Job-seekers			*Vacancies*			
	h^w	λ^w	μ^w	h^e	λ^e	μ^e	mean
(a) Baseline (no duration effects)							
$\log U$	−0.167 (0.039)	−0.147 (0.020)	−0.009 (0.049)	0.442 (0.035)	0.487 (0.051)	−0.009 (0.068)	759
$\log V$	0.405 (0.078)	0.522 (0.050)	−0.152 (0.068)	−0.478 (0.016)	−0.305 (0.048)	−0.152 (0.071)	216
$\alpha + \beta$	1.238 (0.039)	1.374 (0.031)	−0.162 (0.042)	0.964 (0.026)	1.182 (0.005)	−0.162 (0.021)	
Log likelihood	−14528.9	−78380.0	−8696.0	−12797.3	−69871.4	−8696.0	
(b) Pooled, non-parametric baseline							
$\log U$	−0.152 (0.035)	0.024 (0.042)	−0.142 (0.038)	0.426 (0.055)	0.452 (0.011)	−0.008 (0.066)	759
$\log V$	0.410 (0.059)	0.468 (0.072)	−0.111 (0.086)	−0.207 (0.047)	−0.058 (0.052)	−0.146 (0.076)	216
$\alpha + \beta$	1.258 (0.033)	1.492 (0.032)	−0.285 (0.063)	1.218 (0.029)	1.394 (0.047)	−0.174 (0.020)	
Log likelihood	−14337.1	−74055.5	−8433.4	−11185.6	−56274.3	−8666.6	
(c) Random effects, non-parametric baseline[b]							
$\log U$	−0.153 (0.021)	0.045 (0.009)		0.453 (0.030)	0.273 (0.013)		759
$\log V$	0.414 (0.029)	0.366 (0.014)		−0.107 (0.040)	0.208 (0.020)		216
Variance (σ^2)	0.273 (0.101)	1.892 (0.038)		2.199 (0.139)	1.398 (0.029)		
$\alpha + \beta$	1.261 (0.028)	1.411 (0.013)		1.346 (0.042)	1.481 (0.019)		
Log likelihood	−14332.5	−67005.8		−10681.5	−41244.5		
(d) Fixed effects, non-parametric baseline							
$\log U$	−0.096 (0.045)	0.065 (0.012)		0.458 (0.066)	0.106 (0.020)		759
$\log V$	0.385 (0.085)	0.202 (0.023)		0.160 (0.189)	0.525 (0.052)		216
$\alpha + \beta$	1.290 (0.080)	1.267 (0.020)		1.618 (0.169)	1.631 (0.046)		
Log likelihood	−4168.1	−33661.8		−2952.5	−21821.7		
Observations	127987	127987	25267	39762	39762	25267	
Obs. (fixed effects)	14854	63792		14731	29047		

[a] Estimates based on 25267 contacts and 2761 matches between 34657 unemployed job-seeker spells (26113 job-seekers) and 14154 Careers Service job vacancies.

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

Panel (b) is the pooled-Poisson model. The estimates in panels (c) and (d) also allow for unobserved heterogeneity (overdispersion) in the Poisson models. Panel (c) reports the Poisson random-effects model, while panel (d) reports the Poisson fixed-effects model. For each estimate we report the coefficients on U and V. 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.

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 (17). The novelty of our results is that from Equation (3) we can decompose this effect across the contact function and the matching probability. So, for example, in panel (a) the elasticity of h^w with respect to unemployment is -0.167, and that this effect is being driven by the elasticity of U with respect to contacts (-0.147) rather than the probability of a match (-0.009). Even more unusually we can also see the same effects on the other side of the same market. The elasticity of h^e with respect to U is 0.442 and this must be driven again by the increase in the number of contacts (0.487) rather than any effect on the matching probability.[14]

However, panel (a) does not allow for any variation in the baseline hazard in any of the estimates, and therefore estimates of α and β are potentially unreliable. To see why this might be the case, suppose that the "true" matching probability for the unemployed is sharply declining with duration, for whatever reason. Time periods and districts with higher stocks of U will tend to have a larger proportion of long-duration job-seekers with lower exit hazards. Thus the elasticity of h^w and λ^w with respect to U might be negative because of the composition of the pool of job-seekers rather than because of genuine congestion effects.

In panel (b) we can see that although estimates of α and β are quite similar for h^w and h^e, the decomposition is quite different. The elasticity of h^w with respect to U is -0.152, but this is driven by the elasticity of the matching probability (-0.142) rather than the contact rate (0.024). This suggests that

[14]Note that in panel (a) the estimates of μ^w and μ^e are identical because these estimates do not include τ^w or τ^e. The standard errors differ because for estimates of μ^w we cluster on the district of the job-seeker, and for estimates of μ^e we cluster on the district of the vacancy.

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 worker and the firm find that other side's offer acceptable, this is strong evidence that it is the firm effect which dominates. That is, as unemployment increases firms become more selective. We expect μ to be homogeneous of degree zero, so that $\alpha + \beta = 0$, but this is strongly rejected on both sides of the market.

The other potential inconsistency in these estimates arises because of unobserved heterogeneity. Job-seekers and vacancies differ in their "quality", a_i and a_j, in ways that are unobserved in the data. As noted in Section 4.2, this leads to overdispersion in the data. It should be stressed that this will only lead to inconsistent estimates of α and β if the a_i and a_j 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 both as a random and a fixed effect, reported in panels (c) and (d) of Table 2. It turns out that estimates of α and β on the worker's side of the market are quite robust to unobserved heterogeneity at the individual level, whether modelled as a fixed or a random effect. 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 quite different estimate of β. For example, the elasticity of h^e and λ^e with respect to the stock of vacancies moves from -0.207 (0.047) (pooled model) to -0.107 (0.040) (random effects) to 0.160 (0.189) and insignificant (fixed effects). This implies that the unobserved quality of the vacancies in our data are correlated with the stocks of V in each district-month.

What effect does controlling for U, V and unobserved heterogeneity have on the estimates of the baseline hazards? Figure 2 plots the estimated baseline

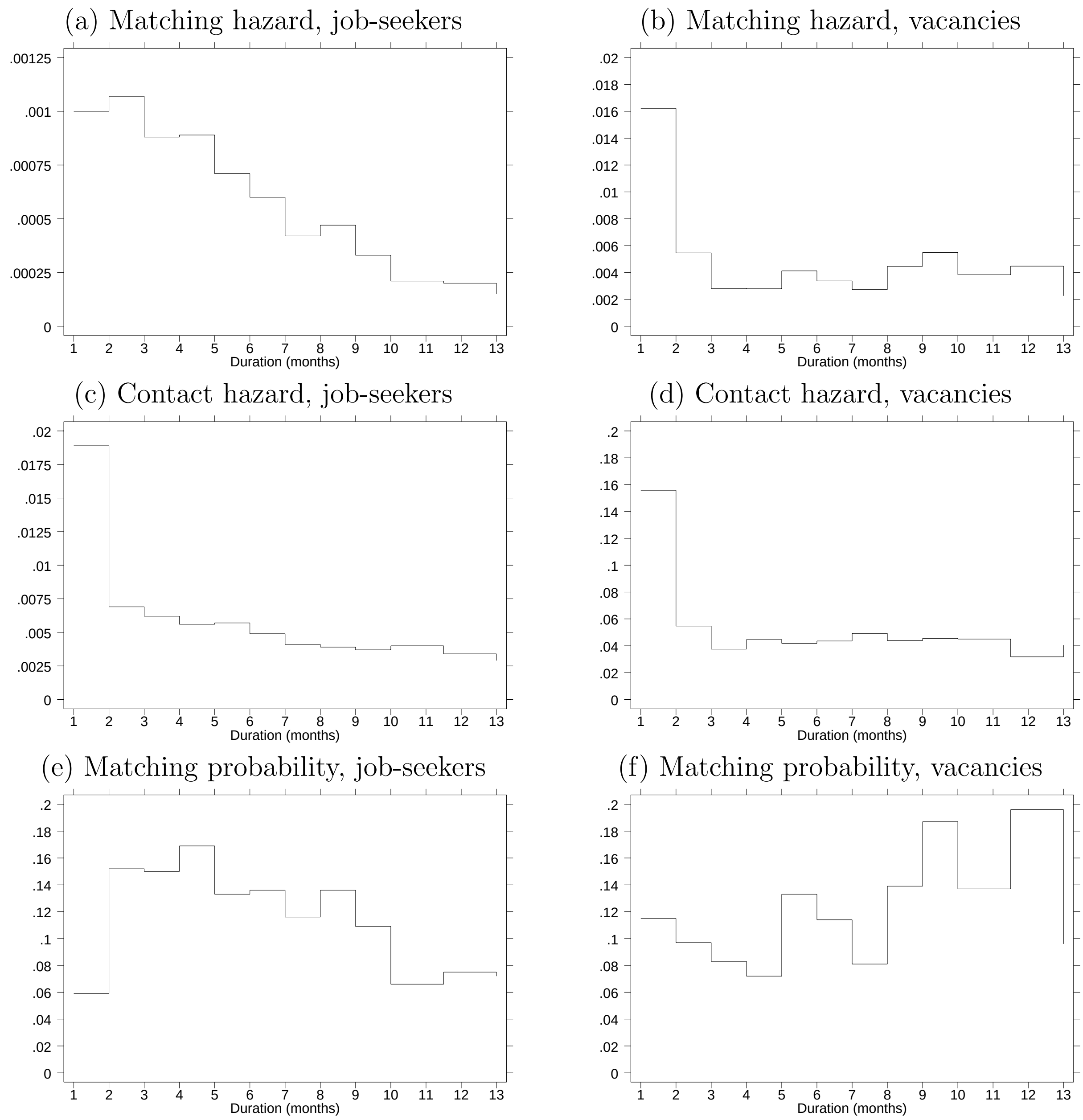

Figure 2: Hazards and matching probabilities, random matching model

hazard from panel (c) in Table 2.[15] The baseline hazards remain very similar to those plotted from the raw data (Figure 1), 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 et al. (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.[16] 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 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 1) and the econometric estimates (Figure 2). 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.

[15] All hazards are plotted using the mean values of U and V.
[16] For a fuller exposition, see Andrews et al. (2003), in particular Section 3.

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 3 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

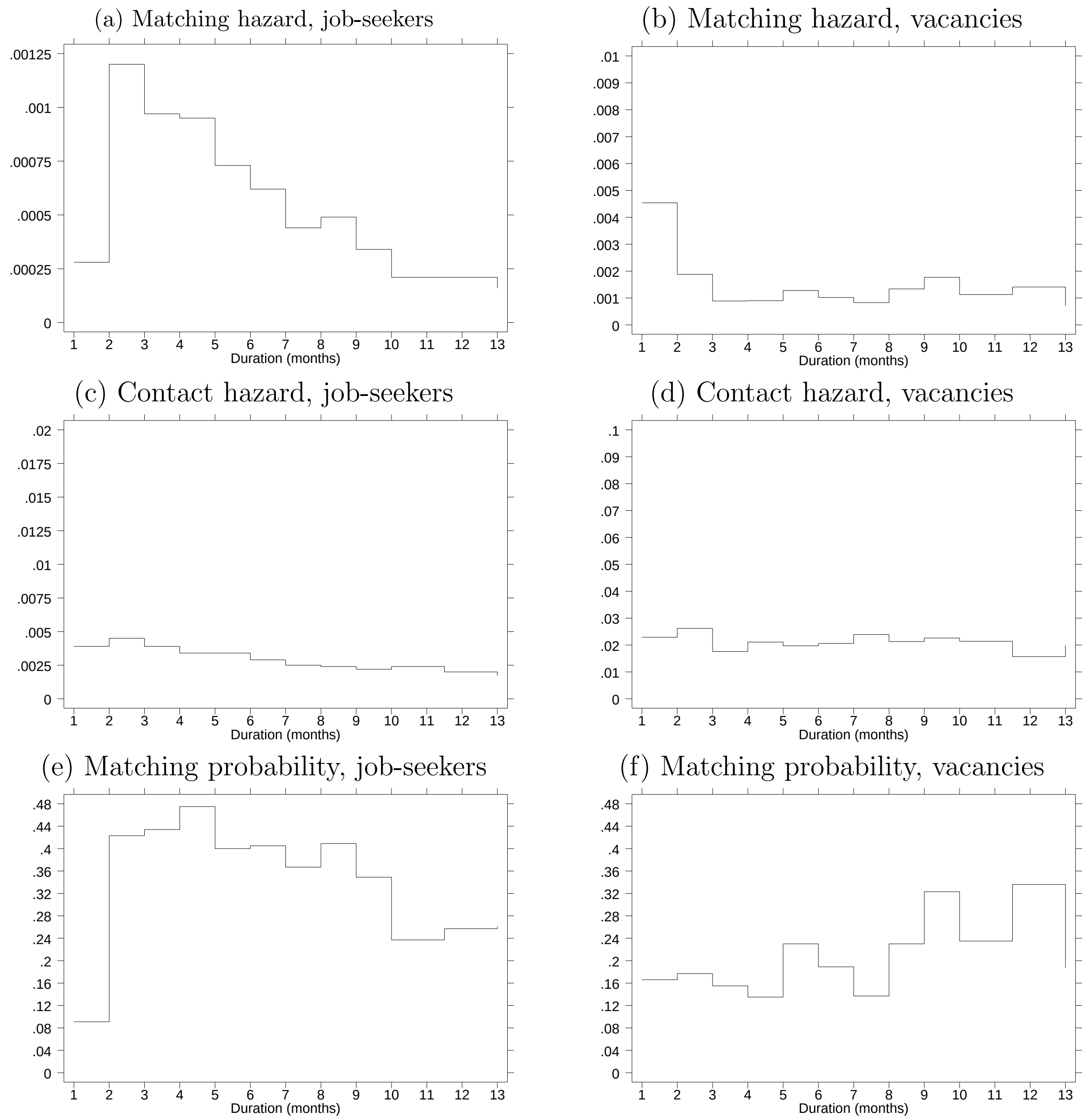

Figure 3: Hazards and matching probabilities, stock-flow model

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.

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.

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.

References

Anderson, P. & Burgess, S. (1997), Empirical matching functions: estimation and interpretation using disaggregate data, Working Paper 5001, National Bureau of Economic Research, February.

Andrews, M., Bradley, S., Stott, D. & Upward, R. (2003), Testing theories of labour market matching, Discussion paper, University of Manchester, March.

Andrews, M., Bradley, S. & Upward, R. (2001*a*), Employer search, vacancy duration, and skill shortages: an analysis of vacancies in the youth labour market, Discussion paper, University of Nottingham, June.

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

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

Blanchard, O. & Diamond, P. (1989), 'The Beveridge Curve', *Brookings Papers on Economic Activity* **1**, 1–60.

Blanchard, O. & Diamond, P. (1992), 'The flows approach to labor markets', *American Economic Review* **82**, 354–59.

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., Coles, M. & van Ours, J. (1994), Temporal aggregation bias in stock-flow models. CEPR Discussion Paper No. 967.

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, January.

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

Gregg, P. & Petrongolo, B. (1997), Random or non-random matching? Implications for the use of the UV curve as a measure of matching effectiveness. CEP Discussion Paper No. 348.

Jackman, R., Layard, R. & Pissarides, C. (1989), 'On vacancies', *Oxford Bulletin of Economics and Statistics* **51**, 377–94.

Lancaster, T. (1979), 'Econometric methods for the duration of unemployment', *Econometrica* **47**, 939–56.

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.

Manning, A. (2001), Monopsony and the efficiency of labour-market interventions. CEP Discussion Paper No. 514.

Marimon, R. & Zilibotti, F. (2000), 'Employment and distributional effects of restricting working time', *European Economic Review* **44**, 1291–1326.

Meyer, D. (1990), 'Unemployment insurance and unemployment spells', *Econometrica* **58**, 757–82.

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. (2002), Scale effects in markets with search, Discussion Paper 3648, CEPR, November.

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.

Winkelmann, R. (1997), *Econometric Analysis of Count Data*, Springer, Berlin.