KERP
2003/02

Starting Wages, Hires and Separations

Paul Bingley and
Gauthier Lanot

Keele Economics Research Papers



KEELE
UNIVERSITY

Keele ▪ April 2003

KERP **Keele Economics Research Papers**

The Keele Economics Department produces this series of research papers in order to stimulate discussion and invite feedback. Copyright remains with the authors.

All papers in the KERP series are available for downloading from the Keele Economics website, via *www.keele.ac.uk/depts/ec/kerp*.

ISSN 1740-231X

Department of Economics ▪ Keele University, Keele, Staffordshire, ST5 5BG, UK
▪ tel: (44) 1782 583091, fax: (44) 1782 717577 ▪ email: economics@keele.ac.uk

Starting Wages, Hires and Separations

by

Paul Bingley (Århus University)
Gauthier Lanot (Keele University)

Date 15 April 2003

Abstract We study the empirical relationship between the hiring rate, separation rate and starting wages. A practical empirical model is set up and estimated on Danish matched employer-employee longitudinal data for the period 1980--1995. We find (1) firm heterogeneity is important in all dimensions of our model: starting wages, time between hires and employment length (2) matching between firms and workers is exogenous (3) friction effects in the determination of wages are very small, which suggests that the Danish labour market is close to competitive.

Keywords linked longitudinal employer-employee data, unobserved heterogeneity, durations, semi-parametric mixtures

J.E.L. Class J31, J63, C33, C41

Address *Paul Bingley*: National Centre for Register-based Research, Tåsingegade 1, Århus University, 8000 Århus C, Denmark. *Gauthier Lanot*: Department of Economics, Keele, Staffordshire ST5 5BG, United Kingdom, Email g.lanot@econ.keele.ac.uk.

Download www.keele.ac.uk/depts/ec/web/wpapers/kerp0302.pdf

1. Introduction[1]

The classical model of the labour market assumes that firms respond instantaneously to changes in the prevailing wage rate by adjusting their labour intake until the marginal product of labour equals the market wage rate. However, assuming that the labour market is imperfect modifies this. Any market imperfection (which may be due to limited information, mobility costs, time to match, for example) creates a difference between marginal productivity and the wage rate at equilibrium. In particular in the case where information about the agents in the labour market is not perfect (the identities of firms looking for workers is unknown, or the identities of workers searching for jobs is unknown, and/or their reservation wages are unknown) and where contracting opportunities are limited (in the sense that firms meet only a limited number of workers per time period) and short lived, frictions may allow firms to set wages permanently below the marginal product (but above some reservation wage).

The original theoretical model proposed by Burdett and Mortensen (1998) describes such a labour market, where worker turnover exists in the presence of a non-degenerate wage distribution at equilibrium. Wage dispersion is directly related to the level of search frictions, or delays in matching job opportunities with workers. Although this model captures many stylised fact from the labour market, with homogeneous workers and firms, one important shortcoming is that the implied equilibrium wage (offer) distribution is right-skewed. Several applied papers introduce productivity heterogeneity in order to closer fit the empirical wage distribution. Van den Berg and Ridder (1998), for example, allow heterogeneity between labour markets; Bontemps, Robin and Van den Berg (1999) allow within labour market, between firm heterogeneity.

Koning, Van den Berg, Ridder and Albaek (2000) take the innovative approach of estimating equilibrium search models without imposing their full structure *a priori*. They set up two models: one allowing worker and firm productivity heterogeneity only between labour markets, and a second allowing firm heterogeneity within labour markets. In the first model, all within-market wage dispersion must be due to search frictions, whereas in the second model this is not the case. The models are estimated using Danish data drawn from the same database as our study. Labour markets are defined by the interaction of education, occupation and industries. Both structural wage regressions derived from the equilibrium search models were rejected by the data, and reduced form wage regressions found no effect of search frictions on wages.

---

[1] Financial support from the Danish Research Foundation, the Danish Social Science Research Council and the Research Foundation of Århus University is gratefully acknowledged. The usual disclaimer applies.

Prevailing empirical strategies are indirect since the history of flows into and out of the firm as well as individual wage rates are mostly unobserved or unused by researchers. Hence most of the existing literature relies on infrequent (usually annual) measurements of firm size and on firm average wages, i.e. it relies on distributional information. It is the tight theoretical specification of the models which allows researchers to proceed and obtain estimates.[2] Not surprisingly, since exogenous frictions present in the theoretical model are the source of the wage distribution, empirical results show that frictions are important.

In this paper we investigate the empirical relationship between observed frictions and (starting) wage rates using a reduced form model. The Danish data we use samples a large number of firms and follows all workers employed at the same firm over a long period of time. For each firm, we construct a history of individual hires and separations as well as the starting wage of any employee. Hence, because we observe the complete history of labour turnover for each firm in our sample, we are able to study empirically the role of turnover (hires and separations) in the determination of wages. A complete description of the relationship between the wage profile and labour turnover is beyond the scope of this paper, and we limit our focus to the determination of starting wages.

At the firm level, both the "thickness" of the labour market in which the firm operates (i.e. the ease with which a firm can find a worker, and a worker can meet a firm), and the firm's policy concerning hires, separations and starting wages may induce an association between the starting wage and the processes that determine hires and separations. The empirical strategy we follow does not allow us to distinguish between these two different interpretations. In other words, the association between "thickness" of the market and starting wages may be determined by market conditions exogenous to the firm and/or the firm's specific pay and employment policy.

Between firms, we are able to compare firms according to starting wages and number of hires and separations that take place in any time period (or equivalently according to the time elapsed between consecutive hires at the firm and employment duration following hires). Hence, in principle, we are able to assess the relationship between the hiring and separation activity of any given firm and the starting wage it is prepared to pay in subsequent periods.

Within the same firm, we wish to compare two hires in terms of the effect on the starting wage of a greater time having elapsed since the previous hire, or in terms of the effect on subsequent employment duration of a higher starting wage. The comparisons are informative about how firm-specific, observable and unobservable,

---

[2] See Canals and Stern (2001) for a further discussion.

characteristics determine the starting wage and how observable and unobservable characteristics at the individual transaction level determine the starting wage and employment duration. We want to measure the association between the current starting wage and the transaction-specific duration since the last hire. Furthermore we want to measure the sensitivity of the new hire's employment duration to the starting wage paid. The first association is indicative of the firms hiring activity, while the second association describes the ability of the firm to retain its employees. Finally our modelling strategy allows us to investigate whether, within any given firm, starting wages are autocorrelated between transactions (given all other firm specific characteristics).

In a recent contribution Postel-Vinay and Robin (2002) use a French matched employer-employee panel to estimate an extension to the model proposed by Burdett and Mortensen (1998). One (among many) of the important empirical consequences of their model is the (theoretically justified) decomposition of the variance of wages as the sum of three components: a term that represents the variance of the individual specific effect, another for the variance of the firm specific effect and a third term for the effect of market frictions. Using a dataset which follows workers during their work history and identifies their workplace, they find that labour market frictions account for over 50% of the total variance of log wages for most occupations, while the share between the variance of the individual effect and the firm effect is occupation dependent. Indeed, for managers the variance of the individual effect can be as much as twice the variance of the firm effect, while the variance of the firm specific effect explains all of the remainder in the case of manual workers. We show how our reduced form model is amenable to a similar decomposition of the total variance of log-wages.

Our findings are threefold : (1) the hypothesis of exogenous matching between firms and workers cannot be rejected, (2) firm heterogeneity is important in all dimensions of our model: starting wages, time between hires and employment length, (3) friction effects in the determination of wages are small.

The remainder of the paper is organised as follows. In the next section we outline our empirical modelling strategy. Then we describe our data source. The fourth section sets up the empirical model, first in its homogenous form and then introducing firm specific unobserved heterogeneity. In section five we present, interpret and discuss our estimation results. Finally we conclude.

2. Modelling Strategy

Our empirical model describes the distribution of the three variables of interest: first the time that elapses between individual hires within firm f, T_f^{H}, second the starting monthly wage w_{fi} of any new hire, i, and the subsequent employment duration of any new hire, T_{fi}^{S}. These three elements together describe the timing of a specific

hire among all of the firm's hires as well as the hiring conditions (the wage), and how long the current hire remains within the firm.

Let τ be individual i's hiring date into firm f, i.e. the transaction date. Note T_{fi}^{H} the duration from the last hire (i.e. the firm last hired a worker before this one in $\tau - T_{fi}^{\mathrm{H}}$), T_{fi}^{S} is the duration in employment (i.e. worker i will separate from firm f in $\tau + T_{fi}^{\mathrm{S}}$) and w_{fi}^{τ} denotes his/her starting wage. Figure 1 illustrates the sequence of events that we are modelling for a single firm. The horizontal axis shows calendar time, continuous lines represent different employment spells with the firm, and dotted lines represent elapsed times between this hiring date[3] and the previous hiring date at the firm. The vertical axis orders these spells according to calendar time of previous hire. Time between hires is positive and hires occur sequentially or may be zero in the case of lumpy (simultaneous) hires, while durations in employment (or until separations) can be simultaneous. Hence spells between hires do not overlap whereas employment spells may overlap.

There is no loss of generality by considering only the joint distribution of the duration between hires and the employment duration. Indeed, in principle, from this joint distribution it is possible to calculate (1) the joint distribution of duration between the last hire or the last separation and the current hire, employment duration for the current hire and the starting wage, or (2) the distribution of duration between the last separation and the current hire, current employment duration and the starting wage. The joint distribution of the duration between hires and the employment duration provides as rich a description of the hiring and separation process as any other joint distribution of the duration between events in the hiring and separation history of the firm. It is possible to deduce the distribution of the latter from the former, and in this sense our modelling approach is general.

More formally, firm f's history until the n th transaction can be written as the list $H_f^{\tau_n} = \left\{\left(T_{fi}^{H}, \hat{T}_{fi}^{S}, w_{fi}^{\tau_i}, X_{fi}^{\tau_i}\right)\right\}_{i=1}^{n}$, where, by convention, $\hat{T}_{fi}^{S}$ equals $+\infty$ if $T_{fi}^{S} + \sum_{i=1}^{n_f} T_{fi}^{H} > \tau_n$ (if by date τ_n individual worker i has not left the firm yet), and T_{fi}^{S} otherwise. This allows us to identify workers still in employment with the firm at the time of transaction. Furthermore, $X_{fi}^{\tau_i}$ collects the observed and unobserved characteristics specific to the transaction that occurred in τ_i (i.e. information specific to the firm and to the individual at that time). Our main modelling

[3] This could be, but it does not have to be, an indication that the firm has a vacancy that it wants to fill. However, hires and separations events can occur frequently within a single firm, and clearly not all events are indicative of the existence (or the absence) of one or many vacancies. We would require richer data to be able to address the issue of vacancies. See Van Ours and Ridder (1992) for an application.

assumption is that the vectors $\left(T_{fn}^{H}, \hat{T}_{fn}^{S}, w_{fn}^{\tau_n}, X_{fn}^{\tau_n}\right)$, which collect all the information describing each transaction, are identically and independently distributed given the history $H_f^{\tau_{n-1}}$ and some firm specific (and transaction invariant) descriptors, say ϑ_f, for all n (conditional independence). If we write $f\left(T_{fn}^{H}, \hat{T}_{fn}^{S}, w_{fn}^{\tau_n}, X_{fn}^{\tau_n} \middle| H_f^{\tau_{n-1}}\right)$ for the conditional probability density function of the vector $\left(T_{fn}^{H}, \hat{T}_{fn}^{S}, w_{fn}^{\tau_n}, X_{fn}^{\tau_n}\right)$ given $H_f^{\tau_{n-1}}$ and ϑ_f, then the joint density of a given firm's history up to the n_f th transaction, $f\left(H_f^{\tau_n} \middle| \vartheta_f\right)$, can be written as

$$f\left(H_f^{\tau_n} \middle| \vartheta_f\right) = \prod_{i=1}^{n} f\left(T_{fi}^{H}, \hat{T}_{fi}^{S}, w_{fn}^{\tau_i}, X_{fn}^{\tau_i} \middle| H_f^{\tau_{i-1}}, \vartheta_f\right).$$

In what follows we make the additional simplifying assumption that workers and firms match in an exogenous manner. To formalise this let $\tilde{X}_{fn}^{\tau}$ be the components of $X_{fn}^{\tau_n}$ which characterise the individual independently from the transaction, while $\hat{X}_{fn}^{\tau_n}$ are the components of $X_{fn}^{\tau_n}$ which characterise the transaction[4]. We have $X_{fn}^{\tau} = \left(\hat{X}_{fn}^{\tau}, \tilde{X}_{fn}^{\tau}\right)$. The exogenous matching hypothesis allows us to decompose the conditional density $f\left(T_{fi}^{H}, \hat{T}_{fi}^{S}, w_{fi}^{\tau_i}, X_{fi}^{\tau_i} \middle| H_f^{\tau_{i-1}}, \vartheta_f\right)$ into two terms

$$f\left(T_{fi}^{H}, \hat{T}_{fi}^{S}, w_{fi}^{\tau_i}, X_{fi}^{\tau_i} \middle| H_f^{\tau_{i-1}}, \vartheta_f\right) = g_1\left(T_{fi}^{H}, \hat{T}_{fi}^{S}, w_{fi}^{\tau_i}, \hat{X}_{fi}^{\tau_i} \middle| H_f^{\tau_{i-1}}, \vartheta_f, \tilde{X}_{fi}^{\tau_i}\right) g_2\left(\tilde{X}_{fi}^{\tau_i}\right),$$

where $g_1\left(.,.,.,.,\hat{X}_{fi}^{\tau_i} \middle| H_f^{\tau_{i-1}}, \vartheta_f, \tilde{X}_{fi}^{\tau_i}\right)$ stands for the joint conditional density of the vector $T_{fi}^{H}, \hat{T}_{fi}^{S}, w_{fi}^{\tau_i}, \hat{X}_{fi}^{\tau_i}$ given $H_f^{\tau_{i-1}}, \vartheta_f, \tilde{X}_{fi}^{\tau_i}$ and $g_2(.)$ stands for the joint density for $\tilde{X}_{fi}^{\tau_i}$. Note that the density $g_2(.)$ does not depend on $H_f^{\tau_{i-1}}$ and/or ϑ_f. It is this independence of the distribution of $\tilde{X}_{fn}^{\tau}$ from $H_f^{\tau_{i-1}}$ and/or ϑ_f which is the exogenous matching assumption above.

The joint density becomes

$$f\left(H_f^{\tau_n} \middle| \vartheta_f\right) = \prod_{i=1}^{n} g_1\left(T_{fi}^{H}, \hat{T}_{fi}^{S}, w_{fi}^{\tau_i}, \hat{X}_{fi}^{\tau_i} \middle| H_f^{\tau_{i-1}}, \vartheta_f, \tilde{X}_{fi}^{\tau_i}\right) \prod_{i=1}^{n} g_2\left(\tilde{X}_{fi}^{\tau_i}\right).$$

This is obviously a strong restriction, which has testable implications. The empirical specification we present in the next section rests on these specific independence and

---

[4] Depending on the data at hand $\hat{X}_{fn}^{\tau_n}$ may be empty.

conditional independence restrictions. One possible test of the restriction can be constructed as follows. Note that the null hypothesis we are testing is:

$$g_2\left(\tilde{X}_{fi}^{\tau_i} \mid Z_f^{\tau_{i-1}}\right) = g_2\left(\tilde{X}_{fi}^{\tau_i}\right),$$

against some unspecified alternative, where $Z_f^{\tau_{i-1}}$ is a all, or a subset of , $\mathcal{G}_f, H_f^{\tau_{i-1}}$. For simplicity let us assume that $Z_f^{\tau_{i-1}}$ is a vector of dimension $(1 \times p)$, and that $\tilde{X}_{fi}^{\tau_i}$ is a vector of dimension $(1 \times q)$.

A possible route for a test is to specify an alternative hypothesis and then test the restriction to the null. A convenient specification for the conditional density under the alternative is then to assume that (at least locally) the conditional density can be written as:

$$g_2\left(\tilde{X}_{fi}^{\tau_i} \mid Z_f^{\tau_{i-1}}\right) = h\left(\tilde{X}_{fi}^{\tau_i} - Z_f^{\tau_{i-1}}\theta\right),$$

where θ is a matrix of dimension $(p \times q)$ and $h(x)$ is a multivariate joint density function.

As a consequence the null hypothesis of exogenous matching becomes a test for $\theta = 0$ against $\theta \neq 0$. A restriction to test is whether $\theta = 0$ is the maximum likelihood estimate for θ given the data at hand, i.e. $\tilde{X}_{fi}^{\tau_i}$ and $Z_f^{\tau_{i-1}}$. The pq conditional moment restrictions we are testing can be written as follows:

$$\mathrm{E}\left[Z_f^{\tau_{i-1}\prime} \otimes \frac{\partial \ln h\left(\tilde{X}_{fi}^{\tau_i} - Z_f^{\tau_{i-1}}\theta\right)}{\partial X'}\bigg|_{\theta=0}\right] = 0.$$

Moreover, if we assume that in the neighbourhood of $\theta = 0$, $\tilde{X}_{fi}^{\tau_i}$ is locally normally distributed with mean $Z_f^{\tau_{i-1}}\theta$ and variance covariance Ω:

$$\frac{\partial \ln h\left(\tilde{X}_{fi}^{\tau_i} - Z_f^{\tau_{i-1}}\theta\right)}{\partial X'} \propto \Omega^{-1}\left(\tilde{X}_{fi}^{\tau_i} - Z_f^{\tau_{i-1}}\theta\right)',$$

then the restriction becomes

$$\mathrm{E}\left[Z_f^{\tau_{i-1}\prime} \otimes \Omega^{-1}\tilde{X}_{fi}^{\tau_i\prime}\right] = 0$$

which, for non-singular Ω, is equivalent to

$$\mathrm{E}\left[Z_f^{T_{i-1}\prime} \otimes \tilde{X}_{fi}^{T_i\prime} \right] = 0 \, .$$

This test of the restriction can then be understood as equivalent to testing whether the $Z_f^{T_{i-1}}$ have some explanatory power in the regression of $\tilde{X}_{fi}^{T_i}$ on $Z_f^{T_{i-1}}$, and can be tested as a set of moment restrictions.

3. Data Description

Denmark's Integrated Database for Labour Market Research (IDA) covers the whole population of residents in Denmark and contains a longitudinally consistent link between employees, establishments and firms[5]. It comprises information from several comprehensive central government registers, and contains, amongst other things, demographic information, employment and income tax records. We draw a 2% sample of private sector firms in 1980 and follow them until 1995 with replacement representative of firm births. All newly hired workers in these firms over this period are sampled. We observe start and end date of employment within this period on a monthly basis, starting wage and some individual characteristics.[6]

The sample is split into four occupation groups: managers, non-managerial white collars, skilled blue collars and unskilled blue collars[7]. Though wage bargaining in Denmark has been increasingly decentralised since the 1980s, historically negotiations were organised roughly along industry-occupational lines[8]. These four occupations best capture the groupings within firm.

We measure the duration between hires within each occupation-firm. Should two or more hires occur in the same month, the duration between them is set to zero, since we are unable to be more precise. One from the bunch of hires is picked at random to be the first, and the duration since last hire for that observation will at least be until the previous period. This is arbitrary, but captures the bunching (29.3%) of observations. Employee characteristics within the occupation-firm are summarised each month that someone is hired. After these summary statistics for the whole occupation-firm have been calculated, we restrict the estimation sample by discarding some observations from the largest firms. If an occupation-firm cell

---

[5] Henceforth we use the generic term "firm" to signify a distinct legal business unit.

[6] For the purpose of calculating summary statistics of firm-occupation composition, which are used as covariates later, we also sample all incumbent workers.

[7] Occupation is defined as the first occupation code for each employment spell. Only 2.1% of spells contain different occupation codes.

[8] See Neumann, Pedersen and Westergaard-Nielsen (1991) and Barth & Zweimuller (1995) for background.

contains 100 employment spells or less, then all spells are kept. Otherwise we sample a random 100 spells for that firm-occupation[9].

It is informative to describe in more detail how each of our three variables of interest are calculated from the raw data, as calculation is not direct and coding and measurement decisions need to be justified with a view to their implications for the modelling exercise.

Employment spells data are derived from records of mandatory pension contributions (ATP). These records comprise a monetary amount paid by employers on behalf of employees to a central fund for a specific period. The amount is a step function of hours worked and the contribution period is weekly, fortnightly or monthly, depending upon the industry and accounting convention of the employer. All employer-employee matches are recorded for each contribution period for which there has been 10 or more hours of work. In principle all such matches can be dated with at least monthly precision. The ATP data that we have access to contains only a subset of matches for each individual: primary and secondary November attachments and primary non-November. A primary attachment at a given point of time is that which gives the employee the highest labour earnings during the calendar year of all attachments at that point[10]. Our three matches capture 77.0% of all spells and 95.7% of all hours worked. Of the three matches 76.0% are primary November and 7.7% are primary non-November. The employment spells that we do not have access to, and those that we could not observe even with full ATP information, would lead to a reduction in the measured duration between two hires for an employer. If unavailable ATP spells were distributed randomly, mean duration would fall from 6.2 to 5.5 months. However, the missing spells are by definition those which are financially less rewarding and shorter: the 23.0% of ATP spells which are unavailable represent 2.5% of hours worked[11].

We are able to date beginning and end points of each attachment provided we make some assumptions. The calendar month is treated as the minimum unit of observation across the sample and any additional information about hours or weeks that may be available for sub-samples is ignored. In any case we are unable to

---

[9] This amounted to sampling 100 spells from 0, 2, 2, 6 firms (out of 273, 427, 357, 385) of mean firm-occupation cell size 0, 373, 191, 251 from managers, white collar non-managerial workers, the skilled and unskilled respectively.

[10] The November/non-November distinction reflects the annual week 48 census point for employment statistics when the data is collected.

[11] It is essentially arbitrary wherever we set the threshold defining the amount of work that constitutes an attachment. We have used the greatest level of detail available to us, but are largely left to argue that the small proportion of labour input that is missing is ignorable. In preliminary estimation the results were found insensitive to dropping spells which were worth less than one week full-time equivalent contributions, of those currently counting as one month. Dropping all spells less than two months begins to significantly change duration coefficients.

observe attachment periods where less than 10 hours per week are worked with the firm. Furthermore, we cannot rule out the possibility that two observations of the same employer-employee match which are observed several months apart are in fact part of a single continuous spell. We make the arbitrary assumption that spells are distinct if separated by a whole calendar year[12], and as a consequence any such short spells within a year are aggregated into one longer one

Monthly wage is calculated from annual labour earnings due to a particular worker-employer match divided by months worked in the match during the year. Months worked is calculated from ATP records, which are censored. Consider the example of a weekly contribution period, where contributions are a step function of hours: 10-19, 20-29, 30-39. No contributions are recorded for those working below 10 hours, and no additional contributions beyond full-time are recorded for overtime work of 40 hours or more. The measurement error induced by using this proxy for hours to calculate monthly wages is certainly non-classical, and addressing these measurement error issues is beyond the scope of this paper. We have compared survey information on hours of work and hourly wages with our approximations, calculated from a different sub-sample drawn from the same register. For a sample of medium to large employers for which a comparison was possible, we find that the rank correlations between the monthly wages in the two datasets is 0.83 and between hours it is 0.79. From this simple exercise we conclude that our wage calculation is not error free, but the magnitude of this error is low.

Occupation is obtained from an annual census of employers and continuous tax records of earnings. At this level of aggregation the groups are well defined as they correspond to distinctions between contributions to different unions, unemployment insurance funds, pension and holiday funds, which the employer administrates.

Firm births and deaths are not considered explicitly here. The condition for a spells' inclusion in the estimation sample as a unit of observation is that we are able to observe at least one previous hire. This allows us to measure lagged dependent variables and aggregates at the firm-occupation level at the time of last hire. Note that it is of no consequence for the model we propose below that the first two transactions we observe for each firm may coincide with the firm's birth. Similarly, for firm death, no distinction is made between a firm which does not hire again before the end of the sample period and a firm which is no longer in business.

Figure 2 presents the starting wage distribution by occupation. Outliers are defined as the top 0.5% of wages, which are recoded to this level[13]. 4.1% of wages are

---

[12] See Westergaard-Nielsen and Jensen (1990) for a discussion of the importance of temporary layoffs in the Danish labour market.

[13] Preliminary analysis, where outlying wages were recoded to missing, gave similar results to those where it was top coded.

within 1% of the agreed minimum wage, and these are omitted from the wage distribution illustration (respectively 0, 4.2, 3.4, and 6.1% for managers, white collars, skilled and unskilled blue collars). Otherwise the distributions appear close to the superimposed log normal.

Figure 3 shows the distribution of durations since last hire. Clockwise from managers in the top left, bunched hires (the leftmost bar on each graph) are 32.5, 28.5, 26.8 and 29.9% of all hires. Mean time since last hire is 5.2, 5.7, 5.7 and 4.4 months respectively. These differences are modest and correspond to the relative sizes of the occupations, in that larger firm-occupation cells hire more frequently.

Figure 4 presents job length distributions. 13.2% of these spells are right-censored (respectively 14.9, 15.6, 13.2 and 10.1% for managers, white collars, skilled and unskilled blue collars). There are only modest reporting spikes at multiples of 12 months. Denmark is amongst OECD countries with the shortest employment lengths (OECD, 1997) and our measures of mean completed tenure reflect this. We find respectively 32, 33, 31 and 27 months.[14]

Figure 5 presents the number of job spells observed per occupation-firm. 15-22% of workers are from firm-occupations where we sample 100 spells, whereas 2% of spells are singletons within firm-occupation.[15] Clockwise from managers in the top left to unskilled blue collars at the bottom left, the mean number of observations per firm is 11.2, 9.9, 10.5 and 13.0.

Table 1 presents a simple analysis of variance for the outcome variables for each occupation separately. The R-squared cells report goodness of fit from a regression of the outcome variable on firm dummies. For starting wages and job length, where we expect individual heterogeneity to dominate, firms account for between a quarter and a third of the variance, whereas about half of the variance in time since last hire is explained by firm heterogeneity. The importance of firm-wise heterogeneity in all outcomes is therefore substantial and of similar magnitude across occupations.

The raw correlation between starting wage and time since last hire is negative for all occupations, and with the exception of skilled workers, job length is positively correlated with the starting wage and negatively with time since last hire. The empirical model which is presented in the next section moves on from these raw

---

[14] These figures are somewhat lower (10-15%) than those usually reported. This is because we consider spells as short as one month, which are not systematically included in the OECD (1997) figures, and because of the censoring which we have made explicit.

[15] These are singletons in the sense that there is one spell remaining after dropping the first spell to calculate lagged dependent variables.

correlations to focus on examining the role of worker turnover in wage determination.

4. Empirical Model

In this section, for the sake of exposition, we set up the basic homogeneous model and extend this to the heterogeneous case. The section concludes with a practical description of the estimation procedure.

4.1. Homogenous Model

A set of firm characteristics x_f (as well as characteristics of the employees in the firm at that time) is measured in $\tau_0 \equiv \tau - T_{fi}^{H}$. We assume that the observed characteristics of the worker to be hired, as measured at the start of his/her employment spell, z_{fi}^{τ}, are exogenous to the duration between hires. Assuming otherwise would require that we provide a model of what constitutes a successful match. In particular, to understand the match we would have to compare the current characteristics of the hired worker with the characteristics of the "average" worker that the firm samples from. This latter information is not available to us.[16] This assumes that firms match with workers in a way, which is unrelated to the observed or unobserved characteristics of the worker or of the firm.

To relate this notation to our earlier section on our empirical methodology, $x_f^{\tau_0}$ captures all the relevant information that we assume is contained in $H_f^{\tau_0}$ while z_f^{τ} collects all the information in $\tilde{X}_{fi}^{\tau}$. In the current homogeneous context we assume that

$$g_1\left(.... \mid H_f^{\tau_0}, \vartheta_f, \tilde{X}_{fi}^{\tau}\right) = g_1\left(.... \mid H_f^{\tau_0}, \tilde{X}_{fi}^{\tau}\right),$$

i.e. there are no firm specific effects that matter.

The data is such that we observe K_1 distinct discrete durations between hires and K_2 different employment durations. Our model of joint duration is therefore based on a discretisation of some continuous time model. We assume that the observed data can be deduced from the appropriate transformation of the following latent model:

$$\ln\left(\mathfrak{I}_{fi}^{H}\right) = x_{fi}^{\tau_0}\beta + \varepsilon_{1,fi},$$
$$\ln\left(\mathfrak{I}_{fi}^{S}\right) = x_{fi}^{\tau_0}\lambda + z_{fi}^{\tau}\eta + \varepsilon_{2,fi},$$
$$\ln w = z_{fi}^{\tau}\gamma + \varepsilon_{3,fi},$$

where $\mathfrak{I}_{fi}^{H}$ and $\mathfrak{I}_{fi}^{S}$ are latent durations. We assume further that given $\left(x_{fi}^{\tau_0}, z_{fi}^{\tau}, \tau, \tau_0\right)$ the joint distribution of the vector of unobserved components $\left(\varepsilon_{1fi}, \varepsilon_{2,fi}, \varepsilon_{3,fi}\right)$ is a multivariate normal with mean zero and variance covariance Σ (the element of which will be made explicit below). Furthermore, we assume the unobserved

---

[16] This is an important point. We provide a test for the exogenous nature of the match in the presentation of our results below.

components are identically and independently distributed given all the past transactions. The observations are derived from a transformation of the latent model according to the following scheme:

$$T_{fi}^{H} = k,\ k \in \{1...K_1\} \text{ if and only if } \delta(k-1,\tau) \le \ln\left(\mathfrak{I}_{fi}^{H}\right) < \delta(k,\tau),$$

$$T_{fi}^{S} = k,\ k \in \{1...K_2\} \text{ if and only if } \zeta(k-1,\tau) \le \ln\left(\mathfrak{I}_{fi}^{S}\right) < \zeta(k,\tau),$$

where $\delta(k,\tau)$ and $\zeta(k,\tau)$ are thresholds function which depend on the observed duration k and on the calendar time τ. Consistency requires that for any τ, $\delta(0,\tau) = -\infty$, $\zeta(0,\tau) = -\infty$, $\delta(K_1,\tau) = +\infty$, $\zeta(K_2,\tau) = +\infty$ and $\delta(k,\tau) < \delta(k+1,\tau)$, $k < K_1$, and $\zeta(k,\tau) < \zeta(k+1,\tau)$, $k < K_2$.

Although we could parameterise $\delta(k,\tau)$ and $\zeta(k,\tau)$ in terms of $(K_1 - 1) \times \#\{\text{dates observed}\}$ and $(K_2 - 1) \times \#\{\text{dates observed}\}$ independent quantities, for the sake of parsimony we specify $\delta(k,\tau)$ and $\zeta(k,\tau)$ somewhat more. First, we summarize the dependence on calendar time, τ, in the (possibly vector valued) functions $g_H(\tau)$ and $g_S(\tau)$. Hence the components of $g_H(\tau)$, say, could be: calendar time, calendar time squared, month of the year, etc. Second, the behaviour of the threshold with respect to k, everything else equal, can be summarized in terms of simple transformations of k. Unfortunately, although obvious candidates, for example $\delta_0 + \delta_1 \ln(k)$, $\delta_0 + \delta_1 k$, or even more generally $\delta_0 + \delta_1 k^{\delta_2}$, provided $\delta_1 > 0$ and $\delta_2 > 0$, are easily shown to satisfy the monotonicity requirement above, early empirical analysis showed that these simple threshold functions do not represent the data in an acceptable way (i.e. a likelihood ratio test against a model where each threshold is a parameter is always rejected). Furthermore, parsimonious flexible forms satisfying the monotonicity requirement are difficult to obtain[17].

We propose the following flexible solution. Note $P(k)$ a polynomial of k of a certain order, say q. Clearly the function such that $d(1) = \gamma_0$, and $d(k) = \gamma_0 + \sum_{i=0}^{k} \exp[P(i)]$, for all $2 \le k \le K-1$, is increasing in k, since $d(k) - d(k-1) = \exp[P(k)]$ is positive by construction. In practice $P(k)$ can be written as a linear form of some well-known family of polynomials. In this application we use the Chebyshev family,[18] $T_n(x)$ defined over $[-1,1]$ with values in$[-1,1]$. For all $2 \le k \le K-1$, $P(k)$ can then be written as:

$$P(k) = \sum_{p=0}^{q} \gamma_{p+1} T_p\left(2\left(\frac{k-2}{K-3}\right) - 1\right),$$

---

[17] Flexibility of the threshold function is directly related to the flexibility of the underlying base line hazard. Van den Berg (1999) argues that a well specified base line hazard function is paramount.

[18] See Press, et al (1993) for a presentation, a discussion of properties and some related numerical methods.

where the parameters $\{\gamma_p\}_{p=0}^{q+1}$ are to be estimated from the data. In what follows, we will write $d_q(k,\Gamma) = \gamma_0 + \exp\left[\sum_{p=0}^{q} \gamma_{p+1} T_p\left(2\frac{k-2}{K-3} - 1\right)\right]$, where $\Gamma = (\gamma_0, \ldots, \gamma_{q+1})$ collects the parameters to be estimated.

In the empirical application we restrict ourselves to the following specifications of the thresholds $\delta(k,\tau)$ and $\zeta(k,\tau)$:

$$\delta(k,\tau) = d_{q_1}(k,\Gamma_H) + g_H(\tau)'\,\Delta_H,$$

$$\zeta(k,\tau) = d_{q_2}(k,\Gamma_S) + g_S(\tau)'\,\Delta_S,$$

for some values q_1 and q_2, where $\Gamma_H, \Gamma_S, \Delta_H$ and Δ_S are conformable vectors of parameters.

Clearly we do not follow the conventional literature on modelling duration data, which relies on proportional hazards. However, the literature is not without examples of approaches similar to ours. For example Han & Hausman (1990) and Meyer (1990) provide two examples of a specification for a piecewise linear hazard rate consistent with the proportional hazard specification in the single spell univariate case, as well as in the competing risk single spell univariate case. Here, we follow the spirit of Han & Hausman's flexible approach. Lillard (1993) discusses simultaneous hazard estimation in a context related to ours.

We use a bivariate normal distribution as the basis for a more flexible description of the distribution of the durations since it allows for the introduction of correlations between the two durations and the hiring wage in a straightforward manner.[19] The correlations themselves can be understood as characteristics of the match between worker and firm.

Given $\left(x_{fi}^{\tau-k_1}, z_{fi}^{\tau}, \tau - k_1, \tau\right)$, the observed duration between hires, T_{fi}^{H}, the employment duration, T_{fi}^{S}, and the starting wage w_{fi}^{τ} are then jointly distributed (for all observations where the durations are above zero and less than the maximum observed[20]) as follows:

---

[19] Abowd and Kang (2002) present an alternative model where an exponential duration and a normal random variable are jointly distributed. The extension of their specification to our case is possible

[20] The expressions for durations of zeros or at the maximum are easily derived, and are available from the authors upon request. The expressions for right truncated employment durations and for unobserved starting wage are similarly straightforward to evaluate.

$$\Lambda_{fi}(\theta) \equiv \Pr\left[\left\{T_{fi}^{\mathrm{H}} = k_1\right\} \cap \left\{w_{fi}^{\tau} = w\right\} \cap \left\{T_{fi}^{\mathrm{S}} = k_2\right\} \mid x_{fi}^{\tau-k_1}, z_{fi}^{\tau}, \tau\right] =$$

$$\left\{\Phi_2\left[\mathrm{u}\left(k_1, \tau - k_1, x_{fi}^{\tau-k_1}, z_{fi}^{\tau}; \theta\right), \mathrm{v}\left(k_2, \tau, x_{fi}^{\tau-k_1}, z_{fi}^{\tau}; \theta\right), \rho_{HS/w}\right] - \right.$$

$$\Phi_2\left[\mathrm{u}\left(k_1 - 1, \tau - k_1, x_{fi}^{\tau-t_1}, z_{fi}^{\tau}; \theta\right), \mathrm{v}\left(k_2, \tau, x_{fi}^{\tau-k_1}, z_{fi}^{\tau}; \theta\right), \rho_{HS/w}\right] -$$

$$\Phi_2\left[\mathrm{u}\left(k_1, \tau - k_1, x_{fi}^{\tau-t_1}, z_{fi}^{\tau}; \theta\right), \mathrm{v}\left(k_2 - 1, \tau, x_{fi}^{\tau-k_1}, z_{fi}^{\tau}; \theta\right), \rho_{HS/w}\right] +$$

$$\left.\Phi_2\left[\mathrm{u}\left(k_1 - 1, \tau - k_1, , x_{fi}^{\tau-t_1}, z_{fi}^{\tau}; \theta\right), \mathrm{v}\left(k_2 - 1, \tau, x_{fi}^{\tau-k_1}, z_{fi}^{\tau}; \theta\right), \rho_{HS/w}\right]\right\}$$

$$\frac{1}{\sigma_w}\phi\left[\frac{\ln w - z_{fi}^{\tau}\gamma}{\sigma_w}\right].$$

where the quantities $\mathrm{u}(.;.)$ and $\mathrm{v}(.;.)$ are defined as:

$$\mathrm{u}\left(k_1, \tau - k_1, x_{fi}^{\tau-k_1}, z_{fi}^{\tau}; \theta\right) \equiv \mathrm{u}\left(k_1\right) =$$

$$\frac{1}{\sqrt{1-\rho_{Hw}^2}}\left(\delta\left(k_1, \tau - k_1\right) - x_{fi}^{\tau-t_1}\beta - \rho_{Hw}\left(\ln w - z_{fi}^{\tau}\gamma\right)\frac{1}{\sigma_w}\right),$$

and

$$\mathrm{v}\left(k_2, \tau, x_{fi}^{\tau-k_1}, z_{fi}^{\tau}; \theta\right) \equiv \mathrm{v}\left(k_2\right) =$$

$$\frac{1}{\sqrt{1-\rho_{Sw}^2}}\left(\zeta\left(t_2, \tau\right) - x_{fi}^{\tau-t_1}\lambda - z_{fi}^{\tau}\eta - \rho_{Sw}\left(\ln w - z_{fi}^{\tau}\gamma\right)\frac{1}{\sigma_w}\right),$$

where θ collects the parameters. Note that the parameters of the latent durations are only identified up to scale, and in the expressions proposed above we have normalised the variances of the latent (log-) durations to unity. The matrix Σ is then such that:

$$\Sigma = \begin{pmatrix} 1 & \rho_{HS} & \rho_{Hw}\sigma_w \\ \rho_{HS} & 1 & \rho_{Sw}\sigma_w \\ \rho_{Hw}\sigma_w & \rho_{Sw}\sigma_w & \sigma_w^2 \end{pmatrix}.$$

Finally, the parameter $\rho_{HS/w}$ is such that $\rho_{HS/w} = \dfrac{\rho_{HS} - \rho_{Hw}\rho_{Sw}}{\sqrt{1-\rho_{Hw}^2}\sqrt{1-\rho_{Sw}^2}}$, where ρ_{HS}, ρ_{Hw} and ρ_{Sw} are the correlations of the latent trivariate normal random unobserved components. The expression above forms the basis for the contribution to the likelihood for each observed "transaction" (i.e. between hire duration, starting wage and employment duration).

The correlations ρ_{HS}, ρ_{Hw} and ρ_{Sw} appear in at least one expression of interest: the conditional expectation of the wage given a duration between hires and a time to separation, which would give the answer to the question "an individual is observed to have been hired after k_1 periods since the last hire and remained with the firm for k_2 periods, what is the likely starting wage he/she received ?". Indeed we have:

$$\mathrm{E}\Big[\ln w \mid \Big\{T_{fi}^{\mathrm{H}} = k_1\Big\} \cap \Big\{T_{fi}^{\mathrm{S}} = k_2\Big\}, x_{fi}^{\tau-k_1}, z_{fi}^{\tau}, \tau\Big] =$$

$$z_{fi}^{\tau}\gamma + \sigma_w \frac{\rho_{Hw} - \rho_{HS}\rho_{Sw}}{1-\rho_{HS}^2} \lambda_H\left(\tilde{\mathrm{u}}\left(k_1\right), \tilde{\mathrm{u}}\left(k_1 - 1\right), \tilde{\mathrm{v}}\left(k_2\right), \tilde{\mathrm{v}}\left(k_2 - 1\right)\right)$$

$$+ \sigma_w \frac{\rho_{Sw} - \rho_{HS}\rho_{Hw}}{1-\rho_{HS}^2} \lambda_S\left(\tilde{\mathrm{u}}\left(k_1\right), \tilde{\mathrm{u}}\left(k_1 - 1\right), \tilde{\mathrm{v}}\left(k_2\right), \tilde{\mathrm{v}}\left(k_2 - 1\right)\right),$$

with: $\tilde{\mathrm{u}}\left(k_1\right) = \delta\left(k_1, \tau - k_1\right) - x_{fi}^{\tau-k_1}\beta$,

$$\tilde{\mathrm{v}}\left(k_2\right) = \zeta\left(k_2, \tau\right) - x_{fi}^{\tau-k_1}\beta - z_{fi}^{\tau}\eta \,.$$

The functions $\lambda_H(.,.,.,.)$ and $\lambda_S(.,.,.,.)$ are bivariate equivalents of the usual inverse Mill's ratio[21] in selection models. In the present context, however, they do not have the interpretation of a correction for selectivity since the dependent variables are generally observed.

The model therefore allows for a dependence of the starting wage on the rate at which hires and separation occurs. In a frictionless labour market we would not expect the wage to be associated with the hiring or the separation processes. i.e. the wage rate should be independent of the time it took for the worker and the firm to meet, and the subsequent employment duration. A test for a labour market with this feature, given homogeneity, would be to test whether both $\frac{\rho_{Hw} - \rho_{Sw}\rho_{HS}}{1-\rho_{HS}^2} = 0$ and $\frac{\rho_{Sw} - \rho_{Hw}\rho_{HS}}{1-\rho_{HS}^2} = 0$. These two hypotheses are equivalent to the hypotheses $\rho_{Sw} = 0$ and $\rho_{Hw} = 0$ (provided $\rho_{HS}^2 < 1$). In this homogenous context, if the alternative is accepted the wages are determined to some extent by the hiring firm.

Other quantities of interest can easily be calculated, although not always analytically. For example it is straightforward to calculate the conditional expectation of the employment duration given the starting wage and the duration between hires, we have:

$$\mathrm{E}\Big[T_{fi}^{S} \mid \Big\{T_{fi}^{H} = k\Big\} \cap \Big\{w_{fi}^{\tau} = w\Big\}, x_{fi}^{\tau-k_1}, z_{fi}^{\tau}, \tau\Big] =$$

$$\sum_{t-1}^{K_2} t \frac{\Pr\Big[\Big\{T_{fi}^{H} = k_1\Big\} \cap \Big\{w_{fi}^{\tau} = w\Big\} \cap \Big\{T_{fi}^{S} = t\Big\} \mid x_{fi}^{\tau-k_1}, z_{fi}^{\tau}, \tau\Big]}{\Pr\Big[\Big\{T_{fi}^{H} = k_1\Big\} \cap \Big\{w_{fi}^{\tau} = w\Big\} \mid x_{fi}^{\tau-k_1}, z_{fi}^{\tau}, \tau\Big]},$$

which can be calculated from the expressions above. This expression can be simplified in some special cases. For example if $\rho_{Sw} = 0$ and $\rho_{HS} = 0$ (which is equivalent to $\rho_{Sw} = 0$ and $\rho_{HS/w} = 0$ whenever $\rho_{Hw}^2 < 1$) it becomes:

$$\mathrm{E}\Big[T_{fi}^{S} \mid \Big\{T_{fi}^{H} = k\Big\} \cap \Big\{w_{fi}^{\tau} = w\Big\}, x_{fi}^{\tau-k_1}, z_{fi}^{\tau}, \tau\Big] =$$

$$\sum_{t-1}^{K_2} t \Pr\Big[\Big\{T_{fi}^{S} = t\Big\} \mid x_{fi}^{\tau-k_1}, z_{fi}^{\tau}, \tau\Big],$$

[21] Their expressions are complex, and are available from the authors upon request.

i.e. the expected employment duration is independent of the starting wage and of the duration between hires. Note that this can occur even if the starting wage depends on the duration between hires.

4.2. Heterogenous Model

Firms, however, may differ in ways, which are directly observable. For example, we can imagine that as part of an implicit, or even explicit (but unobserved to the econometrician) contract, the firm's reward to new hires is related to the time it took to recruit them, and that the reward is related to the length of time the firm expects new workers to remain with it. These correlations between the duration between hires, the duration in employment, and the starting wage are important characteristics of the firm's specific recruitment and pay policies as well as the conditions on the labour market specific to the firm. Theoretical arguments drawn, for example, from Burdett and Mortensen (1998) suggest that such policies may not be unique. Indeed, in this context a firm paying a relatively high wage insures a low separation rate and a high hiring rate and therefore a large firm size, while a relatively low wage implies the exact opposite, many separations, a low hiring rate and a small firm size. It is the existence of both strategies at equilibrium, which ensures a non-degenerate equilibrium wage distribution. However, it is unclear, at least empirically, whether these policies are the only ones followed. It may be the case that firms, all else equal, that pay a higher than average wage, hire relatively rarely and experience frequent separations (relative to other firms in the sample).

The data we use allows us to identify many workers within the same firm and therefore allows us to identify the components of the wage, time between hires and employment duration which are specific to the firm, whatever their causes. This leads us to introduce firm-specific effects in each component of the model (the model is therefore hierarchical with the firm cluster collecting many individual observations). Hence, here we reinstate the dependence of $g_1(.|.)$ on ϑ_f. This creates additional difficulties at the modelling and the estimation stage. In what follows we present a possible specification where we assume that the firm-specific effects collected in ϑ_f are distributed randomly across firms and independently from all other observables and unobservables (the usual random effects hypothesis).

An obvious way of doing this is to allow well-chosen constant terms in each equation to be firm specific. We therefore introduce three random firm specific elements into the model. The thresholds are now dependent on the firm type:

$$\delta_f(k,\tau) = \delta_{f,0} + d_{q_1}(k,\Gamma_H) + g_H(\tau)'\,\Delta_H,$$

$$\zeta_f(k,\tau) = \zeta_{f,0} + d_{q_2}(k,\Gamma_S) + g_S(\tau)'\,\Delta_S,$$

where $\delta_{f,0}$ and $\zeta_{f,0}$ are firm specific constant, and by convention we impose $\delta_{1,0} = \zeta_{1,0} = 0$ [22]. Everything else equal, larger values for $\delta_{f,0}$ or $\zeta_{f,0}$ imply that the firm hires (respectively experiences separations) more often. Finally we assume that the (un-conditional) mean log-wage is firm dependent:

$$\mathrm{E}\left[\ln w_{fi} \mid z_{fi}^{\tau}, \tau\right] = \gamma_{f,0} + z_{fi}^{\tau}\gamma ,$$

where z_{fi}^{τ} is understood to exclude a constant, and γ is reformatted accordingly.

We make the random effect assumptions that the specific effects are not correlated with the history of the collection of explanatories for all equations, and that each realisation of the effects is drawn from a trivariate (unspecified) distribution,

$$\psi(\delta,\zeta,\gamma) \equiv \mathrm{Prob}\left[\delta_0 < \delta, \zeta_0 < \zeta, \gamma_0 < \gamma\right].$$

It is the joint distribution of the firm specific effects, which describes all the elements of the firm's hiring, separation and starting wage policy. By definition, the firm specific effects are time invariant.

It is always possible to decompose the variance of log-wages in order distinguish between components specific to the variation between individuals, the variation between firms and a residual component that in our context can be attributed to frictions. Applying the Law of Iterated Expectations twice:

$$\begin{aligned}
\mathrm{V}\left[\ln w_{fi}\right] &= \mathrm{V}\left[\gamma\right] + \mathrm{V}\left[\varepsilon\right] \\
&= \mathrm{V}_f\left[\mathrm{E}\left[\ln w_{fi} \mid f\right]\right] + && \text{(firm effect)} \\
&\quad \mathrm{E}_f\left[V_{T_f^H, T_f^S}\left[\mathrm{E}\left[\ln w_{fi} \mid T_f^H, T_f^S, f\right] \mid f\right]\right] + && \text{(friction effect)} \\
&\quad \mathrm{E}_f\left[\mathrm{E}_{T_f^H, T_f^S}\left[\mathrm{V}\left[\ln w_{fi} \mid T_f^H, T_f^S, f\right] \mid f\right]\right], && \text{(individual effect)}
\end{aligned}$$

where the subscript f stands for the conditioning with respect to the firm specific random effects. The important term here is the friction effect which describes the expected (between firm) variance of the expected log-wages given between hires and employment durations. It calculates the variance of the mean log-wage among sub-populations with distinct between hires and employment durations for the many firm types. It is understood as a contribution of frictions to the variance of log-wages, since it is equal to zero if log-wages and both durations are uncorrelated given the firm specific effects. The last term aggregates over firms and durations the conditional variance of log-wages.

4.3. Estimation Method

---

[22] This defines the first group as the reference group. Alternatively, the firm specific effects can be understood as unobserved firm specific covariates (here dummies) that should belong to the appropriate list of regressors, here $x_{fi}^{\tau-k_1}$. In either case one group is understood as the reference group.

Note $\Lambda_{f.}\left(\delta,\zeta,\gamma;\tilde{\theta}\right)$ the likelihood for firm *f* (i.e. over all observation in *f*) given that the specific effects take the values (δ,ζ,γ), and where $\tilde{\theta}$ collects all the other parameters, i.e.:

$$\Lambda_{f.}\left(\delta,\zeta,\gamma;\tilde{\theta}\right)=\prod_{i\in f}\Lambda_{fi}\left(\delta,\zeta,\gamma;\tilde{\theta}\right),$$

where $\Lambda_{fi}\left(\delta,\zeta,\gamma;\tilde{\theta}\right)$ is the likelihood for observation *i* in *f*. Conditional on the firm specific effects this has the same expression as the individual likelihood for each individual observation in the homogeneous case presented above.

The estimation problem is the following. We want to estimate the parameters in $\tilde{\theta}$ as well as the unknown distribution $\psi(\delta,\zeta,\gamma)$. Our treatment of the difficulty amounts to the treatment of some trivariate unobserved heterogeneity. The literature suggests that the non-parametric estimation of the joint distribution is possible, although the Maximum Likelihood estimate will be a discrete distribution with a finite number of point of support.[23] Hence, without loss of generality, we assume at the outset that the population consists of *G* homogenous subgroups (a class). Each class, *g*, *g=1...G*, is characterised by the class parameters $\left(\delta_{0,g},\zeta_{0,g},\gamma_{0,g}\right)$ and by a class weight in the total population ϕ_g, $\phi_g>0$. We obviously require that $\sum_{g-1}^{G}\phi_g=1$. The estimation problem can be divided into two problems, first how many classes, *G*, does the data allow for, and second, given the number *G* of classes how to estimate the parameters of the model, i.e. $\tilde{\theta}$ and the collection $\left\{\left(\delta_{0,g},\zeta_{0,g},\gamma_{0,g}\right),\phi_g\right\}_{g-1}^{G}$.

The second problem, i.e. estimation of the parameters of the model with the number of classes fixed, is usually solved in the frequentist approach by way of the EM algorithm (following, for example, the lines of the argument for the univariate case presented in Lancaster, 1991). The contribution to the likelihood of the observed data for a generic firm *f* in this context is:

$$\sum_{g-1}^{G}\Lambda_{f.}\left(\delta_{0,g},\zeta_{0,g},\gamma_{0,g};\tilde{\theta}\right)\phi_g\ .$$

The EM algorithm proceeds instead by considering the likelihood as if the firm's class was known:

$$\prod_{g-1}^{G}\left\{\Lambda_{f.}\left(\delta_{0,g},\zeta_{0,g},\gamma_{0,g};\tilde{\theta}\right)\phi_g\right\}^{\mathbf{1}_{|f\text{ is of type }g|}}\ .$$

Taking logs and then calculating the expectation of the result given what is observed:

$$\sum_{g-1}^{G}\mathrm{E}\left[\mathbf{1}_{|f\text{ is of type }g|}\mid\Theta_f\right]\left\{\log\left(\Lambda_{f.}\left(\delta_{0,g},\zeta_{0,g},\gamma_{0,g};\tilde{\theta}\right)\right)+\log\left(\phi_g\right)\right\},$$

---

[23] See for example Heckman and Singer (1984) and Lancaster (1991).

where $\mathcal{O}_f$ stand for all the observed characteristics of firm f. Bayes' Rule allows us to evaluate $\mathrm{E}\left[\mathbf{1}_{[f \text{ is of type } g]} \mid \mathcal{O}_f\right]$ in terms of the terms $\Lambda_f\left(\delta_{0,g}, \zeta_{0,g}, \gamma_{0,g}; \tilde{\theta}\right)\phi_g$, we have:

$$\mathrm{E}\left[\mathbf{1}_{[f \text{ is of type } g]} \mid \mathcal{O}_f\right] = \frac{\Lambda_f\left(\delta_{0,g}, \zeta_{0,g}, \gamma_{0,g}; \tilde{\theta}\right)\phi_g}{\sum_{h-1}^{G} \Lambda_f\left(\delta_{0,h}, \zeta_{0,h}, \gamma_{0,h}; \tilde{\theta}\right)\phi_h}.$$

These quantities can be understood as the à posteriori probabilities for firm f to belong to type g. For a set of values of the parameters the EM algorithm proceeds first by evaluating these quantities (Expectation-step). This amounts to augmenting the observed data with some likely observations of the unobserved data (here the firm's type). Second, taking these as given, the algorithm maximises the expected likelihood with respect to the parameters $\tilde{\theta}$ and the collection $\left\{\left(\delta_{0,g}, \zeta_{0,g}, \gamma_{0,g}\right), \phi_g\right\}_{g-1}^{G}$ (Maximisation-step)[24]. These two steps are repeated until convergence. It can be shown (Gouriéroux & Monfort, 1995) that the EM algorithm is monotonous, and that it will provide the Maximum Likelihood estimate if it is unique.

The first problem that of allocating firms to classes, can be approached in many different ways. Böhning (1995) presents a review of the various methods available. Here we follow the Vertex Direction Method (VDM)[25].

To summarise we proceed as follows:

i. Estimate the homogenous model presented in the previous section.

ii. From these estimates search within an a priori defined 3-dimensional box using Böhning's VDM for the location parameters of the new class. The VDM suggest a starting value for this new group's share of the total population of firms. If there is no possible direction that improves the likelihood, then stop.

iii. Using the (Conditional) EM algorithm, find the maximum likelihood parameters $\tilde{\theta}$ and the collection $\left\{\left(\delta_{0,g}, \zeta_{0,g}, \gamma_{0,g}\right), \phi_g\right\}_{g-1}^{G}$.

iv. Calculate the asymptotic variance-covariance matrix[26] of the parameters given G.[27]

---

[24] In fact, we break this second stage into two sub stages. Indeed, it is possible to obtain the analytical expression for the update for $\{\phi_g\}_{g=1}^{G}$, given $\{(\delta_{0,g}, \zeta_{0,g}, \gamma_{0,g}), \phi_g\}_{g-1}^{G}$ and the other type invariant parameters from the previous EM steps. Given these values we maximise the expected likelihood as a function of all the other parameters. This amounts to following an extension of the EM algorithm, which is called Expectation/Conditional Maximisation described in Meng & Rubin (1993).

[25] Böhning presents several version of this algorithm. Here we use the improvement he proposes. This relies on a quadratic approximation of the directional derivative. In the context of the simulations we performed in order to validate our code, this proved to be effective.

[26] In the application on real data it turns out that the estimated Hessian calculated for the observed data at the EM estimates can be badly conditioned (in the sense that the ratio of the largest to the smallest eigenvalue is a large number) when the number of firm types becomes large. When this is the

v. Repeat from ii) above, until convergence of the likelihood function.

In practice we limit the algorithm to 12 classes maximum.

The model we propose here therefore allows for a flexible hazard for each individual duration of interest and for unobserved heterogeneity between firms. This is necessary to capture difference between firms in the labour market as well as to capture differences that may occur over time within a single firm but between individuals. In both cases misspecification of the baseline hazard and/or of the unobserved heterogeneity is likely to create spurious correlations between the two measures of turnover and the individual starting wage.

Furthermore, under the maintained hypothesis of conditional independence and given the assumption of exogenous matching, the set of conditioning variables can include lagged dependent variables. Thus our reduced form approach is akin to a SURE model in a regression context, where the parameters of interest are the correlation between the unobservable for each equation (given some set of regressors). Because we are not interested in the potential causal relationship between the contemporaneous dependent variables and since we observe the outcomes completely (except perhaps for the right censoring of employment durations) the usual identification issues do not arise. We therefore claim that under the assumptions described in section 2 the joint density of the two durations and of the starting wage is identified.

The next section discusses the estimation results.

5. Empirical Findings

The preferred empirical specification we present here depends on the complete set of conditioning covariates included (individual observables in $x_{fi}^{\tau-k_1}$ and z_{fi}^{τ}). Given the assumption of conditional independence for the transitory unobservable components, past observations of the time between hires, starting wages, tenure in the firm, as well as some characteristics of the work force in the firm (as they can be observed in $\tau-k_1$ and τ) are justifiably predetermined. Hence we condition on the (log) of the starting wage awarded for the last transaction, the (log) duration between hires and the (log) average tenure in the firm both measured when the last transaction took place ($\tau-k_1$). Including these variables allows for some autocorrelation between transactions within the same firm. Note that the relevant time index here is not calendar time but event time (time of a hire). Furthermore,

---

case we calculate an "approximate" inverse through Singular Value Decomposition "zero-ing" ratios involving negligible eigenvalues as suggested in Press et al. (1993)

[27] See Jamshidian and Jennrich (2000) and Oakes (1999) for formulae.

we include the mean education, gender and age, again measured at $\tau - k_1$. Given exogenous matching, changes in these characteristics mainly result from separation behaviour. However, it is possible that firms react to such changes by increasing (or decreasing) their hiring (or firing) effort. This is our main rationale for their inclusion in the list of conditioning variables (i.e. a firm may consider that an "ageing" stock of workers, say, increases the need for a new hire).

On the individual side, we include the conventional human capital determinants that we can observe: age, age squared, gender, years of education and years of education squared. Finally, we control for year, month-of-year, 11 regions and 6 industries by introducing dummy variables (not reported).

Selected parameter estimates and their standard errors are reported in Table 2. Estimates for the sub-sample of Managers are reported in columns 1-2, Non Managerial White Collars in columns 3-4, Skilled Blue Collars in columns 5-6, and in columns 7-8 for Unskilled Blue Collars. The rows are divided into five blocks, corresponding respectively to the measurements of $\beta, \gamma, \lambda, \eta$ and the elements of Σ.

We selected the number of types for each occupation according to the Akaike Information Criterion (AIC) whereby an increase in the value of the likelihood is penalised by the increase in the number of parameters. This has the effect of reducing the number of types that represent a small proportion of the population. The results shown are insensitive to an increase in the number of types. Note that determining the number of types with the AIC is likely to lead to over-fitting. However, while the Bayesian Information Criterion (BIC) would suggest a smaller number of types than the AIC, the maximum of the likelihood suggests even more types than the AIC[28]. In the example of managers, AIC suggests 9 types are appropriate, BIC indicates 4 types would be enough, while the likelihood value indicates that 12 (or more) types offer a better description of the data. Ultimately, we argue that the AIC provides an intermediate and credible alternative.

5.1. Employment and Wage Dynamics

Strikingly, in each equation, the previous transaction, as characterised by the log duration between hires and log starting wage, as well as the average tenure for the workers currently employed by the firm (these could be understood as the influence of lagged dependent variables) have small effects. The sign and magnitude of these parameters are not sensitive to the number of types, although in some cases their significance is. Indeed, reducing the number of types does in some cases establish significance. This is the case for parameters of the employment duration equation. This phenomenon is not unexpected, since ignoring unobserved heterogeneity is likely to create spurious statistical significance.

---

[28] See Mittelhammer, Judge and Miller (2000) for a discussion.

The duration between hires depends negatively on the starting wage previously paid, and positively on the previous duration between hires. However, in the case of Managers and Unskilled Blue Collars the starting wage paid previously does not have a significant effect. Furthermore, the average tenure of workers with the same occupation currently employed by the firm has no significant effect. The magnitude of these coefficients on lagged dependent variables is such that any shock will only affect the next hire[29]. Regarding other characteristics of the work force, mean education has a negative and significant effect on duration between hires (with the exception of skilled blue collars). The proportion of males and average age are of mixed significance.

Employment duration is largely unaffected by characteristics of the previous hire. The parameters are small in absolute value (never above 0.156 in absolute value, 0.120 for the statistically significant parameters). The coefficients of the lagged dependent variables are all significant for Unskilled Blue Collars only. For Managers, log mean tenure has a negative and significant effect, while for Skilled Blue Collars it is the log of the duration between hires for the previous transaction which is positive and significant. Note again that the degree of autocorrelation (as measured by the coefficient of the lagged variable) is surprisingly small.

Characteristics of workers in the firm at the time of last hire do not play a significant role in the explanation of employment duration, except in the case of Non Managerial White Collars and Unskilled Blue Collars. There, the average age of employees when the last transaction took place is negatively associated with employment duration. Individual age and education are significant too. Those recruits who are older, and have more schooling, exhibit longer employment durations. However, there is no significant effect of gender on employment duration.

The effects of the "lagged" dependent variables on starting wages are small in absolute terms and on the whole insignificant. The effect of the log of lagged starting wage is positive and of similar magnitude (0.1) for all occupations. It is significant for Managers and Non Managerial White Collars only. To interpret the nature of the parameter estimated, assume that the firm hires new employees every month. Assume further that these new employees are all identical. Our specification then implies that the starting wages between three consecutive hires are related as follows

$$\ln\left(\frac{w_t}{w_{t-1}}\right) = \alpha \ln\left(\frac{w_{t-1}}{w_{t-2}}\right).$$

At equilibrium where the firm's wage is constant, the above relationship is clearly verified, whenever $w_{t-1} = w_{t-2}$ then $w_t = w_{t-1}$. Finally it is easily shown that a

[29] Our results suggest that firm specific (permanent) effects are well captured by our model of unobserved heterogeneity. If it was not the case, we would expect values for these autoregression parameters to be closer to unity, i.e. reflecting the presence of a strong permanent effect.

general solution for the equation above is $\ln\left(\frac{w_t}{w_{t-1}}\right) \propto \alpha^t$. For small values of α this quickly converges to zero. Hence, our estimate for α of around 0.1 means that deviations from equilibrium starting wages are quickly dampened.

Log duration between hires at the time of the previous transaction has no significant effect on the current starting wage. Log mean tenure at the time of the previous hire has a small significant effect only on the starting wage of Non Managerial While Collars and of Unskilled Blue Collars, but with opposite signs. Finally, the composition of the workforce has no significant effect on the starting wage except for Managers, where the average age, the average education and the proportion of males in the firm are significant and negative.

Individual characteristics have the expected effect on the starting wage. Males receive a significant premium. Age is significant for all occupations, whereby older workers (everything else equal) receive higher starting wages. Education however has no significant effects, except for Non Managerial White Collars. This is largely a consequence of conditioning on occupation groups.

5.2. Worker-Firm Match Effects

Given the firm specific random effects, there is a positive correlation between unobservables that determine the employment duration and the duration between hires. Indeed the point estimates are always positive, and are insignificant only among Managers. The correlations between individual unobservables which determine the starting wage and the duration between hires do not exhibit any clear pattern between occupations. The correlations are significant for Managers and Skilled Blue Collars (positively so for the former and negatively so amongst the latter), and are not significant in the other cases. Hence, all else equal, while amongst Managers the time between hires and starting wages are positively associated, amongst Skilled Blue Collars they are negatively associated.

There is some evidence at the individual level, given the firm type, that the unobservables which determine the starting wage and the employment duration are significantly correlated, negatively so for White Collars and Skilled Blue Collars, positively so for Unskilled Blue Collars. For Managers the correlation is small and not significant. In any case this suggests that, since we control for unobserved firm-specific effects, unobserved match effects between firms and workers (as captured by the correlation between unobservables) are small.

To summarise, individual correlations between unobservables are small (less than 0.1) in all cases and the pattern of significance depends mainly on occupation. This implies that, for all occupations, the share of the friction effect in the decomposition of the variance of starting wages is very small.

We now turn to a description of the joint distribution of these firm specific effects.

5.3. Firm Heterogeneity

In each occupation the location and relative sizes of the firm types were jointly estimated. We attempt to characterise the distribution of the types graphically in figures 6a, 6b, 7a, 7b, 8a, 8b and 9a, 9b. In the top row of figures *k*a we describe the hazard[30] and the probability density function (represented with log scale on the y axis) of the duration between hires. The bottom row describes the hazard and probability density function of employment durations. The continuous line represents the appropriate function for the exponential distribution, which shares the same average duration. The dotted lines represent the appropriate function for each of the firm types that the estimation procedure identified and measured. The dashed line represents the mean hazard or probability density function for each different type.

Flexible modelling of the base line hazard is important. In practice, we selected the order of the polynomial after separate univariate analyses of the durations without accounting for heterogeneity. The order is selected such that it is the first order that cannot be rejected against a higher order. The orders for the polynomials are given in Table 2. Generally, the base line hazard is not constant. As a consequence the durations do not follow an exponential distribution. Note that in many instances the hazards are first increasing (over durations less than 3 to 4 months) and then decreasing. In the case of employment duration, this suggest that after an initial period, when the firm and the individual assess the match and part company if it is not found to be satisfactory, employment duration exhibits negative duration dependence over a large range (up to about 48 months). This pattern is mitigated by occupation specific factors and by firm heterogeneity. Indeed, in some cases the initial period of positive duration dependence lasts longer (consider Managers or Non Managerial White Collars), while in other cases the range that exhibits negative duration dependence is shorter.

Heterogeneity is clearly important. These differences are notable for small durations, in particular those less than 24 months for the time between hires and less than 50 months for employment durations. It is for these small durations that the samples are most informative (consider for example the rate at which the probability density function decreases as the duration increases). Firms of different types can experience and/or determine widely different instantaneous hiring and separation rates.

Figures *k*b illustrate the relationship between heterogeneity in the labour market in terms of time between hires, employment duration and the starting wages. Each pane is in fact a scatter plot of some of the firm class-specific characteristics (i.e.

[30] This is the ratio of the probability of observing a hire or a separation in period *j* over the probability of not observing such an event at least until period *j*. The probabilities are evaluated at the sample mean of observed characteristics.

transformations of the parameters $\left(\delta_{0,g}, \zeta_{0,g}, \gamma_{0,g}\right)$ evaluated at the sample mean). Each type is a point in this diagram. We represent the relative size of each type by a circle, which is centred on the type's position. The circle diameter is indicative of each type's relative size. The first pane (top left) represents the distribution of employment duration and duration between hires (both in logs) across firm types. The second pane (top right) represents the distribution of duration between hires (in logs and on the same scale as in the previous pane) and the starting wage mark-up above the overall mean. The third pane (bottom right) describes the same information this time between employment duration and starting wage mark-up. In each pane using the estimated values for the locations of the types and their relative sizes, we draw the linear regression lines between the characteristics on the x-axis and the characteristics on the y-axis, the continuous line, (and the reverse regression, the dashed line). In each pane, we indicate below the x-axis the regression coefficients. These regressions are one way to describe the association between the x-axis characteristic and the y-axis characteristic across firm types/labour markets.

Heterogeneity exists in all dimensions[31] (durations and starting wages) as well as in all occupations. In general, no type dominates to the point of making the other types insignificant (for example the main type is not always near the mean of the types as indicated by the intersection of the regression lines). Note that between types the difference can be large. For example, starting wages vary from about 10% to above 15% around the mean. Furthermore, between types, time between hires and employment durations can increase three-fold.

Employment duration and time between hires are always negatively associated (as measured by the regression lines). One measure of this association is elasticity of employment duration with respect to the time between hires. These elasticities evaluate the rate at which on average the employment duration changes relative to the time between hires as we move from one firm type to another. Hence we may conclude that firm types or labour markets where firms hire workers at a lower rate tend on average to keep their workers for shorter spells.

For all occupations, the regression lines suggest that an increase in the duration between hires, as we move from one firm type to another, is associated with a lower starting wage. The corresponding between firm types wage "elasticities" with respect to the duration between hires are small and about -0.08, -0.04, -0.08 and -0.04 for Managers, White Collars, Skilled and Unskilled Blue Collars respectively.

The association between the employment duration and the starting wage can be characterised in a similar way. For most occupations (Managers are the exception) the starting wage is negatively associated with the employment duration and the relevant between firm types wage "elasticities" are small. We find 0.02, -0.03, -0.08

[31] Increasing or decreasing by one the number of types does not change the results qualitatively.

and -0.15 for Managers, White Collars, Skilled and Unskilled Blue Collars respectively.

5.4. Goodness-of-Fit

In all cases the regression lines capture very little of the variability between firm types or labour markets. Although it is clear from the figures *k*b that the firm specific effects are not independently distributed, (the hiring, separation and wages policies are not independent), no obvious pattern presents itself to us.

Though it is difficult to explain the firm effects, it is informative to show the proportion of starting wage variance that can be explained jointly by the endogenously determined firm types and the firm specific time invariant observed characteristics. Table 3 presents the evidence. For each occupation separately and for each of our dependent variables (Starting Wage, Job Length, Last Hire) we regress the individual observations on the set of firm dummies (column headed ANOVA), on the posterior probabilities (column headed P_i), on the posterior probabilities and firm specific time invariant covariates (column headed P_i+X_i). We present the unadjusted R-squared values of these regressions. The results from the ANOVA are clearly the "best" that can be obtained and therefore the explanatory power of the model is judged relative to this saturated model.

Between occupations, the firm's fixed effect regressions explain around 20% of the observed variation of employment durations, about 30% of the variance of starting wages, and around 50% of the variance of duration between hires. By contrast, if we regress observed employment durations on the posterior probabilities only, or on the posterior probabilities and the firm specific time invariant covariates, we explain respectively 8% and 12% of the total observed variance (similar figures on average are 9% and 17% respectively for starting wages, and 30% and 33% respectively for duration between hires). We also calculate the size of the variance that our restricted set of variables explains relative to the size of the variance that a full set of firm dummies explains on their own (computing the ratio of R-squared measures gives the required figures). We find that our set of firm covariates together with the posterior probabilities explain between 50% and 66% of the variance that is explained by the firm dummies. Unsurprisingly these figures tally with those obtained if one attributes to each individual observation the firm average (time invariant mean) for the dependent variable under study and regresses these values on the firm specific covariates and posterior probabilities. Hence, while a substantial proportion of the variance of our dependent variables is not (and cannot be) captured by the firm specific effect, out of the variance that can be captured, our model accounts for a large proportion of the heterogeneity. Note further that firm specific covariates do not make the greatest contribution to the explanatory power of the model, since they account for about a third of the proportion of explained variance. In this sense we feel that our model "fits" the data satisfactorily.

Finally we present a test of our hypothesis that the match between firms and workers is exogenous. Following our discussion in section 2 it is sensible to expect that under the null of exogenous matching the posterior prediction for firm types as well as firm specific characteristics should not be associated with the characteristics of the newly hired workers. The test statistics for each occupation are reported in table 4. The hypothesis of exogenous matching can not be rejected. However, this may be a consequence of the poor discriminatory power of the included observed explanatory variables which would reduce the power of the test.

5.5. Firm Size and Location

Our model of the process of hires and separations should determine firm size. The results we present above show that firms are heterogeneous in the three dimensions of interest. Here we illustrate the explanatory power of these groupings for average firm size. Again we follow a simple approach by regressing average firm size (over the sample period) on the posterior probabilities defined above. Using turnover and wage data we explain on average 25% of the variance in firm size (respectively 24% among managers, 21% white collars, 32% skilled blue collars, 21% unskilled workers). Hence, wage and turnover policies that are followed by the firms in each occupation (or equivalently the specific labour market conditions the firms experience) are associated with a substantial proportion of the variance of firm sizes. Obviously this association is not causal, since other choices made by firms may dictate the policies that the firms follow, and determine, at least in part, firm sizes, and the labour market(s) in which firms are active.

To address the latter issue further, we measure the association between each groupings posterior probability and some measurement of location of firms (approximate longitude and latitude[32]). We regress[33] for each firm and each group the posterior probability of belonging to any given group on the spatial characteristics of the firms. This can be understood as a test for homogeneity of the types in geographical space. Indeed, if firm type is independent of location, spatial coordinates should have no explanatory power. However, this is not case. In all occupation groups and for all firm types, geographical location is (weakly) associated with firm type.

---

[32] Our measure of firm location is crude and is derived from the location of workers' residences. We only observe the municipality of residence for each worker. We know the longitude and latitude of the centre of each municipality. We locate firms where individuals reside on average using the latter coordinates.

[33] The results of this analysis are not presented here, but are available from the authors on request.

6. Conclusion

We have presented an empirical model of the determination of starting wages within four broad occupation groups. In a flexible way, our model accounts for the heterogeneous labour market conditions which firms experience or engineer (the hiring and separation policy). In particular, we model duration between hires at the firm level and employment durations. Furthermore, we allow correlations between the individual unobservables that may determine duration between hires, starting wages and employment durations independently of firm specific conditions.

Unobserved firm heterogeneity is an important component of the joint distribution of duration between hires, employment duration and starting wages. Our analysis reveals that it explains a significant share of the observed variance in starting wages. Indeed a small number of firm types accounts for a large proportion of the variance that can be attributed to firms. Furthermore, we note that the firm effects are not independently distributed, although our analysis reveals no obvious pattern.

For all occupations, we find that both the duration between hires and the employment duration are not well described by exponential durations, in that the hazard estimated using our flexible model is not constant. At the individual level the correlation between these two durations is close to zero. This is a property shared by our estimates of the correlation between the duration between hires and the starting wage and by the correlation between the starting wage and the employment duration. This implies that friction effects are insignificant. In addition we find that the hypothesis of exogenous matching between firms and new workers cannot be rejected. Friction effects in the determination of wages are small, which suggests that the Danish labour market is close to competitive.

We attempt to describe, in a parsimonious way, the dynamics of the starting wage within firm. In particular, we find a significant positive relationship between the starting wage paid to a new worker and the starting wage the firm paid to the last worker previously hired. However, although significant and positive the parameter is small in the sense that any given transaction has limited effects on future transactions. Our earlier conclusion of a competitive labour market in Denmark is clearly consistent with this finding.

Much is left for future research. In particular we want to extend our work on starting wages to the whole wage profile. Moreover, the precise empirical description of labour markets is a necessary task to understand further the pattern shown by the distribution of firm types. In particular even with such rich data we are still unable to describe the population of workers at risk of matching with any given firm.

References

Abowd, J.M., and C. Kang (2002): Simulatneous Determination of Wages Rates and Tenure, **Mimeo Department of Labor Economics Cornell University**.

Barth, E., and J. Zweimuller (1995) Relative Wages Under Decentralised and Corporatist Bargaining Systems, **Scandinavian Journal of Economics**, 3:95, 369-384.

Bontemps, C., J-M Robin and G. Van den Berg (1999), An Empirical Equilibrium Search Model with Search on the Job and Heterogeneous Workers and Firms, **International Economic Review**, 40, 1039-1074.

Böhning, D. (1995), A review of reliable maximum likelihood algorithms for semiparametric mixture models, **Journal of Statistics Planning and Inference**, Vol 47(5), 5-28.

Burdett, K., and D. Mortensen (1998) Wage Differentials, Employer Size and Unemployment, **International Economic Review**, 39, 257-273.

Canals, J. and S. Stern (2001) Empirical Search Models, **Mimeo University of Virginia**.

Gouriéroux, C. and A. Monfort (1995), **Statistics and Econometric Models**, Cambridge: Cambridge University Press.

Han, A., and J.A. Hausman (1990), Flexible Parametric Estimation of Duration and Competing Risk Models, **Journal of Applied Econometrics**, Vol 5, 1-28.

Heckman, J. and B. Singer (1984) A Method of Minimising the Impact of Distributional Assumptions for Duration Data, **Econometrica**, 52, 271-320.

Jamshidian, M., and R.I. Jennrich (2000), Standard Errors for EM Estimation, **Journal of the Royal Statistical Society B**, 62(2), pp257-270.

Koning P, G. Van den Berg, G. Ridder and K. Albæk (2000) The relation between wages and labour market friction: and empirical analysis based on matched worker-establishment data, in Henning Bunzel, Bent Christensen, Peter Jensen and Nick Kiefer (eds.) **Panel Data and Structural Labour Market Models**. Amsterdam.

Lancaster, T. (1991) **The Econometric Analysis of Transition Data**, Cambridge University Press.

Lillard L.A. (1993) Simultaneous Equations for Hazards, **Journal Of Econometrics**, Vol 56, 189-217.

Meng, X.L. and D.B. Rubin (1993) Maximum Likelihood via the ECM Algorithm: A General Framework, **Biometrika**, Vol 80(2), 267-278.

Meyer, B.D. (1990) Unemployment Insurance and Unemployment Spells, **Econometrica**, Vol 58, 757-782.

Mittelhammer, R.C. , G.G. Judge and D. Miller (2000) **Econometric Foundations**, Cambridge University Press, Cambridge.

Mortensen, D and C. Pissarides (1999) New Developments in Models of Search and the Labor Market, in Orley Ashenfelter and David Card (eds.) **Handbook of Labor Economics** Volume 3, 2567-2628.

Neumann, George, Peder J. Pedersen and Niels Westergaard-Nielsen (1991) Long run International Trends in Aggregate Unionisation, in **European Journal of Political Economy**, 7, 249-274.

Oakes, D. (1999), Direct Calculation of the Information Matrix Via the EM Algorithm, **Journal of the Royal Statistical Society B**, 61(2), 479-482.

OECD (1997) **Employment Outlook**. OECD Publications: Paris.

Postel-Vinay, F. and J-M. Robin (2002), Equilibrium Wage Dispersion with Worker and Employer Heterogeneity, **Econometrica**, Vol 70(6), 2295-2350.

Press, W.H., B.P. Flannery, S.A. Teukolsky & W.T. Vetterling (1993), **Numerical Recipes**, Cambridge University Press.

Van den Berg, G. (1999) Empirical Inference with Equilibrium Search Models of the Labour Market, **Economic Journal**, 109, 283-306.

Van den Berg, G. and G. Ridder (1998) An Empirical Equilibrium Search Model of the Labor Market, **Econometrica**, 66, 1183-1221.

Van Ours, J. and G. Ridder (1992) Vacancies and the recruitment of new employees, **Journal of Labor Economics**, 10:2, 138-155.

Westergaard-Nielsen, N. and P. Jensen (1990) Temporary Layoffs in the Danish Labour Market, in J. Hartog, G. Ridder and J. Theeuves (eds.) **Panel Data and Labour Market Studies** , North Holland.

Table 1. Analysis of Covariance for the Raw Data

Statistic	Outcome	Managers	White collar	Skilled	Unskilled
R-squared	Starting wage	0.309	0.272	0.294	0.329
	Last hire	0.503	0.510	0.485	0.490
	Job length	0.247	0.303	0.247	0.267
Correlation	Wage, last hire	-0.135	-0.055	-0.159	-0.030
	Wage, job length	0.179	0.028	-0.031	0.171
	Last hire, job length	-0.129	-0.115	0.028	-0.038
Count	Firms	294	524	421	456
Count	Spells	3308	5246	4433	6063

Table 2. Maximum Likelihood Estimates		Managers (9 types)		Non Managerial White Collars (7 types)		Skilled Blue Collars (7 types)		Unskilled Blue Collars (8 types)	
		Estimates	Standard Err	Estimates	Standard Err	Estimates	Standard Err	Estimates	Standard Err
	Ln Starting wage previous transaction	-0.071	0.057	-0.142	0.050	-0.149	0.055	-0.049	0.068
	Ln Duration between hires previous transaction	0.178	0.033	0.082	0.027	0.175	0.027	0.100	0.023
β	Ln Mean Tenure in $\mathcal{T}$	0.146	0.050	0.077	0.046	-0.032	0.074	0.059	0.049
	Mean Age in $\mathcal{T}$	-0.001	0.006	0.003	0.005	-0.023	0.006	-0.005	0.004
	Mean Education in $\mathcal{T}$	-0.116	0.023	-0.078	0.031	-0.044	0.036	-0.078	0.032
	% Males in $\mathcal{T}$	0.307	0.151	-0.044	0.189	0.151	0.226	-0.177	0.116
	Ln Starting wage in previous transaction	0.020	0.052	-0.065	0.054	0.091	0.065	0.170	0.063
	Ln Duration between hires previous transaction	0.028	0.023	-0.010	0.022	0.065	0.019	0.069	0.020
λ	Ln Mean Tenure in $\mathcal{T}$	-0.120	0.047	0.025	0.050	0.099	0.055	0.156	0.046
	Mean Age in $\mathcal{T}$	-0.010	0.006	-0.008	0.004	-0.007	0.005	-0.010	0.005
	Mean Education in $\mathcal{T}$	0.010	0.024	0.000	0.028	-0.019	0.037	-0.020	0.030
	% Males in $\mathcal{T}$	0.256	0.166	0.098	0.107	0.062	0.158	0.201	0.147
	Individual Age in $\mathcal{T}$	0.143	0.013	0.141	0.012	0.069	0.012	0.137	0.009
	Individual Age squared in $\mathcal{T}$	-0.001	0.000	-0.001	0.000	-0.001	0.000	-0.001	0.000
η	Individual Education in $\mathcal{T}$	0.111	0.061	0.372	0.087	0.183	0.124	0.322	0.078
	Individual Education squared in $\mathcal{T}$	-0.003	0.002	-0.013	0.004	-0.008	0.006	-0.013	0.004
	Individual Male dummy	-0.021	0.063	-0.028	0.040	-0.042	0.081	-0.002	0.050
	Ln Starting wage previous transaction	0.094	0.044	0.089	0.044	0.080	0.063	0.115	0.131
	Ln Duration between hires previous transaction	-0.013	0.008	0.004	0.006	-0.005	0.006	0.000	0.006
	Ln Mean Tenure in $\mathcal{T}$	-0.010	0.013	-0.027	0.012	0.006	0.014	0.024	0.011
	Mean Age in $\mathcal{T}$	-0.004	0.002	-0.002	0.001	0.001	0.001	0.001	0.001
	Mean Education in $\mathcal{T}$	-0.015	0.008	0.015	0.009	-0.004	0.011	0.003	0.007
γ	% Males in $\mathcal{T}$	-0.120	0.049	-0.028	0.038	-0.023	0.046	0.092	0.045
	Individual Age in $\mathcal{T}$	0.050	0.005	0.045	0.004	0.055	0.004	0.028	0.003
	Individual Age squared in $\mathcal{T}$	0.000	0.000	0.000	0.000	-0.001	0.000	0.000	0.000
	Individual Education in $\mathcal{T}$	0.030	0.024	-0.050	0.023	-0.029	0.035	0.023	0.019
	Individual Education squared in $\mathcal{T}$	0.000	0.001	0.003	0.001	0.003	0.002	-0.001	0.001
	Individual Male dummy	0.201	0.023	0.139	0.013	0.143	0.029	0.078	0.016
	σ_w	0.558	0.025	0.446	0.018	0.486	0.031	0.383	0.031
	ρ_{HS}	0.012	0.019	0.028	0.018	0.075	0.018	0.052	0.016
	ρ_{Hw}	0.056	0.019	-0.007	0.016	-0.036	0.017	0.000	0.017
	ρ_{Sw}	0.010	0.024	-0.089	0.025	-0.095	0.025	0.079	0.021
Order of Chebyshev Polynomial : Duration Between Hires		8		10		9		9	
Order of Chebyshev Polynomial : Employment Duration		8		9		10		10	

Table 3. The Explanatory Power of Firm Types

Outcome	Occupation	ANOVA	P_i	$P_i + X_i$
Start wage	Managers	0.309	0.169	0.204
	White collars	0.272	0.092	0.124
	Skilled	0.294	0.071	0.157
	Unskilled	0.329	0.062	0.165
Job length	Managers	0.196	0.083	0.116
	White collars	0.265	0.093	0.138
	Skilled	0.217	0.105	0.133
	Unskilled	0.210	0.064	0.093
Last hire	Managers	0.544	0.319	0.348
	White collars	0.522	0.307	0.326
	Skilled	0.497	0.267	0.289
	Unskilled	0.570	0.281	0.324

Note: Estimated number of firm types is 9, 7, 7 and 8 for managers, white collars, skilled and unskilled respectively.

Table 4. Exogeneity Tests.

	Managers	White collars	Skilled	Unskilled
Test- Stat	0.829	0.458	0.938	0.653
d-o-f	275	265	265	270
Nobs	3300	5225	4424	6062

Figure 1. Event Histories for a Hypothetical Firm

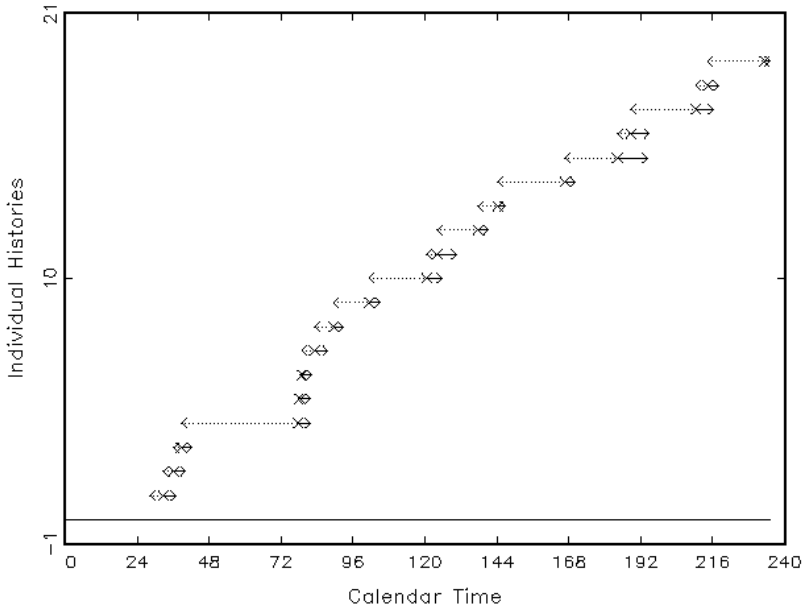


Figure 2.

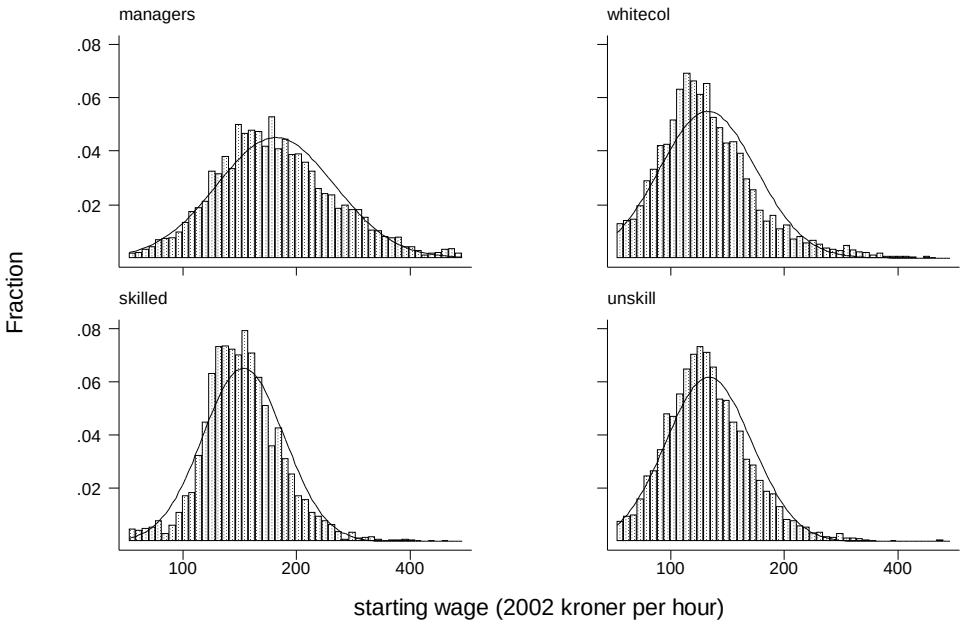


Figure 3.

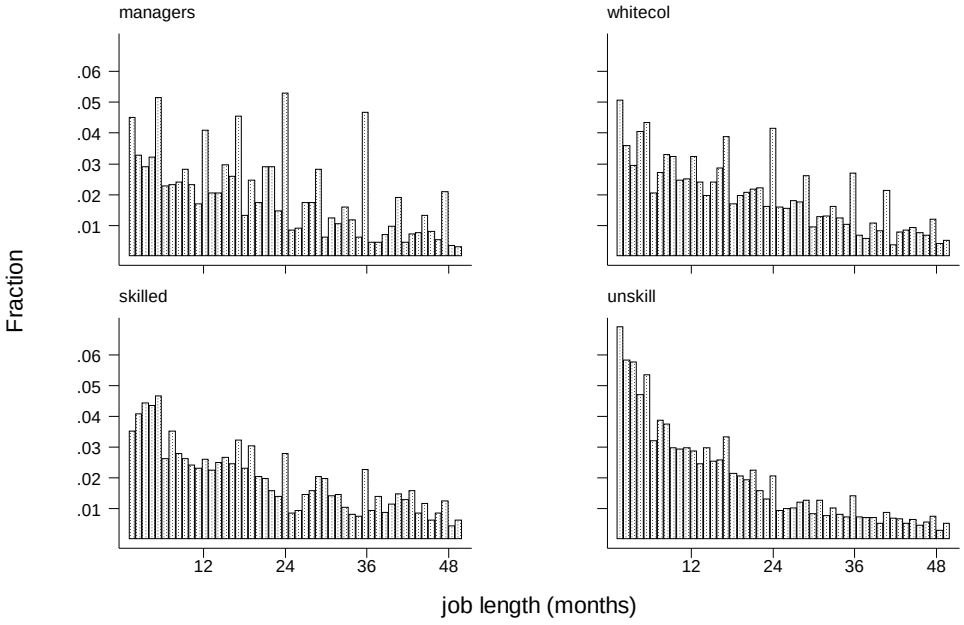


Figure 4.

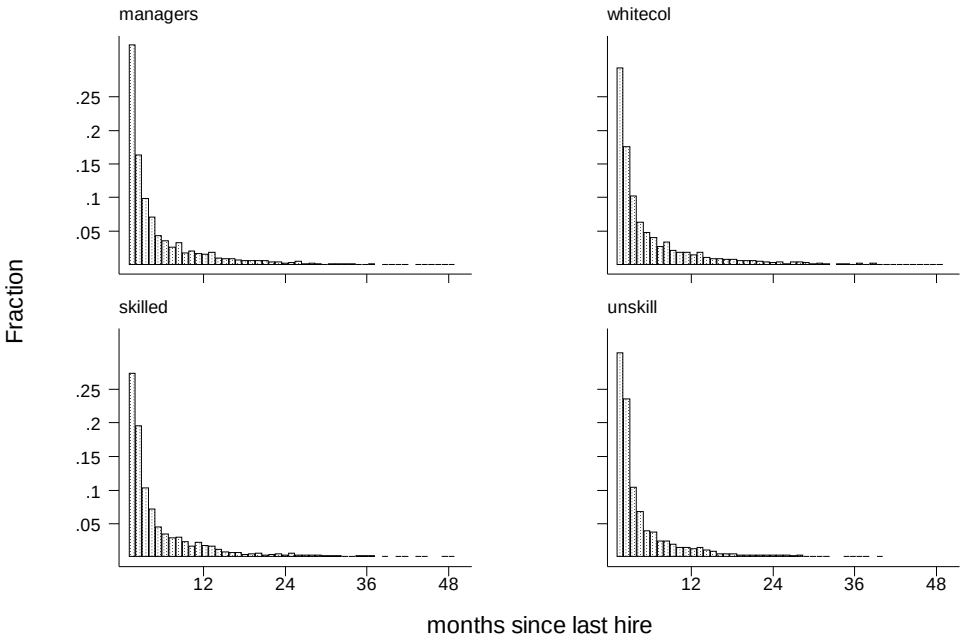


Figure 5.

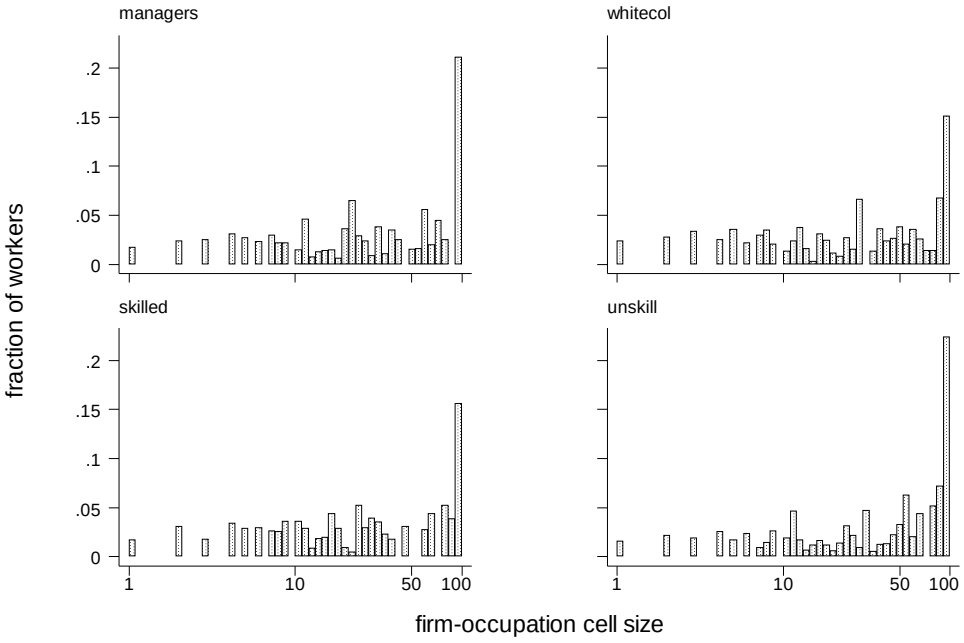


Figure 6a. Durations, Managers.

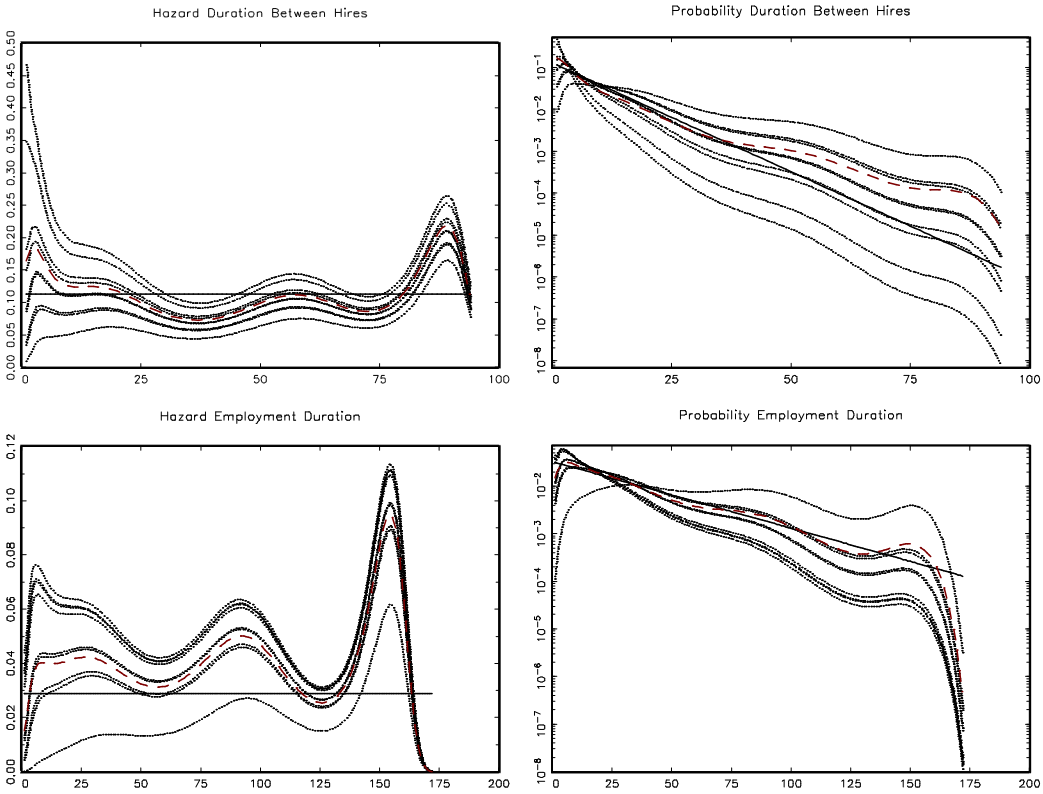


Figure 6b. Policies, Managers.

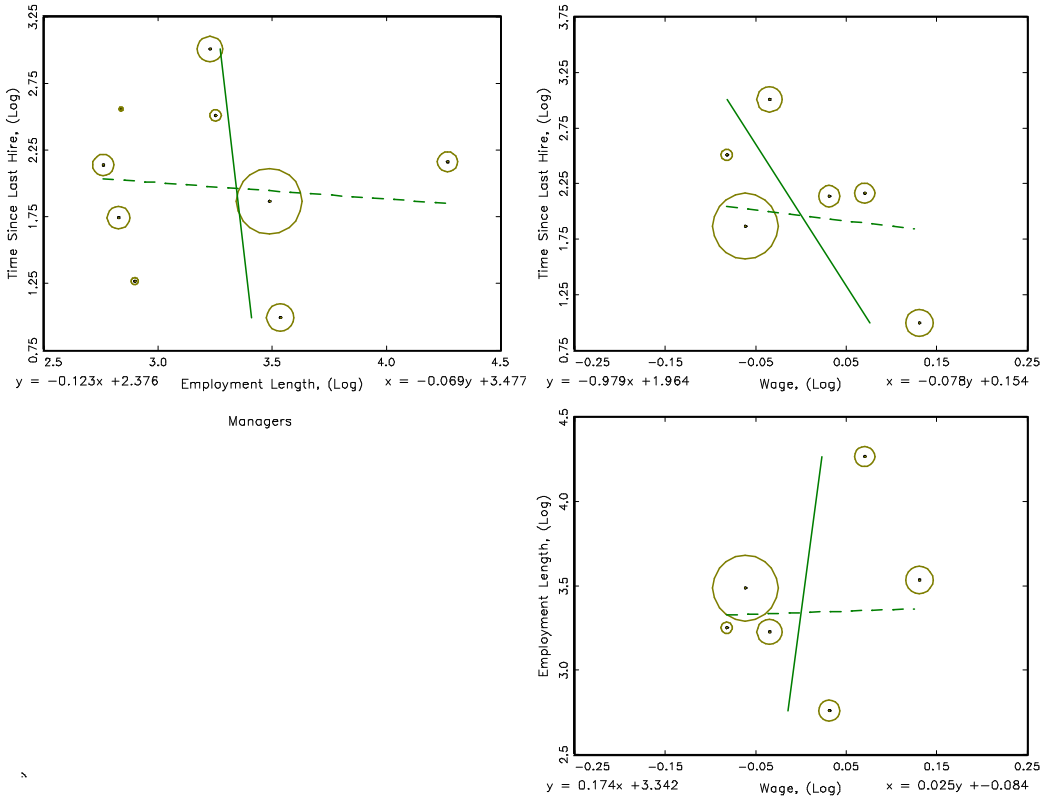


Figure 7a. Durations, Non Managerial White Collars.

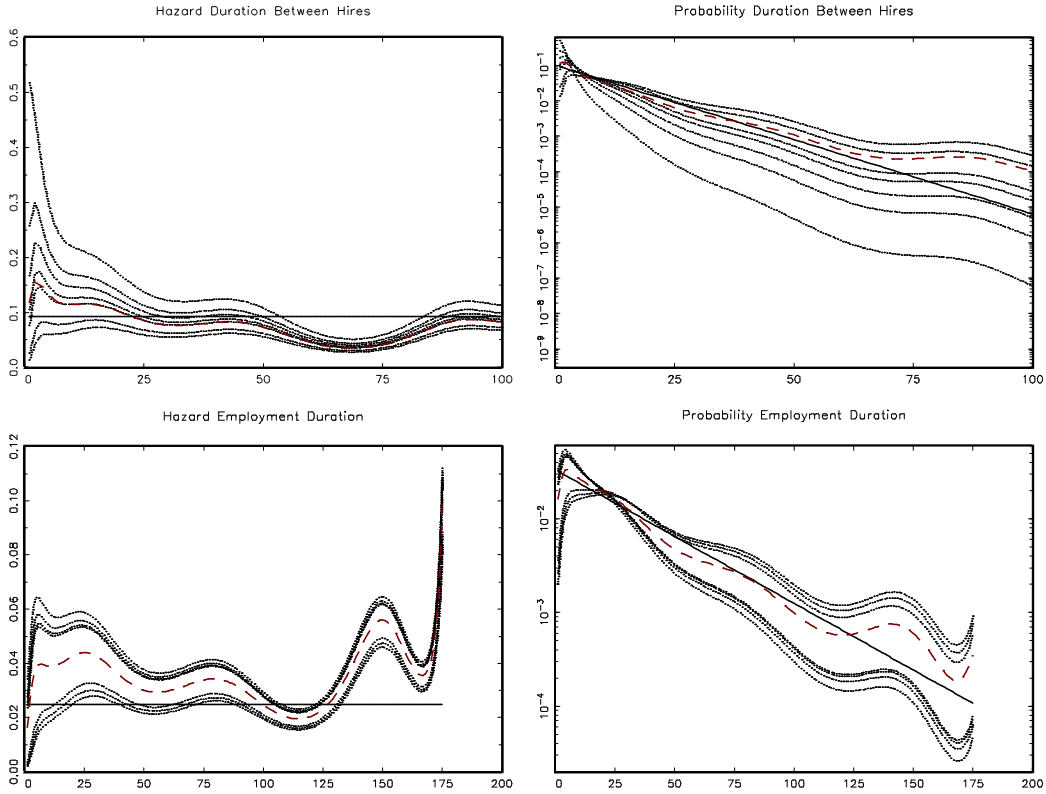


Figure 7b. Policies, Non Managerial White Collars.

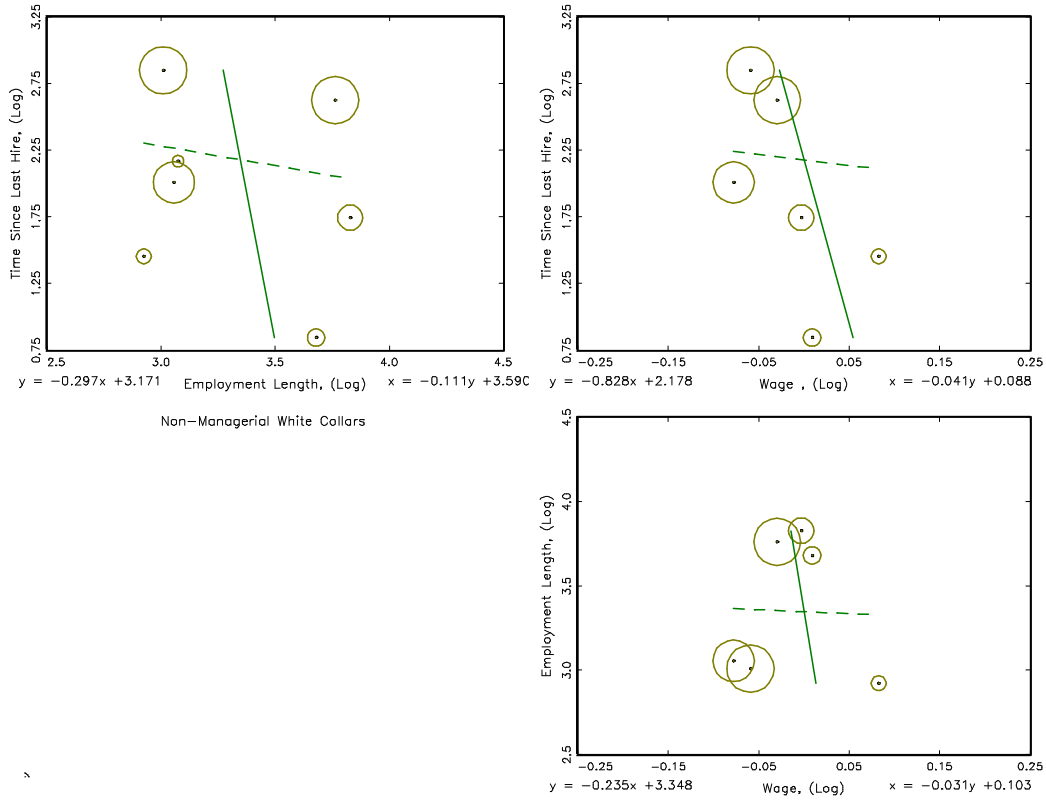


Figure 8a. Durations, Skilled Blue Collars.

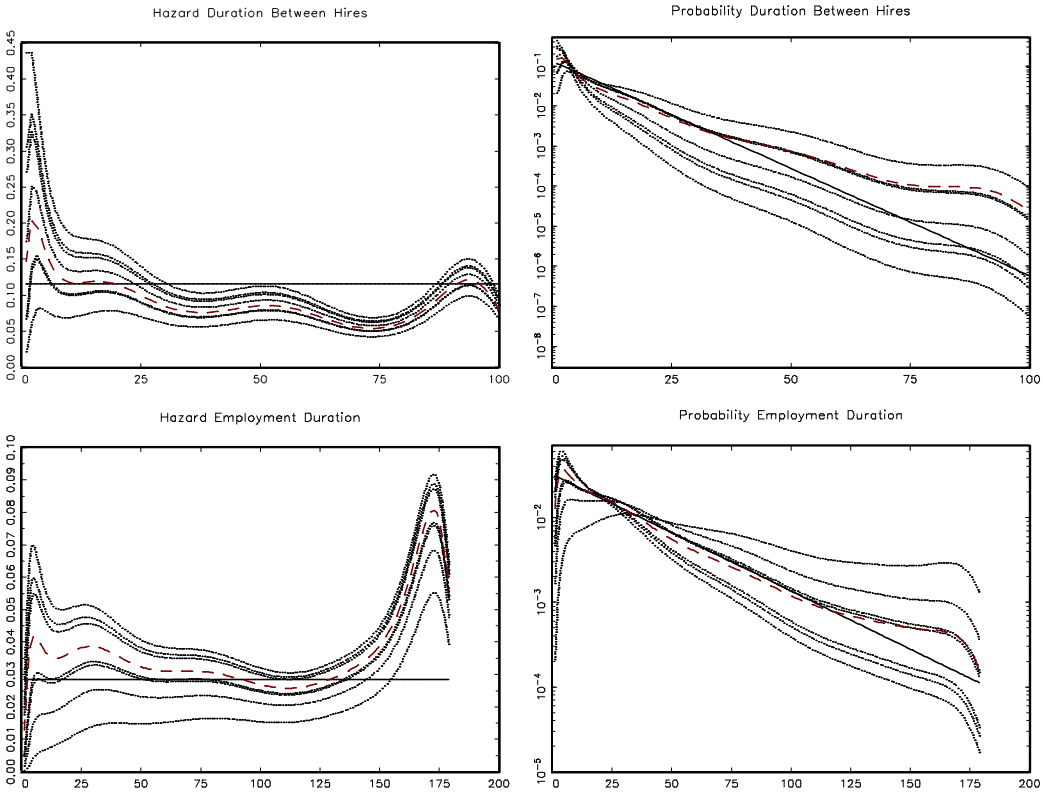


Figure 8b. Policies, Skilled Blue Collars.

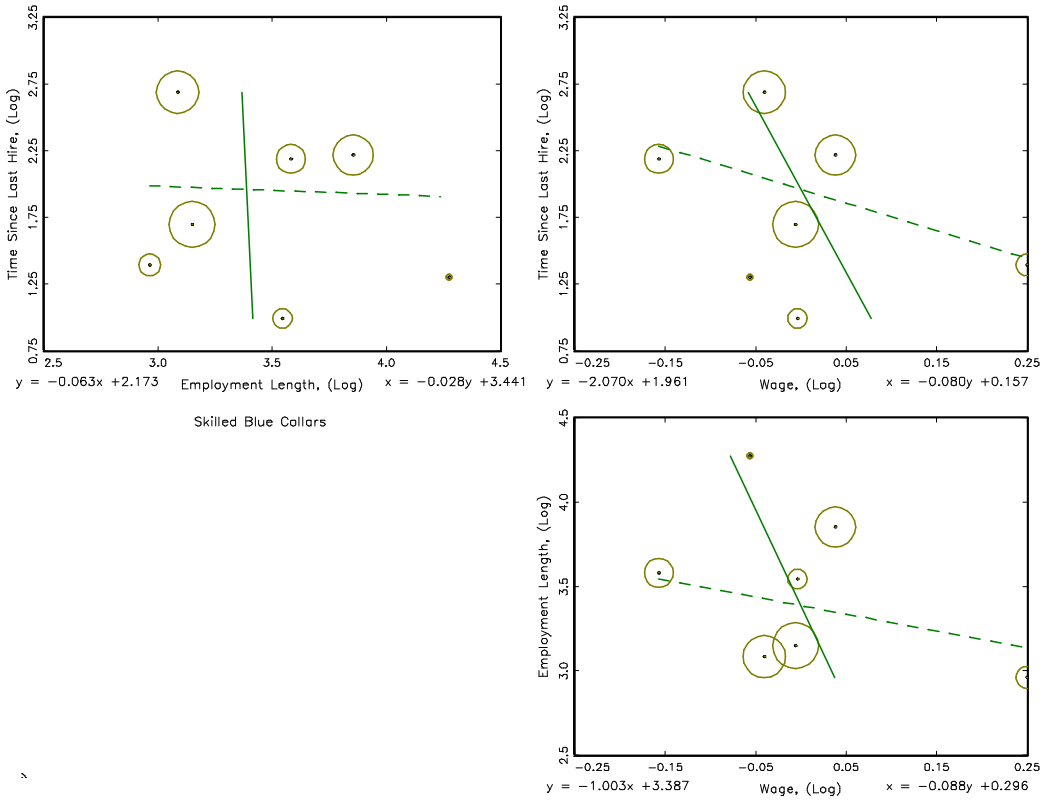


Figure 9a. Durations, Unskilled Blue Collars.

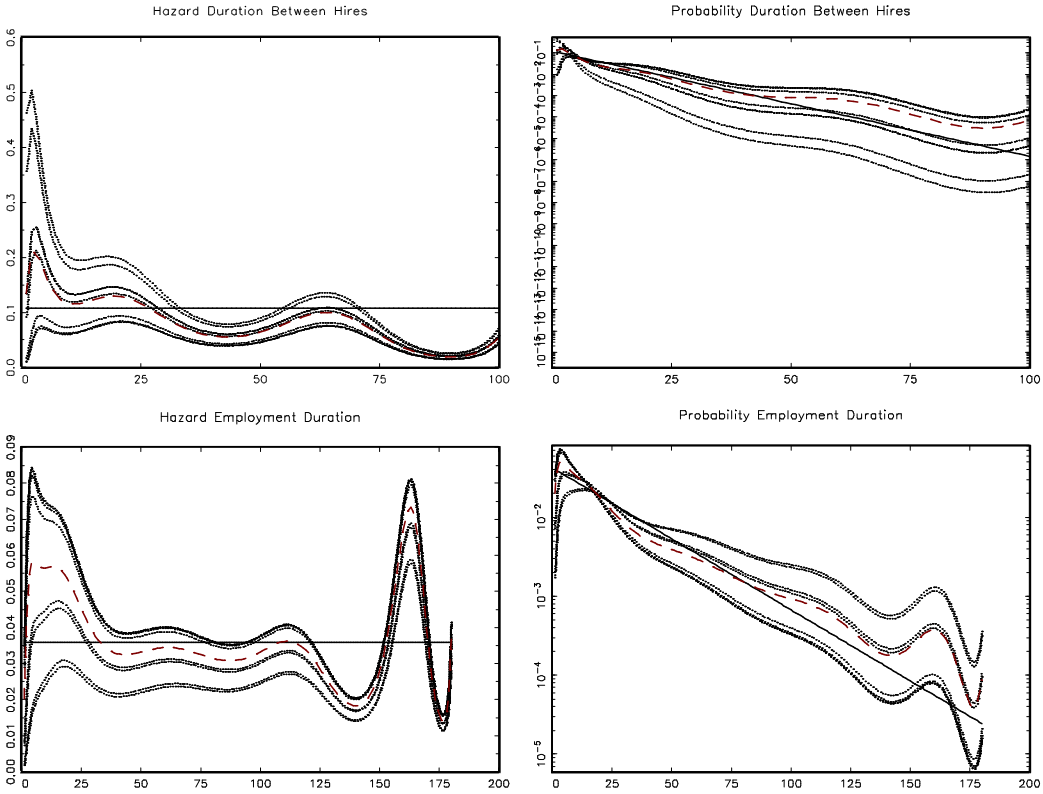
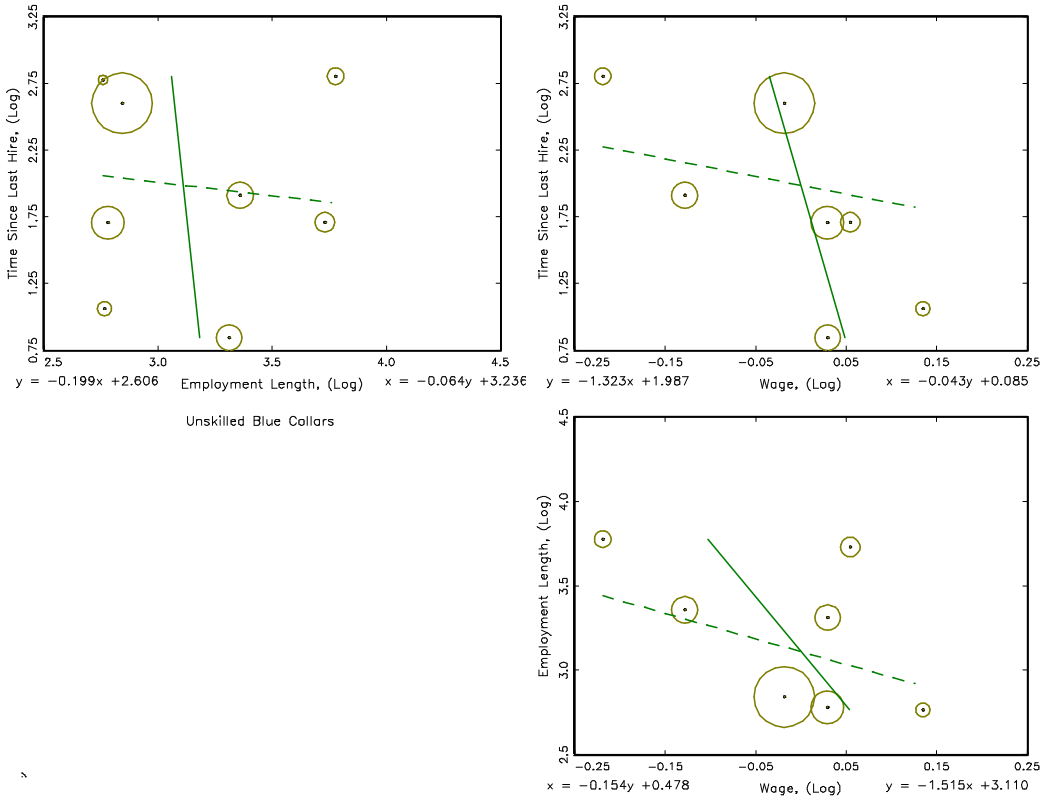


Figure 9b. Policies, Unskilled Blue Collars.



KERP Keele Economics Research Papers — Recent Contributions

All papers in the KERP series are available for downloading from the Keele Economics website, via *www.keele.ac.uk/depts/ec/kerp*.

2003/01 *Analytical Results for a Model of Periodic Consumption*
Clare Kelly and Gauthier Lanot

2002/22 *Asymmetric Contests with General Technologies*
Richard Cornes and Roger Hartley

2002/21 *Damaged Durable Goods*
Jong-Hee Hahn

2002/20 *Unemployment Insurance under Moral Hazard and Limited Commitment: Public vs Private Provision* Jonathan P Thomas and Tim Worrall

2002/19 *The Water Poverty Index: an International Comparison*
Peter Lawrence , Jeremy Meigh and Caroline Sullivan

2002/18 *Price-Capping regulation as a protectionist strategy in developing countries*
Peter Lawrence and Arijit Mukherjee

2002/17 *Licensing and the Incentive for Innovation*
Arijit Mukherjee and Soma Mukherjee

2002/16 *Joint Production Games with Mixed Sharing Rules*
Richard Cornes and Roger Hartley

2002/15 *Sequential Models of Bertrand Competition for Deposits and Loans under Asymmetric Information* Frédérique Bracoud

2002/14 *Consumption Patterns over Pay Periods*
Clare Kelly and Gauthier Lanot

2002/13 *Licensing under Asymmetric Information*
Arijit Mukherjee

2002/12 *Emission Targets and Equilibrium Choice of Technique*
Martin E Diedrich

2002/11 *Dissipation In Rent-Seeking Contests With Entry Costs*
Richard Cornes and Roger Hartley

2002/10 *Advantageous or Disadvantageous Semi-collusion*
Arijit Mukherjee

2002/09 *Licensing in a Vertically Separated Industry*
Arijit Mukherjee

2002/08 *U-shaped Paths of Consumption and Phys. Capital in Lucas-type Growth Models*
Farhad Nili

2002/07 *On the Variance Covariance Matrix of the M.L. Estimator of a Discrete Mixture*
Gauthier Lanot

2002/06 *Monotonicity and the Roy Model*
Arnaud Chevalier and Gauthier Lanot

2002/05 *Capacity Commitment and Licensing*
Arijit Mukherjee

2002/04 *Household Credit and Saving: Does Policy Matter?*
Peter Lawrence

KERP Keele Economics Research Papers — Recent Contributions

All papers in the KERP series are available for downloading from the Keele Economics website, via *www.keele.ac.uk/depts/ec/kerp*.

2002/03 *Innovation, Licensing and Welfare*
Arijit Mukherjee

2002/02 *Historical Schools of Economics: German and English*
Keith Tribe

2002/01 *R&D, Licensing and Patent Protection*
Arijit Mukherjee

2001/09 *Export and Direct Investment as a Signal in Global Markets*
Arijit Mukherjee and Udo Broll

2001/08 *The Welfare Effects of Quality Degradation with Network Externalities*
Jong-Hee Hahn

2001/07 *Cost Padding in Regulated Monopolies*
Spiros Bougheas and Tim Worrall

2001/06 *Is the Unskilled Worker Problem in Developing Countries Going Away?*
Ed Anderson

2001/05 *Does Society Prefer Small Innovation?*
Arijit Mukherjee

2001/04 *Bilateral Merger in a Leadership Structure*
Tarun Kabiraj and Arijit Mukherjee

2001/03 *Imitation, Patent Protection and Welfare*
Arijit Mukherjee and Enrico Pennings

2001/02 *R&D Organization and Technology Transfer*
Arijit Mukherjee and Sugata Marjit

2001/01 *International Joint Venture and the Technology of the Future*
Sugata Marjit, Arijit Mukherjee and Tarun Kabiraj

2000/20 *Gift-Giving, Quasi-Credit and Reciprocity*
Tim Worrall and Jonathan P Thomas

2000/19 *Land Rents and Competitive Equilibrium*
Martin E. Diedrich

2000/18 *Monopoly Quality Differentiation with Top-quality Dependent Fixed Costs*
Jong-Hee Hahn

2000/17 *Time Consistency and Intergenerational Risk Sharing*
Tim Worrall

2000/16 *The Maximum Interest Rate on an Unbalanced Growth Path*
Martin E. Diedrich

2000/15 *Nonlinear Pricing of a Telecomm. Service with Call and Network Externalities*
Jong-Hee Hahn

2000/14 *Rent- seeking by Players with Constant Absolute Risk Aversion*
Richard Cornes and Roger Hartley

2000/13 *Differential Interest Rates on Unbalanced Growth Paths*
Martin E. Diedrich

ISSN 1740-231X

Department of Economics
Keele University
Keele, Staffordshire ST5 5BG
United Kingdom

tel: (44) 1782 583091
fax: (44) 1782 717577
email: economics@keele.ac.uk
web: www.keele.ac.uk/depts/ec/web/