I Z A

DISCUSSION PAPER SERIES

IZA DP No. 3992

Job Durations with Worker and Firm Specific Effects: MCMC Estimation with Longitudinal Employer-Employee Data

Guillaume Horny
Rute Mendes
Gerard J. van den Berg

February 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Job Durations with Worker and Firm Specific Effects: MCMC Estimation with Longitudinal Employer-Employee Data

Guillaume Horny
Bank of France and UC Louvain

Rute Mendes
Tinbergen Institute

Gerard J. van den Berg
VU University Amsterdam, IFAU-Uppsala,
Tinbergen Institute, CEPR, INSEE-CREST and IZA

Discussion Paper No. 3992
February 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3992
February 2009

ABSTRACT

Job Durations with Worker and Firm Specific Effects: MCMC Estimation with Longitudinal Employer-Employee Data*

We study job durations using a multivariate hazard model allowing for worker-specific and firm-specific unobserved determinants. The latter are captured by unobserved heterogeneity terms or random effects, one at the firm level and another at the worker level. This enables us to decompose the variation in job durations into the relative contribution of the worker and the firm. We also allow the unobserved terms to be correlated. For the empirical analysis we use a Portuguese longitudinal matched employer-employee data set. The model is estimated with a Bayesian Markov Chain Monte Carlo (MCMC) estimation method. The results imply that firm characteristics explain around 30% of the variation in log job durations. In addition, we find a positive correlation between unobserved worker and firm characteristics.

JEL Classification: C11, C15, C41, J20, J41, J62

Keywords: job transitions, assortative matching, Gibbs sampling, frailties, dynamic models, matched employer-employee data

Corresponding author:

Gerard J. van den Berg
Department of Economics
VU University Amsterdam
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
E-mail: gjvdberg@xs4all.nl

* We thank François Laisney, Jean-Pierre Florens, Denis Fougère, Bertrand Koebel, and participants in AFSE, SED, ESEM and IZA conferences, for detailed comments. We are grateful to the Portuguese Ministry of Employment (Statistics Department) for the access to the data. Mendes acknowledges financial support by the Portuguese Foundation for Science and Technology. Part of the work was carried out while Horny was at Louis Pasteur University in Strasbourg and while Mendes was at VU University Amsterdam and Collegio Carlo Alberto, Turin.

1 Introduction

The basic stylized facts regarding job durations are well established. For example, the survey by Farber (1999) provides abundant evidence that in OECD countries, long-term employment relationships are common, most new jobs end early, and the probability of a job ending declines with tenure. Unobserved heterogeneity in the probabilities of job exit can largely account for these stylized facts. If workers are heterogeneous in terms of mobility propensities, then the observed job exit rate at any point in time depends on the proportions of those types. Higher mobility workers experience several short spells while lower mobility workers engage in fewer but longer employment relationships. The fact that most new jobs end early is explained by a sufficiently large proportion of high mobility workers. Furthermore, the fact that the probability of job ending is observed to decline with tenure is explained by sorting of the workers into different tenure groups: longer (shorter) tenure groups include a larger proportion of lower (higher) mobility workers.

Since job exit is a decision that involves both the worker and the firm, it is plausible that exit rates are affected simultaneously by characteristics of workers and by characteristics of firms. Whereas the relevance of worker heterogeneity in job durations is well established (see e.g. Farber, 1999, Bellmann *et al.* 2000, and Del Boca and Sauer, 2006), the empirical evidence on the importance of firm heterogeneity is much more limited.[1] It is relevant for a number of reasons to know the relative contributions of worker and firm characteristics as determinants of job durations. First, notice that inequality in society depends on the variation in the characteristics of the jobs that employed individuals have. If the variation in job durations is primarily driven by worker characteristics then the ensuing inequality will be more persistent. Conversely, if the variation is primarily driven by firm characteristics then the restructuring of the market form in a sector can have large effects on inequality in society. Secondly, the results of the analysis are of importance for the econometric analysis of job durations. If unobserved firm heterogeneity is important then the inclusion of very large numbers of worker characteristics to a job duration model does not remove the need to deal with unobserved heterogeneity. Thirdly, the results simply help in improving our understanding of the determinants of job durations and job mobility. (Below we also discuss the relevance for the study of assortative job matching.)

In this paper, we estimate multivariate hazard models for job exits (or, equivalently, job durations), allowing for worker-specific and firm-specific unobserved determinants. We also allow these unobserved terms to be correlated. Worker-specific determinants encompass the propensity of the worker to leave or lose a job, while

[1] Abowd, Kramarz and Roux (2006) include worker and firm unobserved heterogeneity in a model for wages and job mobility and conclude that there is a large amount of heterogeneity among firms and their tenure profiles.

firm-specific determinants can reflect the firm's preference to employ a stable workforce. Furthermore, considering that the matching process between firms and workers may follow some assortative pattern, also in terms of characteristics that are unobservable to the researcher, we allow the unobserved effects of matched firms and workers to be correlated. To our knowledge, this is the first study that allows for such a flexible and precise modelling in job mobility decisions. Obviously, the econometric analysis requires observation of multiple job spells per worker and/or multiple job spells per firm. A firm is cross-sectionally and longitudinally connected to multiple workers, whereas a worker is longitudinally connected to multiple firms. We use a matched employer-employee data set in which both workers and firms are longitudinally followed. The data are from Portugal and are exhaustive for the private sector.

In the econometric analysis we treat the unobserved heterogeneity terms as random effects, one at the firm level and another at the worker level. This is in line with econometric duration analysis with unobserved heterogeneity (see Van den Berg, 2001). Due to right-censoring, fixed-effect panel data methods are not feasible. More to the point, we are interested in the relative contributions of workers' and firms' characteristics in the variation of job durations, and the estimation of the distribution of the random effects enables such a decomposition.

The model structure is such that the unobserved worker and firm effects are neither nested nor independent. In fact, the dependence between the worker and firm effect in a given job creates a major complication for the analysis. If the correlation between the worker and firm random effects is sufficiently high then this entails that the random effects of different workers at a given firm are correlated, and also that the random effects of different firms employing a given worker over time are correlated. As we will show, this is an implication of the required positive semi-definiteness of the correlation matrix of e.g. the random term of a firm and the random terms of two of its workers. A dependence across workers at a given firm and across firms having employed a given worker implies that many observed job spells of different workers and firms are statistically dependent. Indeed, with a sufficiently high mobility and a sufficiently large observation window, all firms and all workers would have jointly dependent random effects. This would make it difficult to apply standard Likelihood-type estimation methods that are used to estimate duration models with one-dimensional random effects. First of all, the computational burden would be insurmountable. Secondly, it is not clear what would be an appropriate asymptotic distribution to obtain reliable standard error estimates. Because of the complex pattern relating the two random effects, we estimate the model using a Bayesian Markov Chain Monte Carlo approach, based on the Gibbs Sampler and in line with Manda and Meyer (2005).[2] In addition, we consider restrictions on the

[2]Robert and Casella, 1999 provide a survey of Markov Chain Monte Carlo methods (MCMC).

correlation of the worker and firm random effects, such that the assumptions of worker random effects being i.i.d. across workers and firm random effects being i.i.d. across firms are not violated in the observation window. Our paper thus contributes to the methodological literature by showing how to handle this complex unobserved heterogeneity structure.[3]

The estimates of the correlation between the worker-specific and the firm-specific unobserved heterogeneity term are informative on the extent to which specific types of firms match with specific types of workers. To see this, consider a firm where the job durations are typically short. Is this only because of high job exit rates at the firm, or is it also because the firm attracts workers who have high job exit rates anyway, i.e. who would also have high job exit rates if employed at firms where job spells are typically long? The former reflects firm heterogeneity whereas the latter leads to a positive correlation estimate. Our paper is therefore connected to the expanding literature on assortative matching of workers and firms. Recent advances in this literature focus on assortative matching in terms of (worker-specific and firm-specific) productivity (see Mendes, Van den Berg and Lindeboom, 2007, and Lopes de Melo, 2008, for empirical analyses based on matched employer-employee data). If the worker and the firm each have a high productivity contribution then the surplus of the match may also be high, and if this is divided amongst them then they may be relatively satisfied with the match, resulting in a low job exit rate. So to the extent that productivity is reflected in job exit rates, we may use the estimate of the correlation to examine whether the job duration data confirm positive assortative matching.

The paper is organized in six sections. The Portuguese matched employer-employee data are described in Section 2. These data have been used before in a number of studies. See e.g. Vieira, Cardoso and Portela (2005), Cardoso and Portela (2005), and Mendes, Van den Berg and Lindeboom (2007), for descriptions and analyses of the data and for summaries of the Portuguese labor market. Section 3 presents the duration models that we estimate. In Section 4, we discuss the estimation method and the choice of the prior distributions. The results are discussed in Section 5. Section 6 concludes.

2 Data

The study is based on *Quadros de Pessoal*, a longitudinal matched employer-employee data set gathered by the Portuguese Ministry of Labor and Solidarity. The data are collected through a report that all firms with registered employees are legally obliged to provide every year. The reported data concern all workers employed by the firm in

[3]See Dostie, 2005 for an analysis with a worker random effect and a job match-specific random effect.

the month in which the survey is collected (March up to 1993, October since 1994). Coverage is low for the agricultural sector and non-existent for public administration and domestic services. On the other hand, the manufacturing and private services sectors are almost fully covered.

An identification code is assigned to every firm when it enters the data set for the first time, while the identification code of the worker is a transformation of his social security number. Based on these identification numbers, one can match workers and firms, and follow both over time to identify job-to-job transitions. To avoid initial-condition problems, we reconstruct the data as if they were collected using flow sampling, by keeping only spells with observed entry. We return to this issue in Section 4.

Since additional checks on the accuracy of the firm identification code are implemented by the Ministry since 1994, we use the data covering the period 1994-2000. These data comprise nearly 385 thousand firms and 4 million workers. For our study we apply a few conditions on characteristics of workers, firms and spells: (a) we discard firms that leave, temporarily or permanently, the market in order to exclude from our analysis job transitions exclusively caused by the closure of the firm; (b) we exclude workers who, at some point, are observed in a non-paid job or in self-employment; (c) we exclude spells with no observed entry and spells terminated by a transition which is not job-to-job. This results in a dataset covering around 338 thousand workers and 55 thousand firms. Descriptives relative to these data are presented under "Full data" in the Appendix . To reduce the computational burden of the estimation procedure, we also use a subsample that we extract such that it has the same worker characteristics as the full data. This involves 6582 firms and 9222 workers. The data confirm the stylized facts that new jobs end early, and that the transition rate decreases with tenure.

Due to the way in which the data are collected, we do not have details on the worker's labor market events between consecutive surveys, nor do we know precisely when in between the survey months a job exit took place. We do identify transitions of workers between firms occurring in time intervals of one year, and we do observe the occurrence of other short spells (job, unemployment or non-participation spells) within that time interval. Table 1 summarizes the number of spells per worker.

Table 1: Number of spells per worker

Number of spells	Full data	Sample
1	90.6%	90.6%
2	8.5%	8.5%
3	0.9%	0.9%
Total	100%	100%

Most workers experience few transitions: only 1 % experienced 3 job transitions in the period 1994-2000. Indeed, we are not investigating temporary but “permanent” employment, with contracts of at least one year.

In our model of job transitions, we use the following observed characteristics of the worker: age, gender, and education. We also observe whether the job is a part-time job or not. Age may capture life-cycle effects. 'Job shopping' tends to take place mainly at an early age, while the worker is not aware of his own abilities or of the characteristics of the labor market (Johnson, 1978). Age is grouped into the categories: 16 - 25, 26 - 35 and 36 - 55 years old. Workers older than 55 were discarded in order to avoid considering also transitions to retirement, which are out of the scope of this analysis. Different degrees of attachment to the labor market, differences in child care and family responsibilities, among other factors, may result in gender differences in terms of job mobility. We also control for education, which is grouped into three categories: primary school, lower secondary, upper secondary and higher education. A part-time indicator is also included because firms facing negative demand shocks may tend to first terminate part-time jobs in order to minimize the loss of specific human capital. Regarding firms, the observed characteristics included in our analysis are economic sector, location, and an indicator for multiple plants.

The wage is also included in the set of controlled observed characteristics influencing the job mobility process. In search models, the wage is often a firm characteristic and is accordingly included as an exogenous variable for the job exit rate. Alternatively, one may think of the wage as being partly determined by job mobility decisions, and so, because of its endogeneity, it should be kept out of the controls included in the job exit rate specification. Also, inclusion of the wage variable as an explanatory variable would complicate the interpretation of the worker and firm random-effects dependence as an indicator of assortative matching. For these reasons, we estimate the models both with and without the wage in the right hand side. Descriptives of firms' and workers' characteristics are presented in the Appendix A.

3 Model

3.1 Discrete-time job duration models

Since we only observe job entry and job exit on an annual basis, we specify the job duration models in discrete time. Specifically, we use the time-aggregated Mixed Proportional Hazard (MPH) model for the hazard function or conditional job exit probability. This is in line with the fact that the underlying processes and transitions are in continuous time.

In our application, a firm is cross-sectionally and longitudinally connected to multiple workers, but a worker is only longitudinally connected to multiple firms. There is thus no hierarchy in the sample: although a firm consists of multiple workers, these workers change between firms when they move to another job. We denote by $i = 1, ..., I$ the firm index and by $j = 1, ..., J$ the worker index. Let the time scale be divided into intervals $]a_{k-1}, a_k]$ where $0 = a_0 < a_1 < \ldots < a_K < \infty$. The discrete-time job duration t_{ijk} is in $\{1, \ldots, K\}$ and indicates a transition observed in $]a_{k-1}, a_k]$. Here, the hazard function is a conditional probability and can be written as:

$$\lambda\left[t_{ijk}|x_{ij}(t_{ij(k-1)}), v_i, w_j\right] = p[a_{k-1} < T \leq a_k | T \geq a_{k-1}, x_{ij}(t_{ij(k-1)}), v_i, w_j], \quad (1)$$

where $x_{ijk}(t_{ij(k-1)})$ are both worker- and firm-specific observed explanatory variables that are potentially time varying, v_i is a random effect at the firm level (more precisely: the effect of unobserved characteristics of firm i on the job exit rate of jobs at firm i; we also call this the unobserved heterogeneity term or frailty) and w_j a random effect at the worker level. Note that the firm random effect is invariant across job spells at the firm, whereas the worker random effect is invariant across different jobs occupied by the worker. Both are time-invariant. The "random effects" assumption states that these are independent of the observed explanatory variables.

We estimate a range of model specifications. The simplest model accounts for observed heterogeneity only. Next, we introduce a worker random effect. The third specification allows for worker and firm random effects that are independent of each other. The most general specification allows the two random effects to be correlated for a matched pair of a worker and a firm.

Following the complementary log-log link function described in Kalbfleisch and Prentice (1980), the discrete-time MPH model without unobserved heterogeneity is defined as

$$\lambda\left[t_{ijk}|x_{ij}(t_{ij(k-1)}), \beta_0, \beta_1\right] = 1 - \exp\Big(- \exp[\beta_{0(k-1)} + x_{ij}(t_{ij(k-1)})'\beta_1]\Big), \quad (2)$$

where $\beta_{0(k-1)}$ is the baseline hazard over the time interval $[a_{k-1}, a_k[$. With a worker random effect we obtain

$$\lambda\left[t_{jk}|x_{ij}(t_{ij(k-1)}), \beta_0, \beta_1, w_j\right] = 1 - \exp\Big(- \exp[\beta_{0(k-1)} + x_{ij}(t_{ij(k-1)})'\beta_1 + w_j]\Big). \quad (3)$$

A discrete-time MPH model with two frailties is defined as:

$$\lambda\left[t_{ijk}|x_{ij}(t_{ij(k-1)}), \beta_0, \beta_1, v_i, w_j\right] = 1 - \exp\Big(- \exp[\beta_{0(k-1)} + x_{ij}(t_{ij(k-1)})'\beta_1 + v_i + w_j]\Big). \quad (4)$$

Let us denote by λ_{ijk} the value of the hazard function (1) at time t_{ijk}. The separation of worker j from firm i at time t_{ijk} contributes to the likelihood as:

$$L^d_{ij}(t_{ijk}|\beta_0, \beta_1, v_i, w_j) = \lambda_{ijk} \prod_{s=1}^{k-1} (1 - \lambda_{ijs}) . \tag{5}$$

A censored spell of length t_{ijk} contributes to the likelihood as:

$$L^c_{ij}(t_{ijk}|\beta_0, \beta_1, v_i, w_j) = \prod_{s=1}^{k} (1 - \lambda_{ijs}) . \tag{6}$$

The full likelihood is thus:

$$L(t|\beta_0, \beta_1, v, w) = \prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{k=1}^{K} \lambda_{ijk}^{\delta_{ijk}} (1 - \lambda_{ijk})^{1-\delta_{ijk}}, \tag{7}$$

where δ_{ijk} is a transition indicator. Likelihood (7) is equivalent to the one of a model treating the δ_{ijk} as Bernoulli draws.

3.2 Dependence between the worker and firm random effects

The individual hazard function as specified above is conditional on unobserved worker and firm characteristics. We proceed by specifying the distribution of these unobserved characteristics, or more precisely, the distribution of the effects of these characteristics on the hazard function.

This includes a specification of the dependence between v_i and w_j. We allow these to be dependent within a job, i.e. the determinants v_i and w_j of the job duration of worker j at firm i are allowed to be dependent, with correlation ρ. At the same time, to keep the mutual dependence between and across workers and firms manageable, we assume that v_i is independent across firms and w_j is independent across workers. As noted in Section 1, this creates a complication. To explain this, we examine the possible correlations between the joint set of the random effects. An $n \times n$ matrix is a correlation matrix of n random variables if and only if the following three conditions are satisfied: the matrix is symmetric, its diagonal elements are equal to one, and the matrix is positive semidefinite. The third condition excludes matrices where most off-diagonal elements are close to -1. Laurence, Wang and Barone (2008) give an overview of the results. For example, let $n = 3$ and let the random variables Y_1, Y_2 and Y_3 have zero mean and unit variance. Then the elements $\rho_{ij} \equiv \mathrm{E}(Y_i Y_j)$ of the correlation matrix satisfy

$$\rho_{12} + \rho_{13} + \rho_{23} \geq -\frac{3}{2}$$

Consider one firm with two workers, so that $v_1 = Y_1, w_1 = X_2, w_2 = X_3$. Then $\rho_{23} = 0$ and $\rho_{12} = \rho_{13} = \rho$, and we obtain that $\rho \geq -3/4$. The condition of positive-semidefiniteness also rules out that one random variable is strongly correlated to other random variables that are uncorrelated with each other. For example, it is not difficult to show that in our case of one firm with two workers, we need to impose that

$$\rho \leq \frac{1}{2}\sqrt{2}$$

With three workers, ρ can not exceed $(1/3)\sqrt{3}$. It is clear that as we allow for more and more workers to have worked at a given firm and for more and more firms to have employed given workers, the range of admissible values of ρ covers smaller and smaller intervals around zero.[4] From this point of view, one should either impose $\rho = 0$ from the outset, or drop the assumption that the random effects are independent across workers and independent across firms. In the latter case, current firms' unobserved characteristics are potentially related to previous firms' unobserved characteristics through the workers' labor market histories, and so on. As a result, many observed job spells of different workers and firms can be statistically dependent. Indeed, with a sufficiently high mobility and a sufficiently large observation window, all firms and all workers could have jointly dependent random effects. In that case, the data would provide a single joint observation of a large number of correlated spells with a potentially large number of correlation parameters for the random effects. This would make it difficult to apply standard maximum-likelihood estimation methods that are commonly used to estimate duration models with one-dimensional i.i.d. random effects. The computational burden would be insurmountable, and it is not clear how to obtain reliable standard errors. Even with simulation methods one has to draw from a joint normal distribution with a dimension equal to the sum of the number of workers and the number of firms, with a structured variance matrix of which each element has to satisfy a number of constraints increasing with the matrix dimension to ensure its positiveness.[5]

We deal with this by pursuing a pragmatic strategy that is justified by the small number of multiple job spells per worker in the data. Virtually all workers have less than 3 job spells. We can also reduce the number of workers per firm in the sample (although this may induce a selection bias, since firms with low v_i will tend to have more workers with long job spells and thus may have a larger workforce). With a

[4]With normally distributed random variables, this particularly is easy to see: a firm j with n workers has $\text{var}(v_i \mid w_1, \ldots, w_n) = \text{var}(v_i)\left(1 - n\rho^2\right)$.

[5]We performed some numerical investigations based on a model with three correlations; one common correlation across the worker random effects, one common correlation across the firm random effects, and our parameter ρ. This results in values of the correlations that are all very close to zero. Note that a common correlation for the fixed effects across workers or firms is probably not very realistic.

finite number of spells per worker and per firm, we may estimate models with non-zero ρ without violating the positive-semidefiniteness of the correlation matrix for the firms and workers in the data. If the number of spells per worker and per firm is small then the bounds on the value of ρ may not be violated in the observation window. In other words, the admissible range of ρ may be sufficiently large in order to obtain an informative estimate of ρ, that is, an estimate that is not too close to zero. We also estimate models with $\rho = 0$, and we can then examine whether the fit of the model is significantly better than in the case of $\rho = 0$. A precisely formulated bound on ρ is hard to obtain due to the variation in observed labor market outcomes across firms and workers and the indirect relation between numbers of spells and the random effects. Moreover, such a bound would be sensitive to outliers in the data. If the estimate of ρ is relatively close to $+1$ or -1 then this suggests violation of the assumption that v_i is i.i.d. across firms and w_j is i.i.d. across workers. Notice that the latter has implications for empirical analysis of job durations with worker data and the empirical analysis of job turnover with firm data, as these types of analyses always assume independent outcomes across workers and firms, respectively.

The above approach can be implemented with Bayesian Markov Chain Monte Carlo (MCMC) estimation methods. First, these methods do not require large sample theory to obtain standard errors. Secondly, these methods do not require numerical integration over multidimensional random effects. In the next section we discuss the estimation method in more detail.

In addition to the sample that we drew from the population (see Section 2; hereafter we call this the "unrestricted sample"), we also draw a subsample with no more than three job spells drawn per worker, and no more than three workers drawn for a given firm (hereafter the "restricted sample"). In other words, the restricted sample contains up to 3 spells per worker and 3 employees per firm. It comprises 6577 firms and 7749 workers. The durations and the number of spells in the full data and the samples are presented in Appendix B. Due to the small number of job-to-job transitions observed in the period 1994-2000, the restriction does not modify statistics of the workers characteristics. However, as our firm identifier refers to a company and not to a specific plant, drawing no more than 3 workers per firm reduces the weight of large companies and especially those with multiple plants in the restricted sample.

When choosing a functional form for the unobserved heterogeneity distribution it is important not to restrict the sign of the correlation between the two random effects. Moreover, it is useful to have a family of distributions where the correlation or covariance of the dependence between the two random effects is a separate parameter. This is why we adopt normal distributions for these random effects. We normalize their means to zero, and we denote their variances by σ_f^2 for v_i and by σ_w^2 for w_j.

4 Bayesian inference

The Bayesian approach augments the assumed model with the prior beliefs on the parameters. We choose proper but uninformative priors. Manda and Meyer (2005) specify a baseline hazard with steps, related through a first-order autocorrelated process, and Grilli (2005) uses a polynomial specification. Due to the sampling scheme, in which durations last less than 6 years, we specify a piecewise constant baseline hazard with unrelated coefficients over the small number of time intervals.[6] The coefficients are given independent gaussian priors with mean 0 and variance 1000.

The precision of each random effect (i.e. σ_f^{-2} and σ_w^{-2}) follows a gamma distribution, and we base our prior elicitation on descriptive statistics. The rate of transition per worker is about 3.5% for the 5th quantile of the duration distribution and 0.9% for the 95th quantile. For 90% of the population, there is at most a fourfold variation between the odds of two workers. The corresponding confidence interval on the rate of transition is thus of width 3, which implies $\sigma_w = 0.5$. We set our prior for the precision σ_w^{-2} to a gamma distribution with expectation 2 and variance 4. Similarly, the rates of transition per firm are in a range from 1.3% to 4% for 90% of the population, implying a gamma prior with expectation 3 and variance 9 for σ_f^{-2}. A uniform distribution over $[-1, 1]$ is specified for ρ, which is the least informative possible prior.

Let us denote by T the vector of durations and by M the number of covariates. Using f as a generic symbol for a density, the joint density of the data and parameters can be expressed as:

$$f(T, \beta_0, \beta_1, v, w) = f(\sigma_f^2) f(\sigma_w^2) f(\rho) \left[\prod_{s=1}^{K-1} f(\beta_{0s})\right] \left[\prod_{m=1}^{M} f(\beta_{1m})\right]$$
$$\left[\prod_{i=1}^{I} f(v_i|\sigma_f^2)\right] \left[\prod_{j=1}^{J} f(w_j|\sigma_w^2, \rho, v_1, \ldots v_I)\right] \left[\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K} L_{ijk}(t_{ijk}|\beta_0, \beta_1, v_i, w_j)\right] \quad (8)$$

Each worker and each firm have their own value of w_j and v_i, respectively. As we sample jobs that commence by a worker's inflow into a firm, and we follow workers over time after that, we effectively assume that this inflow is the underlying population. Consequently, the model assumptions like independence of observed and unobserved explanatory variables relate to this population. The same applies to the joint distribution of the random effects. In reality, the inflow into jobs is also determined by the outflow from jobs, so the distribution of characteristics in the inflow depends on the distribution in the outflow. We abstract from these issues.

[6] We also estimated models using polynomial specification and were led to a 6 degrees polynomial, that is, less parsimonious specifications than with a piecewise constant baseline hazard.

Alternatively, we assume that we observe a stationary process. Notice that the model does not allow for calender time dependence anyway.

The posterior is the ratio of (8) over its integral over the parameter space. Even with all priors being independent, it does not admit an analytical solution. However, we can construct a Markov chain with elements following the posterior distribution and approximate the Bayesian estimator using a Monte Carlo method. Here, the quantities of interest are approached using Gibbs sampling (Gelfand and Smith, 1990), an MCMC method involving draws from the distributions of a given parameter conditional on the other relevant parameters.

We run two MCMC chains for each model. On previous runs, we observed the Markov chains for the parameters σ_f^{-2} and σ_w^{-2} to converge more slowly than those for parameters β and ρ. The starting values for β are thus set at the maximum likelihood estimates in a model without unobserved heterogeneity for both chains. For σ_f^{-2} and σ_w^{-2}, they are set to 1 for the first chain and to 50 for the second one. We set the starting value of ρ to 0 for both chains. We run 50 000 iterations for the models with the two frailties. From convergence plots of the sampled values and Gelman and Rubin (1992) statistics, 20 000 iterations were sufficient for the burn in. The posterior statistics are computed from the post-convergence iterations.

5 Results

5.1 Unobserved heterogeneity

In this section, we focus the results obtained with the restricted sample. Results obtained with the unrestricted sample are reported in the appendix. In this final part of the section, we present the results of two sensitivity analyses. Including the wage as an explanatory variables or not basically gives the same estimates for the other parameters - the estimated coefficients differ, at the maximum, by three hundredth between the two specifications. Below we therefore only present the results for the model with the wage included as an explanatory variable.

For the restricted sample, the estimates of the unobserved heterogeneity distributions are in Table 2. Results for shared parameters are quite similar for the three model specifications, meaning that increasing the unobserved heterogeneity complexity by considering a further parameter does not really affect the other results.[7]

In the three models, the standard deviation of the individual unobserved effects is estimated to be around 0.3 and is significant at the 5% level. In terms of unobserved heterogeneity at firm level, the estimates of the standard deviation are 0.6 and 0.7

[7] A similar remark is found in Horny *et al.* (2005) on a MPH model in continuous time and two random effects. In their study, maximum likelihood results are sensitive to a change in the unobserved heterogeneity structure.

Table 2: Estimates of the standard deviations of the unobserved heterogeneity distributions

Type of heterogeneity	Parameter	Mean	2.5%	97.5%
Correlated frailties				
firm effect	σ_f	0.61	0.48	0.75
worker effect	σ_w	0.29	0.22	0.37
correlation	ρ	0.50	0.29	0.58
Independent frailties				
firm effect	σ_f	0.72	0.58	0.88
worker effect	σ_w	0.26	0.20	0.33
Single frailty				
worker effect	σ_w	0.29	0.22	0.44

(both significant at 5% level) for the model with independent frailties and correlated frailties, respectively, indicating a large amount of unobserved heterogeneity at the firm level. Thus, our results suggest the need to consider unobservable component on both firm and worker level in job transitions analyses. Intuitively, job transition behavior depends on individual unobserved propensity to change jobs and on unobserved retention policies of the firms. The first characteristic is very dispersed across workers as is the second one across firms.

The correlation between the worker and firm effects is estimated to be positive, around 0.50 and significant at the 5% level.[8] Thus, our results suggest that the matching process between firms and workers is, at least partially, based on characteristics unobserved by the econometrician and it tends to follow an assortative pattern. Intuitively, firms with preference for a stable workforce tend hire low mobility workers and high mobility workers tend to search for firms with high workforce turnover. Notice that the estimated value of 0.50 is quite high in the light of the discussion in Subsection 3.2. The value does provide strong evidence for positive assortative matching in terms of job exit determinants, but at the same time it suggest that the maintained assumption of firm-specific (worker-specific) effects being independent across firms (workers) may be violated.

Figure 1 shows the contours and surface of the prior distribution, and Figure 2 the contours and surface of the prior evaluated using the estimates of Table 2. The figures show how the data affect our prior beliefs using Bayes' rule. We use fairly non-informative priors, allowing unlikely values of the parameters to not necessarily have a zero posterior probability.[9] The updated prior has its mass concentrated on

[8]The estimates of the standard deviations of the mixing distributions on the unrestricted sample are in Table 14, in Appendix C. The correlation is positive and significant, however, the assumption of independent v_i and independent w_j is likely to be violated on the unrestricted sample.

[9]Recall that the posterior equals the prior times the likelihood. Assigning a zero prior probability

a smaller support, meaning that information has been extracted from the data and can be used to enrich priors in further analyses.

Figure 1: Contour and surface of the prior mixing distribution

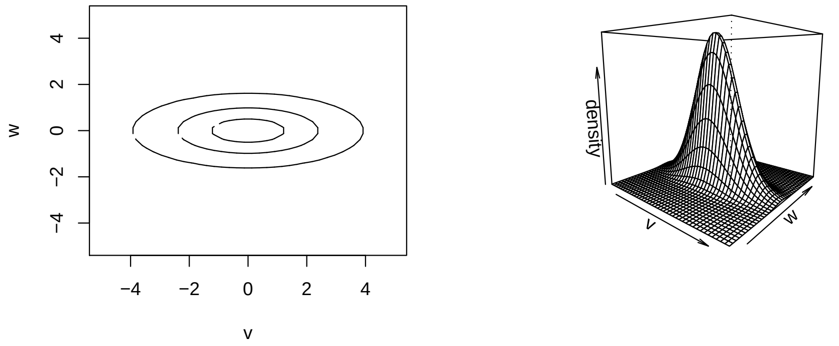

Figure 2: Contour and surface of the updated prior mixing distribution

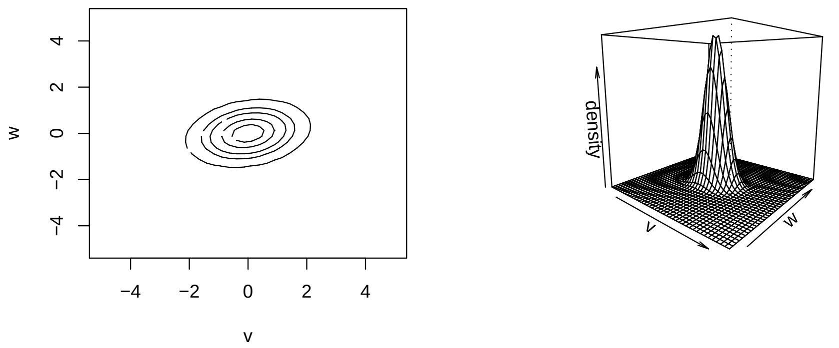

Table 3 depicts the so called "weeding out", i.e., the change, as elapsed duration increases, in the distributions of firm and worker unobserved heterogeneities among survivors.

5.2 Observed heterogeneity

The posterior means for the β coefficients together with information regarding their significance are reported in Table 4. [10] Negative duration dependence is found to be significant in all models, with the probability of separation declining monotonically

on some parameter values leads a zero posterior probability, even if the likelihood reaches its' maximum for these values.

[10]The estimates of the β on the unrestricted sample are in Table 15, in Appendix C. They are similar to the results on the restricted sample.

Table 3: Quantiles of the unobserved heterogeneity realizations

Type of heterogeneity	Min	25%	50%	75%	Max
t=0					
firm effect	-3.26	-0.49	0.00	0.49	3.41
worker effect	-1.11	-0.17	0.00	0.18	1.11
t=T					
firm effect	-3.26	-0.59	-0.12	0.33	2.60
worker effect	-1.11	-0.19	-0.02	0.16	1.05

with tenure (i.e., with the elapsed duration). This suggests that the empirically observed inverse relationship between separation rates and job tenure cannot be fully explained by pure heterogeneity models.

Regarding the controlled worker characteristics, we find that women tend to move less. This result contradicts the findings of many previous studies of job mobility. The main reason could be the fact that the gender difference in terms of mobility rates is changing over time. Indeed, Light and Ureta (1992) find that women's turnover behavior is changing: women belonging to early US birth cohorts appeared to be more mobile than men but this conclusion is reversed when more recent cohorts are considered.

The results for age are relative to the omitted category of workers with 36 to 55 years (the oldest age group considered in our study). Thus, they indicate higher transition probabilities for the younger workers. Notice that, controlling for education, age captures labor market experience and thus these estimates contradict the prediction of no-effect, typical from the pure heterogeneity models. Instead, these estimates can be interpreted under the light of on-the-job search models or models of job shopping. The first type of models predicts that, since the match quality is known ex-ante, more experienced workers are less mobile because they had already time/opportunity to move into high quality matches. Job shopping predicts that mobility decreases with age, as the worker becomes more aware of his own abilities and of the characteristics of the labor market.

Job transitions are also influenced by the education level of the worker. Workers with upper secondary and university education (the reference category in our estimates) are those moving less. The estimate of part-time job effect confirms one of the stylized facts of the empirical job duration literature: part-time job status has a strong positive effect on the probability of job separation.

Looking at the characteristics of the firms, we find some differences across economic sectors and across regions. The North (the reference category) is the region with the lowest job mobility, while Lisbon and Tagus Valley is at the other extreme. In fact, Lisbon is the largest city of the country and has the most developed and dynamic labor market. In terms of sectors, the financial sector exhibits the highest

Table 4: Bayesian estimates of β coefficients

Variable	**None**	**Worker**	**Random Effect(s)**	
			Independent	Correlated
Tenure				
2 years	**-0.59**	**-0.57**	**-0.38**	**-0.46**
3 years	**-0.92**	**-0.89**	**-0.61**	**-0.73**
4 years	**-1.42**	**-1.39**	**-1.06**	**-1.20**
5 years and more	**-2.11**	**-2.07**	**-1.70**	**-1.86**
Worker characteristics				
Female	**-0.31**	**-0.31**	**-0.35**	**-0.33**
Age:				
16 - 25	**0.65**	**0.65**	**0.75**	**0.71**
26 - 35	**0.37**	**0.37**	**0.44**	**0.41**
Education:				
primary school	**0.15**	**0.15**	**0.18**	**0.17**
lower secondary	**0.21**	**0.21**	**0.25**	**0.23**
Part-time	**0.65**	**0.65**	**0.69**	**0.67**
Wage	-0.04	-0.04	-0.04	-0.04
Firm characteristics				
Multiple plants	**0.30**	**0.30**	**0.35**	**0.32**
Region:				
Center	**0.14**	**0.15**	**0.18**	**0.17**
Lisbon and Tagus Valley	**0.38**	**0.39**	**0.46**	**0.43**
Alentejo, Algarve and Islands	**0.25**	**0.25**	**0.33**	**0.30**
Sector:				
Construction	**0.31**	**0.32**	**0.36**	**0.33**
Trade	**0.24**	**0.24**	**0.27**	**0.25**
Financial	**0.51**	**0.52**	**0.60**	**0.55**
Constant	**-2.33**	**-2.38**	**-2.87**	**-2.67**
Number of workers	7749	7749	7749	7749
Number of firms	6577	6577	6577	6577

Note: coefficients in bold are significant at 5% level.

job turnover rates while manufacture (the omitted sector) has the lowest ones.

5.3 Implications

We decompose the variation of the log job durations to separate the influences of three components: the variation due to the firm unobserved heterogeneity, the variation due to the worker unobserved heterogeneity and the variation due to the observed explanatory variables. We simulate the variance by drawing the observed and unobserved heterogeneity from the estimated and observed distributions, input them in likelihood (7) and obtain a precise approximation of var$(\log T_{ijk})$ with a sufficient number of drawings. Table 5 reports the results of the decomposition. In

Table 5: Decomposition of the total variation of the log durations

Source	**Random Effect(s)**	
	Independent	Correlated
observed variables	65%	52%
firm unobserved effect	31%	28%
worker unobserved effect	4%	12%
correlation		8%

the model with independent random effects, the observed characteristics of firms and workers included in the estimated model explain around two thirds of the variation of job durations. The remaining variation is mostly explained by the unobserved heterogeneity at firm level (28%) and the unobserved heterogeneity at worker level explains only 12% of the total variation. Allowing for correlation between the random effects changes mainly the influence of observed covariates, which falls to half of the total variation, and gives closer influences of the firm and worker effects. The correlation between the unobserved worker and firm effects is estimated to explain 8% of the variation on job durations. In sum, the results for the model with the most flexible heterogeneity structure indicate that the unobserved components explain half of the variation in job durations, and, thus, the firm and worker observed explanatory variables are clearly insufficient to capture the heterogeneity in job mobility decisions.

The bayes factor summarizes the evidence provided by the data in favor of one model, and we use it to compare the different models. It is the ratio of the probabilities of the data under the different assumed data generating processes (Lancaster, 2004), and we denote by $B_{1/2}$ the probability of the data under model 1 divided by the probability of the data under model 2. We consider twice the logarithm of the bayes factor, as suggested in Kass and Raftery (1995), which is on the same scale as the deviance and likelihood ratio test statistics. Table 6 displays the bayes factors corresponding to the ratio of the probabilities under the more complex models over the immediately simpler one. We conclude that there is a strong evidence in favor

Table 6: Bayes factors

	$2\ln(B_{w/no})$	$2\ln(B_{i/w})$	$2\ln(B_{c/i})$
Values	409.40	628.69	139.30

of models allowing for two-sided unobserved heterogeneity, among which the model with correlated frailties is the preferred one. The model allowing for a matching based on the unobserved is the most successful at predicting the data.

The building of the Markov chains is computer intensive and we also estimated the models with no unobserved heterogeneities and one random effect by maximum likelihood, using an adaptive Gauss-Hermite quadrature approximations of the mixing distribution when necessary. The gain of speed allowed us to use the full data and results are reported in Appendix D, Table 16. The β estimates are broadly similar to the Bayesian estimates. The likelihood improves with the inclusion of both the worker and the firm effects. Estimates indicate that worker and firm effects variances contribute to 15% and 25%, respectively, to total variance, when included separately.

5.4 Sensitivity analysis

We have performed two sensitivity analysis. The first one tests the sensitivity of the results to the restriction on the number of matches allowed in the sample. The results presented above were obtained from the sample with workers and firms with no more than 3 matches. Below we present results obtained from a sample with no more than 2 spells per worker and 2 no more than employees per firm.

The second sensitivity analysis aims to address the potential endogeneity of wages. We include as an explanatory variable the difference between the starting wage in the current job and the last wage in the previous job. From a job search point of view, we would expect that the larger this difference, the lower the exit rate out of the current job. In such a framework, the level of the wage in any job can be endogenous, but the difference between two jobs is due to the randomness in the matching process. So, in this estimation, we do not deal with wage changes within a job spell.

Table 7: Estimates of the standard deviations of the unobserved heterogeneity distributions - sensitivity analysis

Type of heterogeneity	Parameter	Mean	2.5%	97.5%
Sensitivity analysis: no. of matches				
firm effect	σ_f	0.47	0.31	0.64
worker effect	σ_w	0.27	0.21	0.36
correlation	ρ	0.60	0.30	0.70
Sensitivity analysis: wage endogeneity				
firm effect	σ_f	0.70	0.51	0.88
worker effect	σ_w	0.27	0.21	0.35
correlation	ρ	0.40	0.01	0.57

Results reported in table 7 show that the estimate of the standard deviation of the worker unobserved heterogeneity is insensitive to the number of matches allowed

Table 8: Bayesian estimates for the model with correlated random effects - sensitivity analysis

Variable	Sensitivity analysis	
	no. of matches	wage endogeneity
Tenure		
2 years	**-0.50**	**-0.40**
3 years	**-0.79**	**-0.76**
4 years	**-1.27**	**-1.30**
5 years and more	**-1.89**	...
Worker characteristics		
Female	**-0.32**	**-0.24**
Age:		
16 - 25	**0.70**	**0.57**
26 - 35	**0.37**	**0.32**
Education:		
primary school	**0.17**	**0.29**
lower secondary	**0.22**	0.20
Part-time	**0.65**	**0.72**
Wage	-0.04	...
Wage difference	...	**-0.09**
Firm characteristics		
Multiple plants	**0.33**	0.10
Region:		
Center	0.16	0.14
Lisbon and Tagus Valley	**0.40**	**0.39**
Alentejo, Algarve and Islands	**0.29**	**0.41**
Sector:		
Construction	**0.31**	**0.58**
Trade	**0.22**	**0.38**
Financial	**0.49**	**0.78**
Constant	**-2.58**	**-2.58**
Log-likelihood	-5145	-3708
DIC		7842
Number of workers	7006	5280
Number of firms	6261	4251

Note: coefficients in bold are significant at 5% level.

Table 9: Decomposition of total variation of the log durations with correlated random effects - sensitivity analysis

Source	Sensitivity analysis	
	no. of matches	wage endogeneity
observed variables	61%	48%
firm unobserved effect	21%	34%
worker unobserved effect	12%	12%
correlation	6%	7%

in the sample.[11] The estimate of the standard deviation of the firm effect is instead sensitive to this dimension. In this sample, it is lower - it decreases from 0.6 to 0.47. The decrease in the estimate of this parameter was already observed between the unrestricted and the restricted (up to 3 matches) sample. For what concerns the correlation between the random effects, it increases from 0.5 to 0.6.

The estimates for β coefficients (reported in table 8, column 2) are similar to those obtained with the restricted sample. Regarding the decomposition of total variance of log durations, in this sample (results in 9, column 2), the proportion of variance associated with firm unobserved effect decreases and it increases the proportion associated with observed explanatory variables. The contributions of the worker effects and the correlation are stable.

The inclusion of wage differences as explanatory variable does not affect the estimate of the standard deviation of the worker effect. For all models estimated, this parameter fell always between 0.25 and 0.30. The estimate of the standard deviation of the firm effect is here of 0.70 (0.10 higher than in the core results).[12] The correlation between the random effects is estimated to be somewhat lower (0.40).

Table 8, column 3, reports the β estimates. As expected, the estimate for the wage difference is significantly negative (-0.09), indicating that the larger the difference between the current wage and the wage in the previous job, the smaller the exit probability out of the current job.

The decomposition of total variance of log durations, in this sample (results in 9, column 3), indicates that the proportion of variance associated with the worker effects and with the correlation are stable. The variance coming from the firm side and from the observed characteristics change slightly.

[11] In this section, we present only estimates for the model with 2 correlated random effects. Estimates for the other models are available upon request.

[12] As mentioned before, results for the model without wages were extremely similar to those of the model with wage level, reported as core results. For this reason, estimates without wages were not reported.

6 Conclusion

This paper demonstrates how modern Bayesian Markov Chain Monte Carlo estimation methods can be fruitfully applied to estimate models of job durations with both worker-specific and firm-specific effects. In such models, the various unobserved worker-specific and firm-specific effects are not nested. We also examine the performance of the approach in case the effects are correlated between worker and firm. This expands the set of methods that can be used for the analysis of mobility and matching.

Our results reject a homogeneous view of the labor market, where firms adopt similar workforce management strategies and individuals have similar job change behavior. Instead, the estimates confirm the importance of the unobserved heterogeneity at the individual level, and indicate a large amount of unobserved heterogeneity at the firm level. Indeed, about 30% of the variation in the logarithm of job durations is due to variation in the effects of unobserved firm characteristics. Modelling the unobserved heterogeneity underlying job transitions as coming only from worker observables and unobservables, as is commonly done, is insufficient.

Results for the model allowing for correlation between the two random effects indicate a strong positive correlation. Thus, empirical evidence suggests that employer-employee matching tends to follow an assortative pattern in terms of unobservable characteristics of firms and workers - workers and firms with similar outcomes in terms of job mobility and turnover, respectively, tend to match together. As a topic for further research, it would be interesting to relate these findings to economic models of labor markets with mobility. For example, one may impose or test restrictions from economic models that relate the amount of job search frictions to the wage and the job duration; see e.g. Lise, Meghir and Robin (2008) for such models.

In fact, the magnitude of the estimated correlation is such that it suggests violation of the maintained assumption that worker-specific unobserved effects are independent across workers and that firm-specific unobserved effects are independent across firms. From a theoretical point of view it is plausible that at least the firm-specific unobservables are correlated across firms, since firms sometimes compete directly with a small number of other firms. At the same time, econometric analyses with microdata usually assume independence. Another interesting topic for further research would therefore be to estimate the size of such dependencies. However, it remains to be seen whether is it possible to design parsimonious but sufficiently general correlation matrices between the various random effects, such that the model is still estimable without insurmountable computational problems.

Appendix

A Descriptives of the explanatory variables

Table 10: Firm characteristics

Variable	Full data		Samples			
			Unrestricted		Restricted	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Multiple plants	0.39	0.49	0.38	0.48	0.29	0.45
Sector:						
Mining	0.01	0.08	0.01	0.05	0.01	0.05
Manufacturing	0.37	0.48	0.38	0.49	0.41	0.49
Electricity, gas, water	0.01	0.04	0.01	0.03	0.01	0.03
Construction	0.15	0.35	0.14	0.35	0.15	0.35
Trade, hotels, restaurants	0.30	0.46	0.31	0.46	0.32	0.46
Transport, communication	0.06	0.23	0.05	0.22	0.04	0.19
Finance, insurance and real estate	0.12	0.33	0.12	0.32	0.09	0.28
Region:						
North	0.35	0.48	0.36	0.48	0.38	0.49
Center	0.14	0.35	0.14	0.34	0.15	0.36
Lisbon, Tagus Valley	0.41	0.49	0.43	0.49	0.38	0.49
Alentejo, Algarve	0.06	0.23	0.05	0.21	0.05	0.22
Islands	0.05	0.21	0.03	0.17	0.03	0.17
Number of firms	55 325		6 582		6577	

Table 11: Worker characteristics

Variable	Full data		Samples			
			Unrestricted		Restricted	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Female	0.38	0.49	0.37	0.48	0.36	0.48
Age:						
16 - 25	0.34	0.47	0.34	0.47	0.31	0.46
26 - 35	0.40	0.49	0.40	0.49	0.39	0.49
36 - 55	0.26	0.44	0.26	0.44	0.29	0.45
Education:						
primary school	0.31	0.46	0.30	0.46	0.33	0.47
lower secondary	0.43	0.49	0.44	0.50	0.44	0.49
upper secondary and university	0.27	0.44	0.26	0.44	0.23	0.42
Part-time	0.08	0.27	0.06	0.24	0.06	0.23
Wage	703.80	449.63	699.49	441.88	670.36	419.93
Number of workers	338 445		9222		7749	

B Summary statistics of the durations

Table 12: Observed uncensored spells

Job spell duration	Full data	Samples	
		Unrestricted	Restricted
1	64.0%	68.2%	67.9%
2	19.9%	19.5%	19.8%
3	9.6%	7.9%	8.1%
4 or more	6.4%	4.3%	4.1%
Total	100%	100%	100%

Note: durations are in years.

Table 13: Number of spells per worker

Number of spells	Full data	Samples	
		Unrestricted	Restricted
1	90.6%	90.6%	91.9%
2	8.5%	8.5%	7.5%
3	0.9%	0.9%	0.6%
Total	100%	100%	100%

C Results based on the unrestricted sample

Table 14: Estimates of the standard deviations of the unobserved heterogeneity distributions - unrestricted sample

Type of heterogeneity	Parameter	Mean	2.5%	97.5%
Correlated frailties				
firm effect	σ_f	0.76	0.65	0.89
worker effect	σ_w	0.29	0.22	0.38
correlation	ρ	0.51	0.34	0.58
Independent frailties				
firm effect	σ_f	0.87	0.76	0.98
worker effect	σ_w	0.26	0.20	0.33
Single frailty				
worker effect	σ_w	0.30	0.22	0.41

Table 15: Bayesian estimates - unrestricted sample

Variable	**None**	**Worker**	**Random Effect(s)**	
			Independent	Correlated
Tenure				
2 years	**-0.59**	**-0.57**	**-0.26**	**-0.39**
3 years	**-0.95**	**-0.92**	**-0.49**	**-0.67**
4 years	**-1.34**	**-1.31**	**-0.79**	**-1.00**
5 years and more	**-2.11**	**-2.07**	**-1.49**	**-1.73**
Worker characteristics				
Female	**-0.28**	**-0.28**	**-0.35**	**-0.32**
Age:				
16 - 25	**0.55**	**0.56**	**0.75**	**0.68**
26 - 35	**0.31**	**0.32**	**0.41**	**0.36**
Education:				
primary school	**0.26**	**0.26**	**0.18**	**0.21**
lower secondary	**0.28**	**0.28**	**0.23**	**0.19**
Part-time	**0.56**	**0.57**	**0.66**	**0.62**
Wage	**-0.08**	**-0.08**	**-0.04**	**-0.03**
Firm characteristics				
Multiple plants	**0.21**	**0.21**	**0.32**	**0.27**
Region:				
Center	0.12	0.12	0.19	**0.19**
Lisbon and Tagus Valley	**0.37**	**0.38**	**0.46**	**0.41**
Alentejo, Algarve and Islands	**0.19**	**0.20**	**0.31**	**0.29**
Sector:				
Construction	**0.38**	**0.39**	**0.42**	**0.42**
Trade	**0.32**	**0.33**	**0.34**	**0.30**
Transports	0.07	0.07	**0.34**	**0.33**
Financial	**0.68**	**0.70**	**0.73**	**0.61**
Constant	**-2.41**	**-2.47**	**-3.06**	**-2.75**
Log-likelihood	-7695	-7565	-6130	-6645
DIC				
Number of workers	9222	9222	9222	9222
Number of firms	6582	6582	6582	6582

Note: coefficients in bold are significant at 5% level.

D Frequentist estimates

Table 16: Frequentist estimates - full data

Variable	None	Worker	Firm
Tenure			
2 years	**-0.54**	**-0.48**	**-0.40**
3 years	**-0.76**	**-0.66**	**-0.55**
4 years	**-1.06**	**-0.94**	**-0.82**
5 years or more	**-1.21**	**-1.06**	**-0.91**
Worker characteristics			
Female	**-0.24**	**-0.24**	**-0.29**
Age:			
16-25	**0.49**	**0.51**	**0.59**
26-35	**0.27**	**0.28**	**0.33**
Education:			
primary school	**0.20**	**0.21**	**0.04**
lower secondary	**0.19**	**0.20**	**0.06**
Part time	**0.58**	**0.60**	**0.56**
Wage	**-0.02**	**-0.02**	**0.04**
Firm characteristics			
Multiple plants	**0.18**	**0.18**	**0.25**
Region:			
Center	**0.14**	**0.15**	**0.21**
Lisbon and Tagus Valley	**0.34**	**0.36**	**0.40**
Alentejo and Algarve	**0.38**	**0.40**	**0.40**
Islands	**0.16**	**0.17**	**0.25**
Sector:			
Construction	**0.35**	**0.37**	**0.13**
Trade	**0.30**	**0.31**	**0.17**
Transports	**0.04**	**0.04**	**0.26**
Financial	**0.61**	**0.65**	**0.28**
Constant	**-2.32**	**-2.49**	**-2.46**
Log-likelihood	-303881	-303361	-274440
σ_w	-	0.54	-
σ_f	-	-	0.76
% total var		15	26
Number of workers	338445	338445	338445
Number of firms	55325	55325	55325

Note: coefficients in bold are significant at 1% level.

References

Bellmann, L., S. Bender, and U. Hornsteiner (2000): "Job tenure of two cohorts of young German men 1979 - 1990: An analysis of the (West-)German Employment Statistic Register Sample concerning multivariate failure times and unobserved heterogeneity," Discussion Paper Series 106, IZA.

Cardoso, A., and M. Portela (2005): "The provision of wage insurance by the firm: Evidence from a longitudinal matched employer-employee dataset," Working paper, IZA, Bonn.

Del Boca, D., and R. Sauer (2006): "Life cycle employment and fertility across institutional environments," mimeo.

Dostie, B. (2005): "Job turnover and the returns to seniority," *Journal of Business and Economic Statistics*, 23(2), 192–199.

Farber, H. (1999): "Mobility and stability: The dynamics of job change in labor markets," in *Handbook of Labor Economics*, vol. 3, chap. 37, pp. 2439–2483. O. Ashanfelter and D. Card (eds.), Elsevier, Amsterdam.

Gelfand, A., and A. Smith (1990): "Sampling based approaches to calculating marginal densities," *Journal of the American Statistical Association*, 85, 398–409.

Gelman, A., and D. Rubin (1992): "Inference from iterative simulation using multiple sequences," *Statistical Science*, 7, 457–472.

Grilli, L. (2005): "The random-effects proportional hazards model with grouped survival data: a comparison between the grouped continuous and continuation ration versions," *Journal of the Royal Statistical Society. Series A*, 168, 83–94.

Horny, G., B. Boockmann, D. Djurdjevic, and F. Laisney (2005): "Bayesian estimation of Cox models with non-nested random effects: An application to the ratification of ILO conventions by developing countries," Discussion Paper 05-23, ZEW.

Johnson, W. (1978): "A theory of job shopping," *Quarterly Journal of Economics*, 92(2), 261–278.

Kalbfleisch, J., and R. Prentice (1980): *The Statistical Analysis of Failure Time Data*. John Wiley, New York.

Kass, R., and A. Raftery (1995): "Bayes factors," *Journal of the American Statistical Association*, 90, 773–795.

Lancaster, T. (2004): *An Introduction to Modern Bayesian Econometrics*. Blackwell, Oxford.

Laurence, P., T. H. Wang, and L. Barone (2008): "Geometric properties of multivariate correlation in de Finetti's approach to insurance theory," *Electronic Journal for History of Probability and Statistics*, 4, 1–13.

Light, A., and M. Ureta (1992): "Panel estimates of male and female job turnover behavior: can female nonquitters be identified?," *Journal of Labor Economics*, 10, 156–181.

Lise, J., C. Meghir, and J. M. Robin (2008): "Matching, sorting and wages," Working paper, UCL, London.

Lopes de Melo, R. (2008): "Sorting in the labor market: theory and measurement," Working paper, Yale University.

Manda, S., and R. Meyer (2005): "Age at first marriage in Malawi: a Bayesian multilevel analysis using a discrete time model," *Journal of the Royal Statistical Society. Series A*, 168, 439–455.

Mendes, R., G. J. van den Berg, and M. Lindeboom (2007): "An empirical assessment of assortative matching in the labor market," Working paper, IZA, Bonn.

Robert, C., and G. Casella (1999): *Monte Carlo Statistical Methods*. Springer Verlag, Heidelberg.

Van den Berg, G. J. (2001): "Duration models: specification, identification and multiple durations," in *Handbook of Econometrics*, vol. 5, chap. 55, pp. 3381–3463. J. J. Heckman and E. Leamer (eds.), Elsevier, Amsterdam.

Vieira, J. A. C., A. R. Cardoso, and M. Portela (2005): "Gender segregation and the wage gap in Portugal: An analysis at the establishment level," *Journal of Economic Inequality*, 3, 145–168.