The Anatomy of the Wage Distribution. How do Gender and Immigration Matter?

Rasmus Lentz* Suphanit Piyapromdee† Jean-Marc Robin‡

February 16, 2016

Preliminary and incomplete

Abstract

Workers with similar observed characteristics may have different wage paths because their unobserved characteristics are rewarded differently or because they have different mobility patterns. For example, controlling for observed characteristics, a native worker may have a steeper wage profile than his counterpart immigrant because he has different unobserved characteristics or because he is more likely to be employed at a more productive firm (likewise for male versus female workers). Understanding the contributions of worker and firm heterogeneity to wage and mobility differentials can shed light on many key economic issues such as wage inequality, statistical discrimination and wage assimilation of immigrants.
In this paper, we propose an estimation method that allows for unrestricted interactions between worker and firm unobserved characteristics in both wages and the moving probability. Related to Bonhomme, Lamadon and Manresa (2014) (BLM), our method identifies double sided unobserved heterogeneity through an application of the EM-algorithm. The analysis estimates both wage and mobility patterns using Danish matched employer-employee data and decomposes wage dispersion into mobility and wage variation across firms and workers.

Keywords: On-the-job search; Heterogeneity; Wage distributions

JEL codes: E24; E32; J63; J64

*University of Wisconsin-Madison. E-mail: rlentz@wisc.edu
†University College London. Email s.piyapromdee@ucl.ac.uk
‡Sciences-Po, Paris and University College London. Email jmarc.robin@gmail.com

1 The Model

Workers are indexed by $i \in \{1,...,I\}$ and firms are indexed by $j \in \{1,...,J\}$. For each worker i, we observe (w_{it}, j_{it}) where $t = 1,2,...,T_i$ is either the last week of the year or any week within the year if a transition occurs in that week, and w_{it} denotes the logarithm of the wage rate. The variable $j_{it} \equiv j(i,t) \in \{1,...,J\}$ is the ID of the firm employing worker i at time t. Our framework can easily be extended to allow for control variables such as education or experience. We omit them to simplify the notations. In practice we shall cluster the data by worker's race, education and age group.

We assume that firms can be clustered into L different groups indexed by $\ell \in \{1,...,L\}$ and that workers can be clustered into K different groups indexed by $k \in \{1,...,K\}$. The index ℓ_j is the type of firm j and k_i is the type of worker i.We shall be treating the unobserved firm types $\mathscr{L} = (\ell_1,...,\ell_J)$ as a a fixed effect (i.e. a parameter to be estimated) and the worker type as a random effect mixing conditional distributions of workers' trajectories. Let $\pi = (\pi_1,...,\pi_K)$ denote the proportions of workers' types in the population.

Let $f_{k\ell}(w)$ denote the wage density, conditional on the worker's type k and the employer's type ℓ. We assume w_{it} and $w_{it'}$ independent conditional on $\ell_{j(i,t)}, \ell_{j(i,t')}$.

We further denote the probability for a worker of type k of making a transition from a firm of type ℓ to a firm of type ℓ' (possibly $\ell' = \ell$) at time t as $M_{k\ell\ell'}$, and the probability of staying with the same employer is $\overline{M}_{k\ell} = 1 - \sum_{\ell'=1}^{L} M_{k\ell\ell'}$.

2 EM algorithm

2.1 Likelihood with observed firm types and stationarity

Let us first consider the case where the employers' types ℓ_j are observed. Let $\ell_{it} = \ell_{j(i,t)}$ denote, by some abuse of notation, the type of the firm employing worker i in period t. Let also

$$D_{it} \equiv D(i,t) = \begin{cases} 1 & \text{if } j_{i,t+1} \neq j_{it} \\ 0 & \text{if } j_{i,t+1} = j_{it} \end{cases}$$

indicate an employer change between t and $t+1$.

For a value $\beta = (f, M, \pi)$ of the parameters and a classification $\mathscr{L}$ of firms, the likelihood for one worker i is

$$\sum_{k=1}^{K} L_i(k; \beta, \mathscr{L}),$$

where $L_i(k;\beta,\mathscr{L})$ is the individual likelihood conditional on worker type k, i.e.

$$L_i(k;\beta,\mathscr{L}) = \pi_k m_{k,\ell(i,1)} \prod_{t=1}^{T_i} f_{k,\ell(i,t)}(w_{it}) \prod_{t=1}^{T_i-1} \overline{M}_{k,\ell(i,t)}^{1-D(i,t)} M_{k,\ell(i,t),\ell(i,t+1)}^{D(i,t)}. \tag{1}$$

All the spell likelihoods are multiplied to form the likelihood of observing worker i's spell history.

Note that $m_{k,\ell(i,1)}$ is the probability that a type k worker is with a type-ℓ_{i1} firm at the first observation period $t=1$. These probabilities cannot be assumed structurally independent of the transition probabilities as the identification of sorting relies on mobility patterns. We thus link $m_{k\ell}$ to transition probabilities $M_{k\ell\ell'}$ by assuming them to be equal to the stationary probabilities solving the linear system:

$$m_{k\ell} = m_{k\ell}\overline{M}_{k\ell} + \sum_{\ell'=1}^{L} m_{k\ell'} M_{k,\ell',\ell}. \tag{2}$$

In matrix notations, $\mathbf{m}_k = [m_{k\ell}]_\ell$ is the eigenvector associated to the eigenvalue one of the transition matrix $\mathbf{M}_k^\top$ where:

$$\mathbf{M}_k = \begin{pmatrix} \overline{M}_{k1} + M_{k11} & M_{k12} & \cdots & M_{k1L} \\ M_{k21} & \overline{M}_{k2} + M_{k22} & \cdots & M_{k2L} \\ \vdots & \vdots & \cdots & \vdots \\ M_{kL1} & M_{kL2} & \cdots & \overline{M}_{kL} + M_{kLL} \end{pmatrix}.$$

Alternatively, using the constraint $\mathbf{1}_L^\top \mathbf{m}_k = 1$, where $\mathbf{1}_L = \text{ones}(L,1)$ is the L-vector of ones, one can rewrite $\mathbf{m}_k$ as the solution to

$$\mathbf{B}_k \mathbf{m}_k = \mathbf{1}_L, \quad \mathbf{B}_k = \mathbf{I}_L + (\mathbf{J}_L - \mathbf{M}_k^\top),$$

where $\mathbf{J}_L = \text{ones}(L,L)$. Hence, $\mathbf{m}_k = \mathbf{B}_k^{-1}\mathbf{1}_L$. This procedure guaranties that $\mathbf{m}_k$ sums to one and is a distribution of probability masses.

2.2 A cyclic EM algorithm

The EM algorithm (Dempster et al., 1977) consists of iterating the calculation of the posterior probabilities of worker types (E-step) and the expected log-likelihood maximization using the type probabilities calculated in the E-step (M-step). See Arcidiacono and Jones (2003); Bonhomme and Robin (2009); Arcidiacono and Miller (2011) for recent applications in economics.

For a given value of $\beta = (f, M, \pi)$, the posterior probability of worker i to be of type k given all wages and controls (all the available information) is

$$p_i(k;\beta,\mathscr{L}) = \frac{L_i(k;\beta,\mathscr{L})}{\sum_{k=1}^{K} L_i(k;\beta,\mathscr{L})}. \tag{3}$$

Then, define

$$Q_i(f,\mathscr{L};\beta^{(m)},\mathscr{L}^{(m)}) = \sum_{k=1}^{K} p_i(k;\beta^{(m)},\mathscr{L}^{(m)}) \left[\sum_{t=1}^{T_i} \ln f_{k,\ell(i,t)}(w_{it})\right]. \tag{4}$$

Q_i is the expected log-likelihood given worker i's wage and employer history for a given value $\widetilde{\beta}$ of the parameter. Let also

$$H_i(M,\mathscr{L};\beta^{(m)},\mathscr{L}^{(m)}) = \sum_{k=1}^{K} p_i(k;\beta^{(m)},\mathscr{L}^{(m)}) \Bigg[\ln m_{k,\ell(i,1)} + \sum_{t=1}^{T_i-1} \left\{(1-D_{it})\ln \overline{M}_{k,\ell(i,t)} + D_{it}\ln M_{k,\ell(i,t),\ell(i,t+1)}\right\}\Bigg]. \tag{5}$$

H_i is the expected log-likelihood given worker i's wage and employer history. Finally

The EM algorithm iterates the following steps.

E-step For $\beta^{(m)} = (f^{(m)}, M^{(m)}, \pi^{(m)})$ and $\mathscr{L}^{(m)}$ calculate posterior probabilities $p_i(k;\beta^{(m)},\mathscr{L}^{(m)})$.

M-step Update $\beta^{(m)}$ by maximizing $\sum_i \sum_k p_i(k;\beta^{(m)},\mathscr{L}^{(m)}) \ln L_i(k;\beta,\mathscr{L}^{(m)})$ subject to $\sum_k \pi_k = 1$, that is

$$f^{(m+1)} = \arg\max_f \sum_{i=1}^{I} Q_i(f,\mathscr{L}^{(m)};\beta^{(m)},\mathscr{L}^{(m)}), \tag{6}$$

$$M^{(m+1)} = \arg\max_M \sum_{i=1}^{I} H_i(M,\mathscr{L}^{(m)};\beta^{(m)},\mathscr{L}^{(m)}), \tag{7}$$

$$\pi_k^{(m+1)} = \frac{1}{I} \sum_{i=1}^{I} p_i(k;\beta^{(m)},\mathscr{L}^{(m)}). \tag{8}$$

C-step Update $\mathscr{L}^{(m)}$ by maximizing $\sum_i \sum_k p_i(k;\beta^{(m)},\mathscr{L}^{(m)}) \ln L_i(k;\beta^{(m+1)},\mathscr{L})$ with respect to

$\mathscr{L}$, that is

$$\mathscr{L}^{(m+1)} = \arg \max_{\mathscr{L}=(\ell_1,\dots,\ell_J)} \sum_{i=1}^{I} \sum_{k=1}^{K} p_i(k;\beta^{(m)},\mathscr{L}^{(m)}) \left[\sum_{t=1}^{T_i} \ln f_{k,\ell_{j(i,t)}}^{(m+1)}(w_{it}) \right]$$
$$+ \sum_{k=1}^{K} p_i(k;\beta^{(m)},\mathscr{L}^{(m)}) \left[\ln m_{k,\ell_{j(i,1)}}^{(m+1)} + \sum_{t=1}^{T_i-1} \left\{ (1-D_{it}) \ln \overline{M}_{k,\ell_{j(i,t)}}^{(m+1)} + D_{it} \ln M_{k,\ell_{j(i,t)},\ell_{j(i,t+1)}}^{(m+1)} \right\} \right]. \tag{9}$$

This is a cyclic EM algorithm because in the M-step we maximize with respect to β given $\mathscr{L}^{(m)}$ and in the C-step we maximize with respect to $\mathscr{L}$ given $\beta^{(m+1)}$. In practice we shall iterate the E and M steps keeping the firm classification constant until convergence (or for a few iterations) before running the C-step. The C-step is computationally too costly to be repeated too often.

2.3 Firm classification

In order to initialize the algorithm we follow Bonhomme et al. (2015) and cluster firms with similar observed characteristics. Let z_j denote a vector of characteristics for firm j. For example, z_j may contains the deciles of the distribution of all wages w_{it} such that $j(i,t) = j$, the average job duration at firm j, etc. One can then use any standard classification algorithm, such as k-means for example, to classify firms into L groups.

Next, the C-step is a very difficult optimization problem because of the high dimensionality of the control variable $\mathscr{L} = (\ell_1,\dots,\ell_J)$. Complexity stems from the fact that the posterior likelihood in (9) is non separable in $\mathscr{L} = (\ell_1,\dots,\ell_J)$ because of the terms $M_{k,\ell_{j(i,t)},\ell_{j(i,t+1)}}$. We suggest to iterate the following firm-specific optimizations,

$$\ell_j^{(p+1)} = \arg\max_{\ell_j} \sum_{i=1}^{I} \sum_{k=1}^{K} p_i(k;\beta^{(m)},\mathscr{L}^{(m)})) \left[\sum_{t=1}^{T_i} \ln f_{k,\ell_j}^{(m+1)}(w_{it}) \times \mathbf{1}\{j(i,t) = j\} \right]$$
$$+ \sum_{i=1}^{I} \sum_{k=1}^{K} p_i(k;\beta^{(m)},\mathscr{L}^{(m)}) \left[\ln m_{k,\ell_j}^{(m+1)} \times \mathbf{1}\{j(i,1) = j\} \right]$$
$$+ \sum_{i=1}^{I} \sum_{k=1}^{K} p_i(k;\beta^{(m)},\mathscr{L}^{(m)}) \left[\sum_{t=1}^{T_i-1} \left\{ (1-D_{it}) \ln \overline{M}_{k,\ell_j}^{(m+1)} + D_{it} \ln M_{k,\ell_j,\ell_{j(i,t+1)}^{(p)}}^{(m+1)} \right\} \times \mathbf{1}\{j(i,t) = j\} \right]. \tag{10}$$

In practice it suffices to iterate until we find a value $\mathscr{L}^{(m+1)}$ such that

$$\sum_{i=1}^{I}\sum_{k=1}^{K} p_i(k;\beta^{(m)},\mathscr{L}^{(m)})\ln L_i(k;\beta^{(m+1)},\mathscr{L}^{(m+1)}) > \sum_{i=1}^{I}\sum_{k=1}^{K} p_i(k;\beta^{(m)},\mathscr{L}^{(m)})\ln L_i(k;\beta^{(m+1)},\mathscr{L}^{(m)}).$$

3 Empirical specification

3.1 Wage distribution

Wages are assumed lognormal given match type. Specifically,

$$f_{k\ell}(w) = \frac{1}{\sigma_{k\ell}}\varphi\left(\frac{w-\mu_{k\ell}}{\sigma_{k\ell}}\right), \tag{11}$$

with $\varphi(x) = (2\pi)^{-1/2}e^{-x^2/2}$. This specification of the log-wage mean allows for a match-specific mean $\mu_{k\ell}$ and variance $\sigma_{k\ell}^2$.

The M-step update 6 takes the following form:

$$\mu_{k\ell}^{(m+1)} = \frac{\sum_{i=1}^{I} p_i(k;\beta^{(m)})\sum_{t=1}^{T_i} w_{it}}{\sum_{i=1}^{I} p_i(k;\beta^{(m)})T_i}, \quad \sigma_{k\ell}^{(m+1)} = \frac{\sum_{i=1}^{I} p_i(k;\beta^{(m)})\sum_{t=1}^{T_i}[w_{it}-\mu_{k\ell}^{(m+1)}]^2}{\sum_{i=1}^{I} p_i(k;\beta^{(m)})T_i}.$$

3.2 Transition probabilities

The probability for a worker of type k of a transition from a firm of type ℓ to a firm of type ℓ' at time t is

$$M_{k,\ell,\ell'} = \lambda_\ell \nu_{\ell'}\frac{\gamma_{k,\ell'}}{\gamma_{k\ell}+\gamma_{k,\ell'}}.$$

This is a generalization of the Bradley-Terry model (see e.g. Agresti, 2003; Hunter, 2004). Parameter $\lambda_\ell \in [0,1]$ is the probability of being poached when the current employer is of type ℓ. Parameter $\nu_{\ell'} \geq 0$, with $\sum_{\ell'=1}^{L}\nu_{\ell'} = 1$, is the probability that the poacher is a firm of type ℓ'. Parameter $\gamma_{k\ell}$, with $\sum_{\ell=1}^{L}\gamma_{k\ell} = 1$, measures the quality of the match (k,ℓ). [1]

[1] We also tried the specification

$$M_{k,\ell,\ell'} = \lambda_\ell \nu_{\ell'}\frac{\gamma_{k,\ell'}}{\theta\gamma_{k\ell}+\gamma_{k,\ell'}},$$

where $\theta > 0$ measures the incumbent's advantage and parametrizes mobility within the same group of firms. However it appeared difficult to disentangle θ from λ.

3.3 M-step update for transition probabilities

In the M-step of the EM algorithm, we maximize the part of the expected likelihood that refers to transitions, i.e.

$$\widetilde{H}(M;\beta^{(m)}) \equiv \sum_{k=1}^{K}\sum_{\ell=1}^{L}\left\{\overline{n}_{k\ell}(\beta^{(m)})\ln\overline{M}_{k\ell}+\sum_{\ell'=1}^{L}n_{k\ell\ell'}(\beta^{(m)})\ln M_{k\ell\ell'}\right\},$$

where

$$\begin{aligned}\overline{n}_{k\ell}(\beta^{(m)}) &= \sum_i p_i(k;\beta^{(m)})\#\{t:D_{it}=0,\ell_{it}=\ell\},\\ n_{k\ell\ell'}(\beta^{(m)}) &= \sum_i p_i(k;\beta^{(m)})\#\{t:D_{it}=1,\ell_{it}=\ell,\ell_{i,t+1}=\ell'\},\end{aligned}$$

where $\#\{\}$ denotes the cardinality of a set. This likelihood is similar to the likelihood of a Bradley-Terry model except that when the incumbent firm ℓ wins we do not know against which ℓ'. The likelihood is thus rendered more nonlinear by the presence of the term in $\ln\overline{M}_{k\ell}$. An MM algorithm can still be developed as follows.

Because the logarithm is concave, we can minorize $\ln\overline{M}_{k\ell}$ as follows. Let $P_{k\ell\ell'} \equiv \frac{\gamma_{k\ell}}{\gamma_{k\ell}+\gamma_{k\ell'}}$ denote the probability that the incumbent firm be preferred to the poacher. Then, with obvious notations,

$$\begin{aligned}\ln\overline{M}_{k\ell} &= \ln\left(1-\sum_{\ell'=1}^{L}M_{k\ell\ell'}\right) = \ln\left(1-\lambda_\ell+\sum_{\ell'}\lambda_\ell\nu_{\ell'}P_{k\ell\ell'}\right)\\ &\geq \frac{1-\lambda_\ell^{(s)}}{\overline{M}_{k\ell}^{(s)}}\ln\left(\frac{1-\lambda_\ell}{1-\lambda_\ell^{(s)}}\overline{M}_{k\ell}^{(s)}\right)+\sum_{\ell'=1}^{L}\frac{\lambda_\ell^{(s)}\nu_{\ell'}^{(s)}P_{k\ell\ell'}^{(s)}}{\overline{M}_{k\ell}^{(s)}}\ln\left(\frac{\lambda_\ell\nu_{\ell'}P_{k\ell\ell'}}{\lambda_\ell^{(s)}\nu_{\ell'}^{(s)}P_{k\ell\ell'}^{(s)}}\overline{M}_{k\ell}^{(s)}\right).\end{aligned}$$

Note that both sides of the inequality are equal if $\beta=\beta^{(s)}$ (no parameter change).

Let

$$\widetilde{n}_{k\ell\ell'}^{(s)} = \overline{n}_{k\ell}^{(m)}\frac{\lambda_\ell^{(s)}\nu_{\ell'}^{(s)}P_{k\ell\ell'}^{(s)}}{\overline{M}_{k\ell}^{(s)}}.$$

This is the predicted number of times that home beats visitor ℓ'. Given initial values $\lambda_\ell^{(s)},\nu_{\ell'}^{(s)}$ one can update $\gamma^{(s)}$ so as to maximize

$$\sum_{k=1}^{K}\sum_{\ell=1}^{L}\sum_{\ell'=1}^{L}\left\{\widetilde{n}_{k\ell\ell'}^{(s)}\ln\frac{\gamma_{k\ell}}{\gamma_{k\ell}+\gamma_{k\ell'}}+n_{k\ell\ell'}\ln\frac{\gamma_{k\ell'}}{\gamma_{k\ell}+\gamma_{k\ell'}}\right\},$$

subject to the normalization $\sum_{\ell=1}^{L}\gamma_{k\ell}=1$. Now, because

$$-\ln(\gamma_{k\ell}+\gamma_{k\ell'}) \geq 1-\ln(\gamma_{k\ell}^{(s)}+\gamma_{k\ell'}^{(s)})-\frac{\gamma_{k\ell}+\gamma_{k\ell'}}{\gamma_{k\ell}^{(s)}+\gamma_{k\ell'}^{(s)}}$$

(see Hunter, 2004), we can instead maximize

$$\sum_{k=1}^{K}\sum_{\ell=1}^{L}\left(\sum_{\ell'=1}^{L}(\widetilde{n}_{k\ell\ell'}^{(s)}+n_{k\ell'\ell})\right)\ln\gamma_{k\ell}-\sum_{k=1}^{K}\sum_{\ell=1}^{L}\sum_{\ell'=1}^{L}\left((\widetilde{n}_{k\ell\ell'}^{(s)}+n_{k\ell\ell'})\frac{\gamma_{k\ell}+\gamma_{k\ell'}}{\gamma_{k\ell}^{(s)}+\gamma_{k\ell'}^{(s)}}\right).$$

That is (taking special care to indices),

$$\gamma_{k\ell}^{(s+1)} \propto \left(\sum_{\ell'}(\widetilde{n}_{k\ell\ell'}^{(s)}+n_{k\ell'\ell})\right)\left[\sum_{\ell'}\frac{\widetilde{n}_{k\ell\ell'}^{(s)}+n_{k\ell\ell'}+\widetilde{n}_{k\ell'\ell}^{(s)}+n_{k\ell'\ell}}{\gamma_{k\ell}^{(s)}+\gamma_{k\ell'}^{(s)}}\right]^{-1},$$

where $X_\ell \propto Y_k$ means $X_\ell = Y_\ell/\sum_\ell Y_\ell$, that is $\gamma_{k\ell}^{(s+1)}$ should sum to one over ℓ.

Update $\lambda^{(s)}$ by maximizing

$$\sum_{\ell=1}^{L}\left(\sum_{k=1}^{K}\left(\overline{n}_{k\ell}\frac{1-\lambda_\ell^{(s)}}{\overline{M}_{k\ell}^{(s)}}\right)\ln(1-\lambda_\ell)+\sum_{k=1}^{K}\sum_{\ell'=1}^{L}\left(\widetilde{n}_{k\ell\ell'}^{(s)}+n_{k\ell\ell'}\right)\ln\lambda_\ell\right).$$

Let

$$\lambda_\ell^{(m+1)}=\left[\sum_{k,\ell'}\left(\widetilde{n}_{k\ell\ell'}^{(m)}+n_{k\ell\ell'}^{(m)}\right)\right]\left[\sum_{k}\left(\overline{n}_{k\ell}^{(m)}\frac{1-\lambda_\ell^{(m)}}{\overline{M}_{k\ell}^{(m)}}\right)+\sum_{k,\ell'}\left(\widetilde{n}_{k\ell\ell'}^{(m)}+n_{k\ell\ell'}^{(m)}\right)\right]^{-1}.$$

Finally update $\nu^{(s)}$ by maximizing

$$\sum_{\ell'=1}^{L}\left[\sum_{k=1}^{K}\sum_{\ell=1}^{L}\left(\widetilde{n}_{k\ell\ell'}^{(s)}+n_{k\ell\ell'}\right)\right]\ln\nu_{\ell'} \quad \text{s.t.} \quad \sum_{\ell'=1}^{L}\nu_{\ell'}=1.$$

That is

$$\nu_\ell^{(s+1)}=\frac{\sum_{k=1}^{K}\sum_{\ell'=1}^{L}\left[\widetilde{n}_{k\ell'\ell}^{(s,m)}+n_{k\ell'\ell}^{(m)}\right]}{\sum_{\ell=1}^{L}\sum_{k=1}^{K}\sum_{\ell'=1}^{L}\left[\widetilde{n}_{k\ell'\ell}^{(s,m)}+n_{k\ell'\ell}^{(m)}\right]}.$$

4 Monte Carlo simulations

TBC.

A Simulation notes

Specification

Workers differ by x distributed as a standard log-normal variable. Firms differ by y also distributed as a log-normal random variable. A match (x, y) produces surplus $S(x, y) = xy - (x - y)^2$. The log-wage is $w = \ln[S(x, y)/2] + \varepsilon$, where ε is normally distributed. The probability of preferring a firm y to a firm z is $S(x, y)/[S(x, y) + S(x, z)]$.

Link structure and initialization

There are I individuals and J firms. First assign firm ID's to firm type classes. Assign a type $k \in \{1, ..., K\}$ and a type $\ell \in \{1, ..., L\}$ uniformly to all workers and firms. Then assign $\ln x = \Phi^{-1}(k/K)$ to all workers of type k and $\ln y = \Phi^{-1}(\ell/L)$ to all firms of class ℓ.

Let χ_ℓ be the vacancy intensity of firms type ℓ. In that case, the probability of drawing a type-ℓ firm is given by,

$$\nu_\ell = \frac{J_\ell \chi_\ell}{\sum_{\ell'} J_{\ell'} \chi_{\ell'}}.$$

For each worker draw an initial match firm type from $m_{k\ell}$, conditional on k. For the ℓ realization, draw a firm id j from group of ℓ type firm id's. With this procedure, there will be no simulation noise in the number of firms by firm class and there will be no noise in the number of workers by worker type class.

Durations, transitions and wages

Simulate spell lengths where the spell hazard in a job (k, ℓ) is given by $h_{k\ell} = \sum_{\ell'} M_{k\ell\ell'} = 1 - \overline{M}_{k\ell}$. First simulate the duration. The probability of a duration n is geometrically distributed, $\Pr(dur = T) = (1 - h_{k\ell})^{T-1} h_{k\ell}$, with CDF $G(T) = 1 - (1 - h_{k\ell})^T$. Duration is found by inverting the CDF. Take a uniformly distributed random number, u, invert the CDF,

$$u = 1 - (1 - h_{k\ell})^T \Leftrightarrow T \ln(1 - h_{k\ell}) = \ln(1 - u) \Leftrightarrow T = \frac{\ln(1 - u)}{\ln(1 - h_{k\ell})} + 1.$$

Conditional on the spell ending, then simulate the destination. The probability that the worker moves to a type ℓ' firm conditional on moving is $M_{k\ell\ell'}/h_{k\ell}$. After the fact, for each year of the duration of the simulated spell, simulate wages.

References

Agresti, A. (2003). *Categorical Data Analysis*. Wiley Series in Probability and Statistics. Wiley.

Arcidiacono, P. and J. B. Jones (2003, 05). Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm. *Econometrica 71*(3), 933–946.

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 79*(6), 1823–1867.

Bonhomme, S., T. Lamadon, and E. Manresa (2015). A Distributional Framework for Matched Employer Employee Data. mimeo, University of Chicago.

Bonhomme, S. and J.-M. Robin (2009, 01). Assessing the equalizing force of mobility using short panels: France, 1990-2000. *Review of Economic Studies 76*(1), 63–92.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological) 39*(1), pp. 1–38.

Hunter, D. R. (2004, 02). Mm algorithms for generalized bradley-terry models. *Ann. Statist. 32*(1), 384–406.