A Distributional Framework for Matched Employed Employee Data

S. Bonhomme, T. Lamadon, and E. Manresa, ECMA2019

Presenter: Zhongji Wei

December 4, 2023

Question Description & Concepts

- Some questions in Labor
  - What causes earning dispersion : individual or firm?
  - The nature of sorting pattern : why someone in some firm?
- Some key concepts for explanation
  - Matched Data : who in which firm
  - Heterogeneity and Interaction
  - Sorting and Complementarity, Becker (1973)
  - Two-sided Unobserved Heterogeneity
  - Distributional framework: explain from identified distribution

Contents

Introduction

Introduction: Previous Approach

- Approach: identify the contribution of worker and firm (2-sided unobserved) heterogeneity to earning dispersion
- Two angles: reduced and structural
- Reduced: Two Way Fixed Effect, AKM(1999)

$$y_{it} = \mu_y + (x_{it} - \mu_x) + \underbrace{\theta_i}_{\text{Pure Person Effect}} + \underbrace{\psi_{J_{(i,t)}}}_{\text{Pure Firm Effect}} + \varepsilon_{it}$$

  - Lack of firm-worker interaction, restrics complementarity
  - Static: lack of previous earning & firm dependence

Introduction: Previous Approach

- Structural: full-specified theoretical models
  - Example: wage posting, bargaining
  - Portray the interaction between worker and firms
  - Empirical challenge: worker $\times$ firms, curse of dimensionality
  - May be driven by functional form

Introduction: This Paper

- This paper: an empirical framework to reconcile both angles
  - Allowing complementarities, sorting, and dynamics
  - Dimension reduction: discrete firm class and worker type

$$Y_{it} = \underbrace{\rho_t Y_{i,t-1}}_{Dynamic} + a_{1t}(k_{it}) + \underbrace{a_{2t}(k_{i,t-1})}_{Dynamic} + \underbrace{b_t(k_{it})}_{Interact} \alpha_i + X_{it}' c_t + v_{it}$$

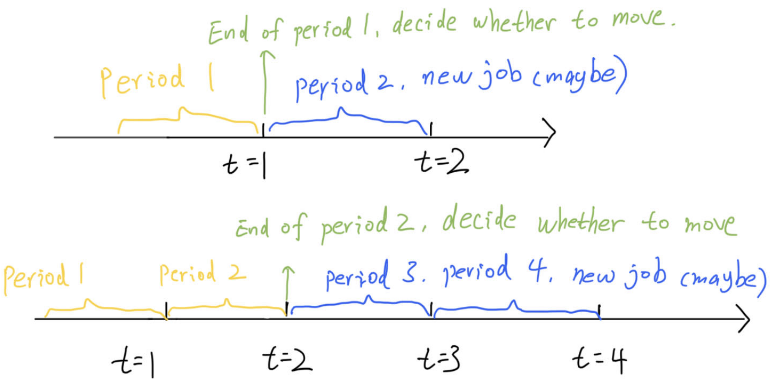
- Identification of income and worker distribution
- 2-step estimation
  - Classification: k-means for firm grouping, given worker type
  - Estimation: estimate key parameters using MLE

Framework of Analysis

Framework: basic settings

- Firms
  - J firms into K classes, $k_{it} = k(j_{it}) \in \{1, 2, ..., K\}$, $j_{it} \in \{1, 2, ..., J\}$
- Workers
  - N workers, at period t, worker i's "state": $(Y_{it}, j_{it}, m_{it}, X_{it})$
  - Earning, firm class, job moving choice at the end of period, other characteristics
  - Each worker belongs to a type α_i time-invariantly
- History of Z_{it}: $Z_i^t = (Z_{i1}, ..., Z_{it})$
- What to find (identify)
  - Given (k, α), the earning distribution
  - The proportion of type α worker in class k firm

Framework: Timing



Framework: Static Model

- "Model": what determines (the distribution of) a variable
- **Assumption 1.1 (Mobility determinant)**:

$$m_{it} \sim F_m(\cdot \mid \alpha_i, k_i^t, m_i^{t-1}, X_i^t), \perp Y_i^t$$

Similar for $k_{i,t+1}$ and $X_{i,t+1}$

- **Assumption 1.2 (Serially Independence)**

$$Y_{it+1} \sim F_Y(\cdot \mid \alpha_i, k_{it+1}, X_{it+1}, m_{it} = 1), \perp Y_i^t, k_i^t, m_i^{t-1}, X_i^t$$

- Example, reduced to AKM when $b_t(k) = 1$, $K = J$

$$Y_{it} = a_t(k_{it}) + b_t(k_{it})\alpha_t + X_{it}' c_t + \varepsilon_{it} \tag{1}$$

where $E(\varepsilon_{it} \mid \alpha_i, k_i^T, m_i^T, X_i^T) = 0$

Framework: Dynamic Model

- Introduce dynamic using first-order Markov Property
- **Assumption 2.1 (Mobility determinant)**

$$m_{it} \sim F_m(\cdot \mid Y_{it}, \alpha_i, k_i^t, m_i^{t-1}, X_i^t), \perp Y_i^{t-1}, k_i^{t-1}, m_i^{t-1}, X_i^{t-1}$$

Similar for k_{it+1} and X_{it+1}

- **Assumption 2.2 (Serial Dependence)**

$$Y_{it+1} \sim F_Y^{t+1}(\cdot \mid Y_{it}, \alpha_i, k_{it+1}, k_{it}, X_{it+1}, m_{it}), \perp Y_i^{t-1}, k_i^{t-1}, m_i^{t-1}, X_i^t$$

- Example

$$Y_{it} = \rho_t Y_{it-1} + a_{1t}(k_{it}) + a_{2t}(k_{it-1}) + b_t(k_{it})\alpha_i + X_{it}' c_t + v_{it} \quad (2)$$

where $E(v_{it} | \alpha_i, k_i^t, m_i^{t-1}, Y_i^{t-1}, X_i^t) = 0$, $m_{it-1} = 1$.

Framework: Theoretical base

It can include varieties of models

- Non-linear wage function: $w(\alpha_t, k_t, \varepsilon_t)$
  - Just static setting
- Time effect: boom or bust
  - Let Markovian to be non-homogeneous
- Match-specific heterogeneity and observable potential wage $\tilde{Y}_{t+1}$
  - Jointly Markov: $(Y^*_{t+1}, k^*_{t+1}, \tilde{Y}_{t+1}) \sim F_{Y,k,\tilde{Y}}(\cdot,\cdot,\cdot|\alpha, Y_t, k_t)$
- Outside this framework
  - Non-markov: permanent-transitory earning;
  - Comment: distribution may be useful, but story of power absent
  - Unemployment state not considered

Identification

Identification: what to identify?

Generally speaking, four targets to identify:

- Move from here to there or not: $p_{kk'}(\alpha)$
- The proportion of each type of worker in a firm class: $q_k(\alpha)$
  - Above two: sorting pattern
- Earning if leave: $F^m_{k'\alpha}(y)$
- If not leave: $F_{k\alpha}(y)$
  - Above two: complementarities

Note: Class k to be estimated, type α can be arbitrary labeled

Identification: intuition of conditions

The key to identification this is a rank condition

- Consider workers moving from class k' to k and vice versa, then:

$$Y_{i1} = a(k') + b(k')\alpha_i + \varepsilon_{i1} \quad Y_{i2} = a(k) + b(k)\alpha_i + \varepsilon_{i2}$$

and

$$Y_{i1} = a(k) + b(k)\alpha_i + \varepsilon_{i1} \quad Y_{i2} = a(k') + b(k')\alpha_i + \varepsilon_{i2} \tag{3}$$

then

$$\frac{b(k')}{b(k)} = \frac{E_{kk'}(Y_{i2}) - E_{k'k}(Y_{i1})}{E_{kk'}(Y_{i1}) - E_{k'k}(Y_{i2})} \tag{4}$$

where $E_{kk'} = E(\cdot | k_{i1} = k, k_{i2} = k', m_{i1} = 1)$

Identification: intuition of conditions

- To identify the interaction effect $\frac{b(k')}{b(k)}$(explain), we need:

$$E_{kk'}(\alpha_i) \neq E_{k'k}(\alpha_i) \tag{5}$$

i.e.

$$E_{kk'}(Y_{i1} + Y_{i2}) \neq E_{k'k}(Y_{i1} + Y_{i2}) \tag{6}$$

which can be empirically tested

- 6 in fact is a rank condition

Identification: static

Let type α to be discrete, $F_z^m(\cdot) = F(\cdot|z, m = 1)$

- For a job mover from k to k', we have:

$$Pr[Y_{i1} \leq y_1, Y_{i2} \leq y_2 | k_{i1} = k, k_{i2} = k', m_{i1} = 1]$$
$$= \overbrace{\sum_{\alpha=1}^{L} \underbrace{F_{k\alpha}(y_1) F_{k'\alpha}^m(y_2)}_{independence} p_{kk'}(\alpha)}^{\text{Take expectation}} \tag{7}$$

- $F_{k'\alpha}^m(y_2)$: log-earnings' cdf in period 2, for α worker, k' firm
- $p_{kk'}(\alpha)$: proportion of α workers among those from k to k'
- $F_{k\alpha}(y_1)$: log-earnings' cdf in period 1, for α worker in k firm

Identification: static

- And log-earnings' cdf in period 1 in k firm:

$$Pr[Y_{i1} \leq y_1 | k_{i1} = k] = \sum_{\alpha=1}^{L} F_{k\alpha}(y_1) q_k(\alpha) \tag{8}$$

- $q_k(\alpha)$: proportion of α workers in k firm
- Question: In what conditions can they be well-identified?

Identification: static

Definition (Connecting cycle of length R)

A pair of sequences of classes $(k_1, ..., k_R)$ in period 1, $(\tilde{k}_1, ..., \tilde{k}_R)$ in period 2,$k_{R+1} = k_1$, s.t.

- $p_{k_r,\tilde{k}_r}(\alpha) \neq 0$
- $p_{k_{r+1},\tilde{k}_r}(\alpha) \neq 0$

, $\forall (r, \alpha) \in \{1, ..., R\} \times \{1, ..., L\}$

- How to understand this?
- Both "stay" and "leave" are possible
- Communicate and accessible
- $\tilde{k}$ is like a re-ordering

Identification: static

Assumption 3:Mixture Model, Static

- **Assumption 3.1**:(Accessibility and communicativeness) $\forall k \neq k' \in \{1, ..., K\}$, $\exists$ connecting cycle $(k_1, ..., k_R)$ and $(\tilde{k}_1, ..., \tilde{k}_R)$,s.t. $\exists r, k_1 = k,\ k_r = k'$, and scalar $a(1), ..., a(L)$ are distinct, where:

$$a(\alpha) = \frac{p_{k_1,\tilde{k}_1}(\alpha)...p_{k_R,\tilde{k}_R}(\alpha)}{p_{k_2,\tilde{k}_1}(\alpha)...p_{k_1,\tilde{k}_R}(\alpha)}$$

Identification: static

- **Assumption 3.2**:(Rank condition) $\exists$ finite sets including M y_1,y_2, s.t. $\forall r \in \{1, ..., R\}$, matrix $A(k_r, \tilde{k}_r)$ and $A(k_{r+1}, \tilde{k}_r)$ have rank L, where $A_{R\times R}(k, k')$ has (y_1, y_2) element:

$$Pr[Y_{i1} \leq y_1, Y_{i2} \leq y_2 | k_{i1} = k, k_{i2} = k', m_{i1} = 1]$$

Identification: static

Theorem (Well-identification)

Let $T = 2$ and Assumptions 1,3 hold. Suppose firm classes are observed. Then, up to labeling of types α, $F_{k\alpha}$ and $F^m_{k'\alpha}$ are identified for $\forall(\alpha, k, k')$.
$\forall(k, k')$, k, $p_{kk'}(\alpha)$, $q_k(\alpha)$ is identified for all α, for the same labeling

- Label of α can be arbitrary
- Identification is up to (α, k, k') with high degree of freedom

Identification: dynamic

- Backward and forward cdf $G^b_{y_3,k',\alpha}(y_4)$ and $G^f_{y_2,k,\alpha}(y_1)$
  - Impact of previous and future job and earning
  - Can be recovered from data; Interpret "forward"
- Proportion $p_{y_2,y_3,k,k'}(\alpha)$, then we have:

$$
\begin{aligned}
&Pr[Y_{i1} \leq y_1, Y_{i4} \leq y_4 | Y_{i2} = y_2, Y_{i3} = y_3, \\
&k_{i1} = k_{i2} = k, k_{i3} = k_{i4} = k', m_{i1} = 0, m_{i2} = 1, m_{i3} = 0] \\
&= \sum_{\alpha=1}^{L} G^b_{y_3,k',\alpha}(y_4) G^f_{y_2,k,\alpha}(y_1) p_{y_2,y_3,k,k'}(\alpha) \qquad (9)
\end{aligned}
$$

And finally

$$
\begin{aligned}
&Pr[Y_{i1} \leq y_1, Y_{i2} \leq y_2 | k_{i1} = k_{i2} = k, m_{i1} = 0] \\
&= \sum_{\alpha=1}^{L} G^f_{y_2,k\alpha}(y_1) F_{k\alpha}(y_2) q_k(\alpha) \qquad (10)
\end{aligned}
$$

Identification: dynamic

- This identify the pattern of
  - Earning distribution for job movers who move at the end of period 2
  - Income distribution of all workers in period 1
- With similar conditions to static case, it well-identifies the distribution

Two-step Estimation

What to estimate

- Estimate the classification of firms
- Using the estimated class, recover the distribution of earnings and workers by parameter estimation
- EM algorithms plays an important role for both steps

Introduction to EM algorithm

Steps (**E**xpectation **M**aximization algorithm)

Input: observed data $x = (x^{(1)}, ..., x^{(m)})$*, joint distribution* $p(x, z|\theta)$*, conditional distribution* $p(z|x, \theta)$*, maximum iteration* J

- *Step 1: randomly initialize parameter* θ *by* θ^0
- *Step 2: for* j *in* $1 : J$*:*
  - **E-step**: *calculate*

$$Q_i(z^{(i)}) \equiv P(z^{(i)}|x^{(i)}, \theta)$$

  - **M-step**: *get maximal* θ*:*

$$\theta \equiv arg \max_{\theta} L(\theta) = \sum_{i=1}^{m} \sum_{z^{(i)}} Q_i(z^{(i)}) log P(x^{(i)}, z^{(i)}|\theta)$$

Repeat E and M until converge, output θ

Discussion about EM algorithm

- EM is a kind of heuristic algorithm, all estimation below use it
- May be sensitive to initial parameter: initialize many times
- Must converge to a stationary point
- If $L(\theta, \theta^j) \equiv \sum_{i=1}^{m} \sum_{z^{(i)}} P(z^{(i)}|x^{(i)}, \theta^j) logP(z^{(i)}|x^{(i)}, \theta)$ convex, then global maximal

Firm Classification: k-Means clustering

- Assume that firms' heterogeneity is only in class level, we have:

$$Pr[Y_{i1} \leq y_1 | j_{i1} = j] = \sum_{\alpha=1}^{L} F_{k\alpha}(y_1) q_k(\alpha) \tag{11}$$

- Partition J firms into K classes, K exogenous by solving the weighted k-means problem

$$\min_{k(1),\ldots,k(J),H_1,\ldots,H_K} \sum_{j=1}^{J} n_j \int (\hat{F}_j(y) - H_{k(j)}(y))^2 d\mu(y) \tag{12}$$

- $\hat{F}$: empirical distribution of j, In practice: get empirical distribution by griding the support of j by percentiles
- $H_{k(j)}$: targeted distribution we want to find
- y can be log-earning, or other variables

Recover distribution: static

- Already have estimated $\hat{k}_{it}$
  - $\hat{k}(j) \overset{J\to\infty}{\to}$ population one, $\forall$ labeling
- Maximize the log-likelihood below:

$$\sum_{i=1}^{N_m} \sum_{k=1}^{K} \sum_{k'=1}^{K} \mathbf{1}\{\hat{k}_{i1} = k\}\mathbf{1}\{\hat{k}_{i2} = k'\}$$

$$ln(\sum_{\alpha=1}^{L} \underbrace{p_{kk'}(\alpha;\theta_p)}_{\theta_p:\ \text{prop param log-norm},} \quad \underbrace{f_{k\alpha}(Y_{i1};\theta_f) f^m_{k'\alpha}(Y_{i2};\theta_{f^m})}_{\theta_f,\ \theta_f^m:\ (k,\alpha)\ \text{specific mean-var}} \quad) \tag{13}$$

- Interpret: Likelihood of worker in α moves from k to k' and gets income Y_{i1}, Y_{i2} before and after. N_m: number of job-movers.

Recover distribution: static

- After getting $\hat{\theta}_f$, maximize:

$$\sum_{i=1}^{N}\sum_{k=1}^{K}\mathbf{1}\{k_{i1} \hat{=} k\} ln(\sum_{\alpha=1}^{L} \underbrace{q_k(\alpha;\theta_q)}_{\theta_q:\ \text{prop parameter}} f_{k\alpha}(Y_{i1};\hat{\theta}_f)) \tag{14}$$

- Interpret: Likelihood of all workers' earning pattern in period 1
- Parameter vector $(\hat{\theta}_f, \hat{\theta}_{f^m}, \hat{\theta}_p, \hat{\theta}_q)$ characterizes the sorting and complementarity pattern

Recover distribution: dynamic

- A specific parametric form for the G^f, G^b defined before:

$$E[Y_{i1}|Y_{i2},k,\alpha] = \mu_{1k\alpha} + \rho_{1|2}Y_{i2}$$
$$E[Y_{i4}|Y_{i3},k',\alpha] = \mu_{4k'\alpha} + \rho_{4|3}Y_{i3}$$

  - μ:(k,α) specific heterogeneity, ρ: lasting effect of earning

$$E[Y_{i2}|\alpha,k,k'] = \mu_{2k\alpha} + \xi_2(k')$$
$$E[Y_{i3}|\alpha,k,k'] = \mu_{3k'\alpha} + \xi_3(k)$$

  - ξ: effect of future and previous job on α worker
- All similar to static, but this conditional expectation change

Recover distribution: dynamic

Similarly, we maximize this log-likelihood:

$$\sum_{i=1}^{N_m} \sum_{k=1}^{K} \sum_{k'=1}^{K} \mathbf{1}\{\hat{k}_{i2} = k\}\mathbf{1}\{\hat{k}_{i3} = k'\} \times \ldots \times$$

$$ln(\sum_{\alpha=1}^{L} p_{kk'}(\alpha;\theta_p) \underbrace{f^f_{Y_{i2},k\alpha}(Y_{i1};\hat{\rho}_{1|2},\theta_{f^f})}_{\text{dist. w. forward effect}} \underbrace{f^m_{kk'\alpha}(Y_{i2},Y_{i3};\theta_{f^m})}_{\text{job-mover}}$$

$$\underbrace{f^b_{Y_{i3},k',\alpha}(Y_{i4};\hat{\rho}_{4|3},\theta_{f^b}))}_{\text{dist. w. backward effect}} \quad (15)$$

- The likelihood of job-mover's pattern throughout 4 periods

Recover distribution: dynamic

And after getting $(\hat{\rho}, \hat{\theta}_{f^b}, \hat{\theta}_{f^s})$, we have non-mover's likelihood:

$$\sum_{i=1}^{N}\sum_{k=1}^{K}\mathbf{1}\{\hat{k}_{i2}=k\}\times ln(\sum_{\alpha=1}^{L} q_k(\alpha;\theta_q) f^f_{Y_{i2k\alpha}}(Y_{i1};\hat{\rho}_{1|2},\hat{\theta}_{f^f}) f^s_{k\alpha}(Y_{i2},Y_{i3};\theta_{f^s}) f^b_{Y_{i3},k',\alpha}(Y_{i4};\hat{\rho}_{4|3},\hat{\theta}_{f^b}) \tag{16}$$

- Parameter vector $(\hat{\theta}_p, \hat{\theta}_{f^f}, \hat{\theta}_{f^m}, \hat{\theta}_{f^b}, \hat{\theta}_q, \hat{\theta}_{f^s}, \hat{\rho})$ characterizes the sorting, complementarity, dynamic pattern
- $\hat{\xi}$ is also about dynamic, can be estimated similarly in $f^m_{kk'\alpha}(Y_{i2}, Y_{i3}; \xi, \theta_{f^m})$

Recover distribution: dynamic

- An effect distribution based on the estimation

$$\underbrace{Var(E(Y_{i3}|k_{i2}))}_{\text{total}} = Var(E[E(Y_{i3}|k_{i3}, k_{i2})|k_{i2}])$$

$$= \underbrace{Var(E[E(Y_{i3}|k_{i3})|k_{i2}])}_{\text{network effect}}$$

$$+ \underbrace{Var(E(Y_{i3}|k_{i2})) - Var(E[E(Y_{i3}|k_{i3})|k_{i2}])}_{\text{state dependence effect}}$$

where conditional expectations are from estimation

  - Network effect: from the link between current and previous job
  - State dependent effect: from the current job

5. EMPIRICAL RESULTS I: STATIC MODEL

Data and Sample

- Data
  - The authors use Sweden administrative data covering the entire working age population in Sweden between 1997 and 2008.
- Sample selection
  - males working in the private sector
  - Static model: 2002 and 2004
    - keep workers who are both fully employed in the same firm in 2002 and fully employed in the same firm in 2004, and firms with at least one fully-employed worker during the period.
  - Dynamic model: 2001–2002 and 2004–2005,
    - fully-employed in the same firm in 2001–2002 and 2004–2005.

Firm Classes

- estimate firm classes using a weighted k-means algorithm with 10,000 randomly generated starting values.
- use firms’ cdfs of 2002 log-earnings on a grid of 20 percentiles of the overall log-earnings distribution. --> identify latent heterogeneity
- weight measurements by firm size.

Firm Classes

TABLE I

DATA DESCRIPTION, BY ESTIMATED FIRM CLASSES[a]

10 classes of firms

Class:	1	2	3	4	5	6	7	8	9	10	All
Number of Workers	16,868	50,906	74,073	76,616	80,562	66,120	105,485	61,272	47,164	20,709	599,775
Number of Firms	5808	6832	4983	5835	3507	4149	3672	3467	2886	2687	43,826
Mean Firm Reported Size	12.43	20.92	42.68	28.47	65.06	32.30	60.08	51.24	54.16	50.86	37.59
Number of Firms ≥ 10 (Actual Size)	160	1034	1519	1357	1192	930	999	855	632	415	9093
Number of Firms ≥ 50 (Actual Size)	7	87	260	225	270	162	245	183	147	52	1638
% High School Drop Out	28.5%	27.8%	25.9%	26.8%	22.2%	23.8%	18.9%	12.9%	6.1%	3.2%	20.6%
% High School Graduates	61.3%	63.4%	62.3%	63.3%	59.1%	62.7%	58.4%	49.3%	34.9%	25.6%	56.7%
% Some College	10.2%	8.8%	11.8%	9.9%	18.7%	13.5%	22.8%	37.8%	59.0%	71.2%	22.7%
% Workers Younger Than 30	24.3%	19.5%	19.8%	17.5%	18.6%	15.4%	13.8%	14.3%	15.0%	14.3%	16.8%
% Workers Between 31 and 50	54.1%	54.6%	55.0%	56.2%	56.0%	57.6%	58.5%	58.9%	60.0%	64.2%	57.2%
% Workers Older Than 51	21.7%	25.9%	25.1%	26.3%	25.5%	27.0%	27.6%	26.8%	25.0%	21.5%	26.0%
% Workers in Manufacturing	24.3%	39.3%	46.8%	53.0%	51.5%	52.0%	53.0%	40.3%	31.5%	7.6%	45.4%
% Workers in Services	39.3%	32.1%	23.3%	19.7%	14.4%	15.0%	16.0%	29.7%	52.1%	72.6%	25.3%
% Workers in Retail and Trade	26.4%	19.0%	24.9%	10.6%	29.3%	7.9%	8.4%	17.7%	14.8%	18.7%	16.7%
% Workers in Construction	9.9%	9.6%	5.1%	16.8%	4.9%	25.1%	22.5%	12.3%	1.5%	1.1%	12.6%
Mean log-Earnings	9.69	9.92	10.01	10.06	10.15	10.16	10.24	10.36	10.50	10.77	10.18
Variance of log-Earnings	0.101	0.054	0.085	0.051	0.102	0.051	0.077	0.096	0.109	0.173	0.124
Skewness of log-Earnings	-1.392	-0.709	0.345	0.019	0.576	0.433	0.474	0.703	0.385	1.001	0.582
Kurtosis of log-Earnings	7.780	14.093	9.017	15.565	7.788	14.763	10.033	8.141	6.651	6.984	7.400
Between-Firm Variance of log-Earnings	0.0462	0.0044	0.0036	0.0018	0.0032	0.0016	0.0016	0.0045	0.0057	0.0435	0.0475
Mean log-Value-Added per Worker	12.40	12.58	12.69	12.69	12.84	12.75	12.87	12.94	13.03	13.18	12.74

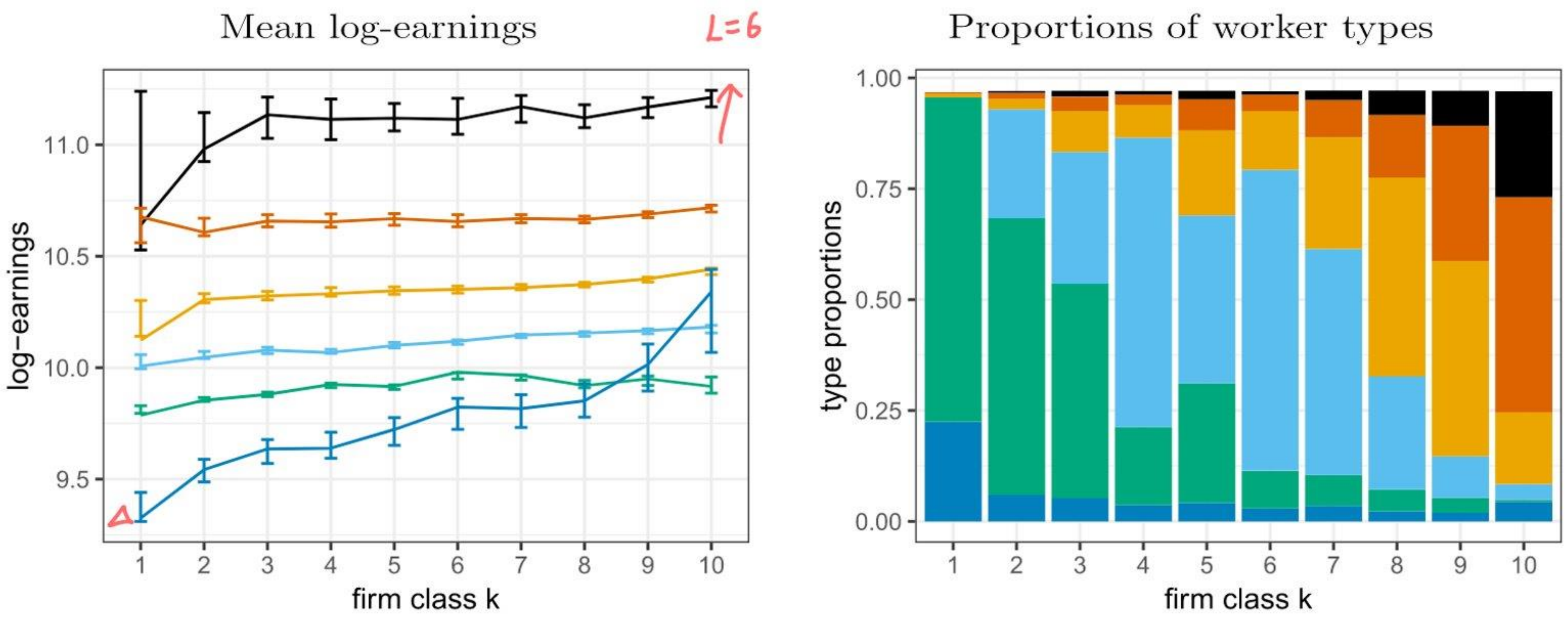
[a]The table corresponds to males fully employed in the same firm in 2002 and 2004, for firms that are continuously present in the sample. The "actual size" is the number of workers per firm in our sample. All numbers in the table correspond to 2002.

The firm classes are ordered according to mean log-earnings in each class.

The between-firm-class log-earnings variance is 0.0421, that is, 89% of the overall between-firm variance.

Wages, Worker Heterogeneity, and Firm Heterogeneity

FIGURE 2.—Main parameter estimates of the static model.



θ_f

$$\sum_{i=1}^{N}\sum_{k=1}^{K}\mathbb{1}\{\widehat{k}_{i1}=k\}\ln\left(\sum_{\alpha=1}^{L}q_k(\alpha;\theta_q)f_{k\alpha}(Y_{i1};\widehat{\theta}_f)\right).$$

θ_q

- **Results of Left panel**
- The results show clear evidence of worker heterogeneity.
- They also show some variation between firm classes, although to a lesser extent.
- Moreover, lower-type workers appear to gain the most from working in a higher-wage firm. This suggests the presence of some complementarity between firms and lower-type workers.

- **Results of Right panel**
- It shows strong sorting between worker type and firm classes.

- Overall, the two graphs in Figure 2 suggest that variation in log-earnings between firm classes is mainly due to firms employing different workers. --> complementarity and sorting.

Variance Decomposition and Reallocation

TABLE II

VARIANCE DECOMPOSITION AND REALLOCATION EXERCISE IN THE STATIC MODEL[a]

worker effect | firm effect | residual

Variance Decomposition ($\times 100$)					
$\frac{\mathrm{Var}(\alpha)}{\mathrm{Var}(y)}$		$\frac{\mathrm{Var}(\psi)}{\mathrm{Var}(y)}$	$\frac{2\,\mathrm{Cov}(\alpha,\psi)}{\mathrm{Var}(y)}$ interaction	$\frac{\mathrm{Var}(\varepsilon)}{\mathrm{Var}(y)}$	$\mathrm{Corr}(\alpha,\psi)$
60.03	>	2.56	12.17	25.24	49.13
(0.85)		(0.16)	(0.39)	(0.59)	(0.86)

Reallocation Exercise ($\times 100$)				
Mean	Median	10%-Quantile	90%-Quantile	Variance
0.50	0.58	2.60	-1.24	-1.12
(0.10)	(0.11)	(0.19)	(0.31)	(0.11)

[a] Estimates of the static model, on 2002–2004. In the top panel, α denotes the worker effect, and ψ denotes the firm effect, in the linear regression $Y = \alpha + \psi + \varepsilon$. In the bottom panel, we report differences in means, quantiles, and variances of log-earnings between two samples: a counterfactual sample where workers are randomly reallocated to firms, and the original sample. The results

$$\mathrm{var}(\alpha+\phi)=\mathrm{var}(\alpha)+\mathrm{var}(\phi)+2\mathrm{cov}(\alpha,\phi)$$

- **Estimate from Regression model**
- **Results of Variance decomposition**

1. First, worker heterogeneity explains substantially more variation in earnings than firm heterogeneity.
2. The part explained by the covariance is substantial.
3. The correlation between worker and firm effects is 49%, which suggests the presence of strong sorting between workers and firms.

- **Results of Variance decomposition**

1. The positive mean impact (0.5%, small but statistically significant) shows the presence of complementarities.
2. differences in medians and 10% and 90% percentiles of log-earnings: the bottom of the distribution would tend to benefit in the random allocation, whereas the top would be hurt.
3. the reduction in variance: log-earnings are less dispersed in the random allocation

Robustness Analysis

TABLE III

VARIANCE DECOMPOSITION (×100), STATIC MODEL, ROBUSTNESS CHECKS[a]

	$\frac{\mathrm{Var}(\alpha)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\psi)}{\mathrm{Var}(y)}$	$\frac{2\,\mathrm{Cov}(\alpha,\psi)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\varepsilon)}{\mathrm{Var}(y)}$	$\mathrm{Corr}(\alpha,\psi)$
A. Baseline Model					
	60.0	2.6	12.2	25.2	49.1
B. Varying the Number of Firm Classes					
$K = 3$	65.9	1.4	8.7	24.1	45.3
$K = 20$	57.6	2.9	12.2	27.3	47.4
C. Varying the Number of Worker Types					
$L = 3$	13.5	13.4	12.2	60.8	45.4
$L = 5$	55.5	3.2	12.8	28.5	47.9
$L = 9$	53.3	3.5	13.1	30.1	48.2
D. Other Mixture Specifications					
Mixture-of-Mixtures	62.6	2.4	11.3	23.8	46.4
Log-Earnings Residuals	49.6	3.5	10.6	36.3	40.0
Firms With $\leq$50 Workers	52.6	4.0	17.0	26.4	59.1
Firms With > 50 Workers	60.1	2.6	7.9	29.4	31.6
Fully Nonstationary	60.5	2.8	12.5	24.3	48.3
E. Regression Models					
Interactive	56.5	2.1	10.8	30.6	50.2
Linear	60.6	1.7	10.0	27.6	48.5
F. Other Classifications					
Classify Using Means	57.1	4.2	13.7	25.0	43.8
Split by Percent of Movers	67.2	1.5	8.5	22.9	42.5
Split by Value Added	55.4	3.4	12.5	28.7	46.0
G. Reclassification (Starting at Baseline)					
1 Iteration	55.9	4.1	13.3	26.7	43.9
10 Iterations	56.3	4.1	12.5	27.2	41.2

[a]Estimates of the static model, on 2002–2004. α denotes the worker effect, and ψ denotes the firm effect, in the linear regression $Y = \alpha + \psi + \varepsilon$. The results are obtained using 1,000,000 simulations. The various specifications are described in the text.

Overall, these estimates confirm the stability of the variance decomposition results.

6. EMPIRICAL RESULTS II:DYNAMIC MODEL

Parameter Estimates of dynamic model

TABLE IV

PARAMETER ESTIMATES OF THE DYNAMIC MODEL[a]

Earnings Effects $\xi_2(k')$ of Future Firm Classes								
$k'=2$	$k'=3$	$k'=4$	$k'=5$	$k'=6$	$k'=7$	$k'=8$	$k'=9$	$k'=10$
−0.005 (0.008)	0.004 (0.009)	0.005 (0.011)	0.022 (0.012)	0.002 (0.011)	0.015 (0.010)	0.009 (0.011)	0.016 (0.012)	0.023 (0.012)
Earnings Effects $\xi_3(k)$ of Past Firm Classes								
$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$
0.051 (0.016)	0.038 (0.015)	0.045 (0.015)	0.061 (0.016)	0.040 (0.017)	0.072 (0.015)	0.058 (0.018)	0.087 (0.016)	0.090 (0.017)

In addition, we specify the mean of (Y_{i2}, Y_{i3}) for job movers between classes k and k' as $(\mu_{2k\alpha} + \xi_2(k'), \mu_{3k'\alpha} + \xi_3(k))$. The term $\xi_2(k')$ reflects that, conditional on moving between k and k', mean log-earnings before the move can differ with the firm of destination, due to the presence of *endogenous mobility*. The term $\xi_3(k)$ reflects that the previous firm is allowed to have a direct effect on log-earnings after a move, through the presence of *state dependence*. Neither of those effects is allowed for in the static version of the model.

- ξ2(k’) effect is quantitatively small.
- ξ3(k) effect appears empirically quite large in the dynamic model.

Parameter Estimates of dynamic model

Persistence Parameters ρ			
$\rho_{1\|2}$	$\rho^{m}_{3\|2}$	$\rho^{s}_{3\|2}$	$\rho_{4\|3}$
0.227	0.246	0.681	0.651
(0.009)	(0.044)	(0.022)	(0.004)

$$\sum_{i=1}^{N_m}\sum_{k=1}^{K}\sum_{k'=1}^{K}\mathbf{1}\{\widehat{k}_{i2}=k\}\mathbf{1}\{\widehat{k}_{i3}=k'\}\times\ldots$$

$$\ln\left(\sum_{\alpha=1}^{L}p_{kk'}(\alpha;\theta_p)f^{f}_{Y_{i2},k\alpha}(Y_{i1};\widehat{\rho}_{1|2},\theta_{ff})f^{m}_{kk'\alpha}(Y_{i2},Y_{i3};\theta_{f^m})f^{b}_{Y_{i3},k'\alpha}(Y_{i4};\widehat{\rho}_{4|3},\theta_{f^b})\right), \quad (15)$$

and:

$$\sum_{i=1}^{N}\sum_{k=1}^{K}\mathbf{1}\{\widehat{k}_{i2}=k\}\ln\left(\sum_{\alpha=1}^{L}q_k(\alpha;\theta_q)f^{f}_{Y_{i2}k\alpha}(Y_{i1};\widehat{\rho}_{1|2},\widehat{\theta}_{ff})f^{s}_{k\alpha}(Y_{i2},Y_{i3};\theta_{f^s})f^{b}_{Y_{i3},k'\alpha}(Y_{i4};\widehat{\rho}_{4|3},\widehat{\theta}_{f^b})\right).(16)$$

Persistence estimates are higher for job stayers than for job movers.

Variance Decomposition and Reallocation

Compare static model with dynamic model

TABLE II

VARIANCE DECOMPOSITION AND REALLOCATION EXERCISE IN THE STATIC MODEL[a]

Variance Decomposition (×100)				
$\frac{\mathrm{Var}(\alpha)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\psi)}{\mathrm{Var}(y)}$	$\frac{2\,\mathrm{Cov}(\alpha,\psi)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\varepsilon)}{\mathrm{Var}(y)}$	$\mathrm{Corr}(\alpha,\psi)$
60.03 (0.85)	2.56 (0.16)	12.17 (0.39)	25.24 (0.59)	49.13 (0.86)
Reallocation Exercise (×100)				
Mean	Median	10%-Quantile	90%-Quantile	Variance
0.50 (0.10)	0.58 (0.11)	2.60 (0.19)	−1.24 (0.31)	−1.12 (0.11)

TABLE V

VARIANCE DECOMPOSITION AND REALLOCATION EXERCISE IN THE DYNAMIC MODEL

Variance Decomposition (×100)				
$\frac{\mathrm{Var}(\alpha)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\psi)}{\mathrm{Var}(y)}$	$\frac{2\,\mathrm{Cov}(\alpha,\psi)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\varepsilon)}{\mathrm{Var}(y)}$	$\mathrm{Corr}(\alpha,\psi)$
60.27 (1.30)	4.24 ↑ (0.65)	13.40 (0.37)	22.09 (0.69)	41.90 ↓ (2.35)
Reallocation Exercise (×100)				
Mean	Median	10%-Quantile	90%-Quantile	Variance
0.26 (0.28)	0.80 (0.19)	2.57 (0.70)	−3.24 (0.57)	−1.05 (1.03)

[a] Estimates of the dynamic model, on 2001–2005. See the notes to Table II.

The result is quite similar to the one we obtained using the static model, with some differences.

Robustness Analysis

TABLE VI

VARIANCE DECOMPOSITION ($\times 100$), DYNAMIC MODEL, ROBUSTNESS CHECKS[a]

	$\frac{\mathrm{Var}(\alpha)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\psi)}{\mathrm{Var}(y)}$	$\frac{2\,\mathrm{Cov}(\alpha,\psi)}{\mathrm{Var}(y)}$	$\frac{\mathrm{Var}(\varepsilon)}{\mathrm{Var}(y)}$	$\mathrm{Corr}(\alpha, \psi)$
			A. Baseline model		
	60.3	4.2	13.4	22.1	41.9
			B. Persistence parameters estimated in first differences		
	58.7	4.5	13.0	23.9	40.1
			C. Varying the number of firm classes		
$K = 3$	63.5	3.2	10.9	22.4	38.4
$K = 15$	58.7	4.9	13.8	22.6	40.8
			D. Varying the number of worker types		
$L = 3$	18.7	16.6	10.7	54.0	30.2
$L = 5$	57.8	5.5	13.6	23.2	38.0
$L = 9$	60.0	4.8	13.4	21.9	39.5
			E. Regression models		
Interactive	60.8	4.9	14.9	19.4	43.1
Linear	72.9	2.9	11.2	13.0	38.9
			F. Reclassification (starting at baseline)		
1 iteration	58.0	5.8	12.9	23.3	35.0
10 iterations	57.9	7.1	11.5	23.3	28.4

[a] Estimates of the dynamic model, on 2001–2005. See the notes to Table III.

6.2. Dynamic Effects

The estimates of the dynamic model challenge some assumptions commonly made in empirical work, such as **exogenous mobility** and the absence of **state dependence** in firm effects.

Endogenous Mobility

TABLE VII

TRANSITION PROBABILITIES ($\times$100), BY CONDITIONAL DECILE OF PREVIOUS EARNINGS[a]

to

from	All movers	$k' = 1$–3	$k' = 4$–7	$k' = 8$–10
	overall mobility rate		All	
$k = 1$–3	2.20 (0.03)	0.84 (0.02)	1.06 (0.03)	0.30 (0.02)
$k = 4$–7	1.86 (0.03)	0.39 (0.02)	1.03 (0.02)	0.44 (0.01)
$k = 8$–10	2.90 (0.06)	0.47 (0.06)	1.00 (0.03)	1.43 (0.02)
			First Conditional Decile of Earnings 低收入<10%	
$k = 1$–3	3.41 (0.21)	1.53 (0.09)	1.52 (0.12)	0.37 (0.05)
$k = 4$–7	3.20 (0.17)	0.77 (0.09)	1.74 (0.09)	0.69 (0.05)
$k = 8$–10	4.92 (0.31)	0.93 (0.15)	1.74 (0.14)	2.25 (0.13)
			Tenth Conditional Decile of Earnings 高收入 >90%	
$k = 1$–3	2.76 (0.20)	0.95 (0.08)	1.42 (0.13)	0.40 (0.06)
$k = 4$–7	1.82 (0.10) no effect	0.35 (0.04)	1.05 (0.06)	0.42 (0.03)
$k = 8$–10	2.03 (0.14)	0.29 (0.06)	0.71 (0.06)	1.03 (0.08)

[a]We report the probability of changing job, overall and by destination firm class k', for each origin firm class k. In the top panel, we show results for all workers. In the middle and bottom panels, we show results for workers in the first and tenth deciles of log-earnings Y_{i2}, respectively, conditional on worker type α_i and current firm class $k_{i2} = k$. The estimates are obtained using the dynamic model,

Study how the current wage affects a worker's decision to move, and which firm he/she moves to.

Evidence of endogenous mobility: workers are more likely to move when paid less.

State Dependence and Network Effects

TABLE VIII

DECOMPOSITION OF THE SHARE OF VARIANCE OF LOG-EARNINGS OF JOB MOVERS EXPLAINED BY THE FIRM CLASS OF THEIR PREVIOUS EMPLOYER ($\times 100$)[a]

	Total	Network Effect	State Dependence
Year After the Move (2004)	2.57 (0.75)	0.84 (0.14)	1.74 (0.67)
Two Years After the Move (2005)	2.04 (0.49)	0.98 (0.16)	1.05 (0.39)

[a] Estimates of the dynamic model, on 2001–2005. In the first column, we show the share of the within-worker-type variance of log-earnings of job movers explained by the firm class of their previous employer. In the second and third columns, we decompose this share of variance into two components, "network effect" and "state dependence," which we define in the text. The estimates are based on 1,000,000 simulations. We report parametric bootstrap standard errors in parentheses (computed using 200 replications).

- The network effect is driven by mobility patterns across the network of firms.
- The state dependence effect: the past firm may have a direct effect on the worker's wage after a job move.

7. CONCLUSION

- In this paper the authors propose a framework to allow for two-sided unobserved heterogeneity in matched employer employee data sets.
- The author’s application to Swedish administrative data shows that an additive model provides a good first-order approximation to the variance structure of log-earnings, while at the same time showing a strong association between worker and firm heterogeneity and a small relative contribution of firms to earnings dispersion.
- Using the dynamic model, they find that endogenous mobility, state dependence and network effects have an impact on earnings after a job move.
- Moreover, the authors show this strategy by reducing dimension of firm heterogeneity to a smaller number of classes is helpful in alleviating small-sample biases arising from low mobility rates.

Comments are welcome