The Anatomy of Sorting: Evidence from Danish Data

Rasmus Lentz[1] Suphanit Piyapromdee[2] Jean-Marc Robin[3]

[1]UW-Madison & The Dale T. Mortensen Centre, [2]UCL & IFS, [3]Sciences Po & UCL

November 11, 2019

1. Introduction

Introduction

1. What is the assignment of workers to firms with respect to **unobserved characteristics** filtered out from wages?
2. Literature finds modest firm contribution and limited sorting between worker and firm characteristics.

How is sorting realized?

- At entry? Via job-to-job mobility? What is the effect of job destruction?
- This question is largely unexplored.

Measuring sorting

- Fixed effects, Abowd, Kramarz, Margolis (1999, FR), Abowd, Creecy, Kramarz (2002, US), Woodcock (2015, AUS), Iranzo, Schivardi, Tosetti (2008, IT), de Melo (2018, Brazil), Card, Heining, Kline (2013, GER), Song, Price, Guvenen, Bloom, Wachter (2018, US)
- Discrete types, Bonhomme, Lamadon, Manresa (2017)

In this paper we will follow BLM's discrete mixture approach.

Results

Fixed effect AKM usually find big contribution of worker effect, rather small contribution of firm effect, small or negative contribution of covariance between worker and firm effects, and small residual variation.

		AKM											BLM
		FR	US1	AU	US2	IT	BR	DE		US3		DK	SW
		76-87	84-93	90-97	90-99	81-97	95-05	85-91	02-09	80-86	07-13	87-13	02-04
Residual	$\mathrm{E}\sigma^2$	15.8	9.3	5.0	9.3	15.1	7.0	7.7	5.0	20.3	14.9	39.0	25.2
Person effect	$\mathrm{V}a$	76.9	81.6	66.3	63.7	43.9	60.0	61.0	51.2	47.5	52.8	42.9	60.1
Firm effect	$\mathrm{V}b$	30.2	19.2	37.0	15.4	13.1	26.9	18.5	21.2	16.0	11.9	11.6	2.5
Cross effect	$2\mathrm{Cov}(a,b)$	-27.2	-2.0	-22.4	0.62	2.1	3.2	2.3	16.4	1.6	7.1	3.3	12.2
Match effect	$\mathrm{V}\widetilde{\mu}$				5.1			2.6	2.3				
Observed heterogeneity	$V\overline{\mu}$	6.8	52.0	3.1	4.0	7.5	3.0	10.7	2.8	7.6	7.2	1.8	
	$2\mathrm{Cov}(a,x)$	-3.1	-69.0	9.4	0.64	15.5		-2.6	0.70			0.88	
	$2\mathrm{Cov}(b,x)$	0.7	9.0	1.7	1.25	2.6		1.6	1.7			0.52	
Sorting	$\mathrm{Corr}(a,b)$	-28.3	-2.5	-22.7	1.0	4.4	4.0	3.4	24.9	2.9	14.2	7.4	49.1

Why are the contributions of firm effects and sorting so small?

2. The model

Data structure

- Workers: $i \in \{1, ..., N\}$. $\sim$ 4m workers.
- Firms: $j \in \{1, ..., J\}$. $j = 0$ denotes non-employment. $\sim$ 400k firms.
- t denotes the observation occurrence (week). 1985-2013.
- Observations (w_{it}, j_{it}, x_{it}), $t = 1, 2, \ldots, T_i$.
  - $j_{it} \in \{0, 1, ..., J\}$ is the ID of the firm employing worker i in week t.
  - x_{it} are potential experience (4 cat), and tenure (2 cat).
  - w_{it} is worker i's log-wage rate at time t.
- We also observe z_i: gender and education (3 cat.)
- Firms differ in observable characteristics, e.g. private/public sector (ζ_j).
- 5 distinct time periods: 89-93, 94-98, 99-03, 04-08, 09-13. Periods 1 and 5 are throughs.

Unobserved heterogeneity

- Firms clustered into L groups indexed by $\ell \in \{0, 1, \ldots, L\}$. Non-employment is $\ell = 0$.
  - Unobserved firm types ℓ treated as fixed effects.
  - Let $F = (\ell_1, \ldots, \ell_J)$ denote a given firm classification.
- Workers clustered into K groups indexed by $k \in \{1, ..., K\}$.
  - Unobserved worker types k treated as random effects.
- Workers and firms stay the same type for ever, but link between wages/mobility and types can change flexibly in x.

Likelihood

Lots of conditional independence:

$$L(\beta, F) = \sum_{k=1}^{K} \frac{\pi(k|z_i)\, m^0(k, \ell_{i1}|x_{i1})}{q(\ell_{i1}|F)} \times \prod_{t=1}^{T} f(w_{it}|k, \ell_{it}, x_{it}, w_{i,t-1}, \ell_{i,t-1}, x_{i,t-1})$$
$$\times \prod_{t=1}^{T-1} M(\neg|k, \ell_{it} x_{it})^{1-D_{it}} \left(\frac{M(\ell_{i,t+1}|k, \ell_{it}, x_{it})}{q(\ell_{i,t+1}|F)}\right)^{D_{it}}$$

where

- $\pi(k|z_i)$ is probability of worker type given education and gender
- $m^0(k, \ell_{i1}|x_{i1})$ is initial matching probability given current date
- f is wage distribution given match type
- M is transition probability
- $q(\ell|F) = \#\{j : \ell_j = \ell\}/J$ is the share of type-ℓ firms (so $1/q(\ell|F)$ is proportional to the probability that *this* particular firm j_{it} be drawn)
- D_{it} indicates a mobility (EE, EU or UE)

Empirical specification – Wages

- $f_{k\ell}(w|x)$ denotes the wage density, conditional on worker type k and employer type ℓ.
- We assume

$$w_{it} = \mu\left(k, \ell_{it}, x_{it}\right) + v_{it}$$
$$v_{it} = \rho v_{i,t-1} + \epsilon_{it}, \quad \epsilon_{it} \sim N\left(0, \sigma^2\left(k, \ell_{it}, x_{it}\right)\right)$$

- x includes potential experience (4 cat), and tenure (2 cat).

Empirical specification – Transition probabilities

- Probability that type k worker moves from ℓ to ℓ' firm type,
$$M(\ell'|k,\ell,x) = \lambda_k(x)\nu_{\ell'}(x)P_{k\ell\ell'}(x)$$
  - λ_k: worker search intensity
  - $\nu_{\ell'}$: sampling rate of firm types (sums to one)
  - $P_{k\ell\ell'}$: Probability that ℓ to ℓ' transition is executed.
- We assume a Bradley-Terry specification, with $P_{k00} = 0$ and
$$P_{k\ell\ell'}(x) = \frac{\gamma_{k\ell'}(x)}{\gamma_{k\ell}(x) + \gamma_{k\ell'}(x)}, \quad \sum_{\ell=1}^{L} \gamma_{k\ell} = 1,$$
where $\gamma_{k\ell}$ measures the quality of the match (k,ℓ).
- $M(\neg|k,\ell,x) = 1 - \sum_{\ell'=0}^{L} M(\ell'|k,\ell,x)$ is probability of staying with same employer.
- u-e and e-u transition probabilities completely unrestricted:
$$M(\ell'|k,0,x) = \psi_{k\ell'}(x), \quad M(0|k,\ell,x) = \delta_{k\ell}(x).$$

Classification EM algorithm

- For given K, L iterate the following two steps:
  1. For a given firm classification $F^{(s)} = (\ell_1^{(s)}, ..., \ell_J^{(s)})$, find $\beta^{(s)}$ the ML estimate of $\beta = (\underbrace{\mu, \sigma, \rho}_{\text{wages}}, \underbrace{\lambda, \nu, \gamma, \delta, \psi}_{\text{mobility}}, \underbrace{\pi, m^0}_{\text{initial}})$ using the EM algorithm.
     - We developed a special MM algorithm to estimate the nonlinear mobility model inside each M-step.
     - In practice we start we the unconstrained BLM model until it converges, then we switch to the constrained model.
     - This allows to test the restriction using a LR test. We do not reject the constraints.
  2. Update firm classification by maximizing the expected log likelihood given observations and current values of $\beta^{(s)}$ (like estimating F in the M-step)
- Repeat with different starting values (30) and keep the global maximum. (We verified that the best five local MLE give similar estimates.)
- Repeat the whole procedure for different K and L and find the “elbow” to estimate K, L.

3. Worker and firm types

- Composition
- Selection

Number of groups and labeling

- We select $K = 24$ worker types, $L = 12$ firm types.
- Linear projection,

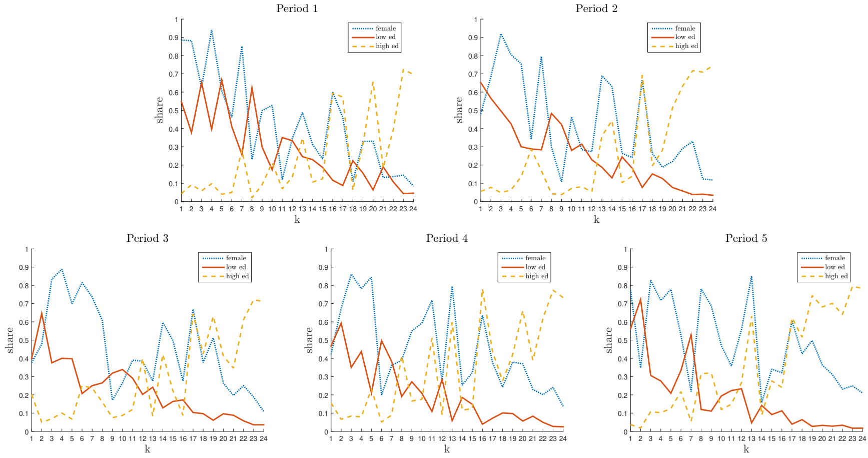
$$\mu_{k\ell}(x) = \overline{\mu}(x) + a_k + b_\ell + \widetilde{\mu}_{k\ell}(x),$$

  where $\overline{\mu}(x)$ contains tenure*experience interactions.
- Order worker types by a_k.
- Order firm types by b_ℓ.

3.1. Composition

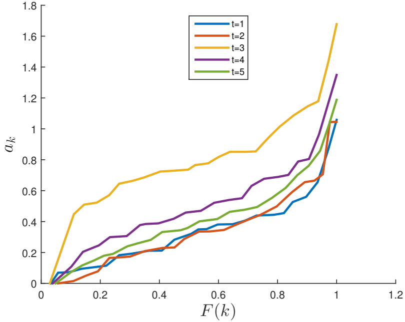
Worker k types by education and gender

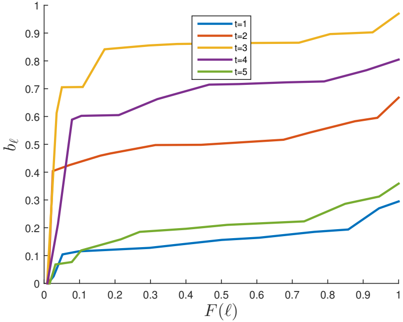
Different ways of grouping unobserved ability, education and gender
Still, fewer low ed and females in larger ks



Worker and firm fixed effects over time

Flat firm fixed effects for larger firms (for $\ell > 4$)
Less dispersion of worker and firm fixed effects in downturns





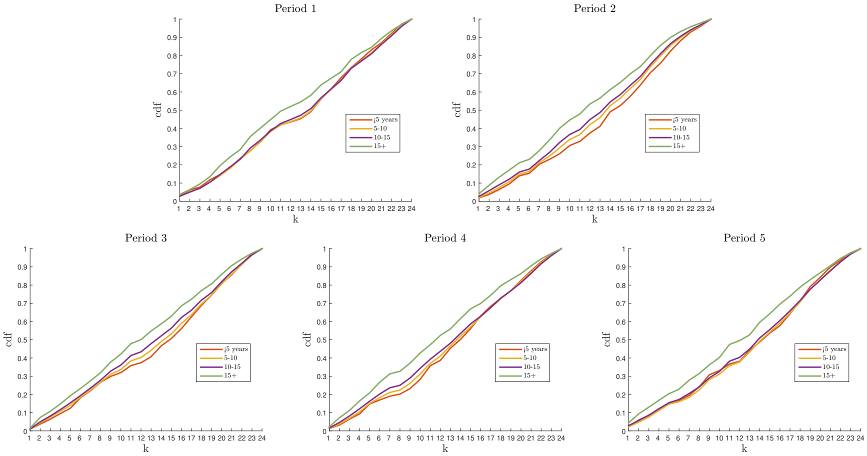
Firm types (period 1)

ℓ	no info	public	private	no. firms	avg size/yr	avg inflow/yr	avg outflow/yr	sd size/yr	sd inflow/yr	sd outflow/yr
1	0.02	0.05	0.93	16688	2.54	0.36	0.44	2.41	0.73	0.90
2	0.02	0.09	0.90	30079	1.81	0.15	0.13	0.88	0.31	0.29
3	0.02	0.04	0.94	18293	3.79	0.66	0.67	2.73	0.72	0.78
4	0.02	0.10	0.88	28569	4.00	0.59	0.59	2.43	0.67	0.70
5	0.13	0.70	0.17	23	14184.45	1630.90	1575.91	11349.18	1030.81	1100.91
6	0.02	0.10	0.88	19773	10.15	1.81	1.82	5.72	1.59	1.73
7	0.09	0.55	0.36	240	781.79	106.20	104.68	392.25	52.75	55.96
8	0.27	0.33	0.40	63	2982.31	344.66	370.17	1317.84	173.84	177.40
9	0.02	0.10	0.88	9269	32.99	5.52	5.46	22.61	4.60	5.00
10	0.04	0.20	0.76	814	207.66	30.00	29.12	93.96	17.24	18.73
11	0.04	0.08	0.88	408	530.42	67.68	69.98	594.69	71.26	61.72
12	0.03	0.05	0.92	3037	46.79	7.23	8.34	41.86	7.44	10.27

3.2. Selection

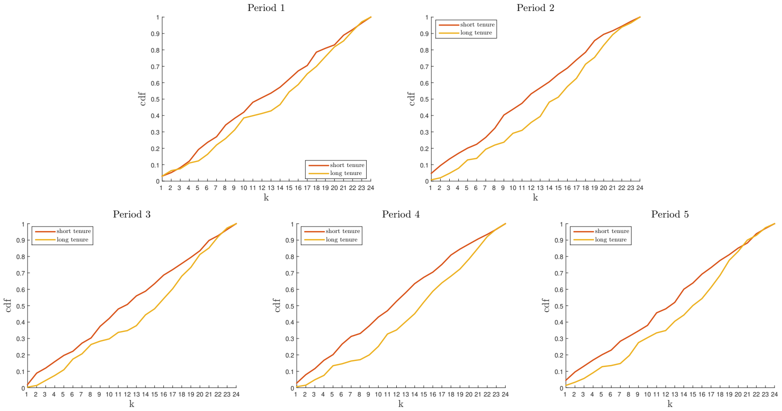
CDF of worker types by age

Older workers tend to be less able
Clearer age effect during the booms



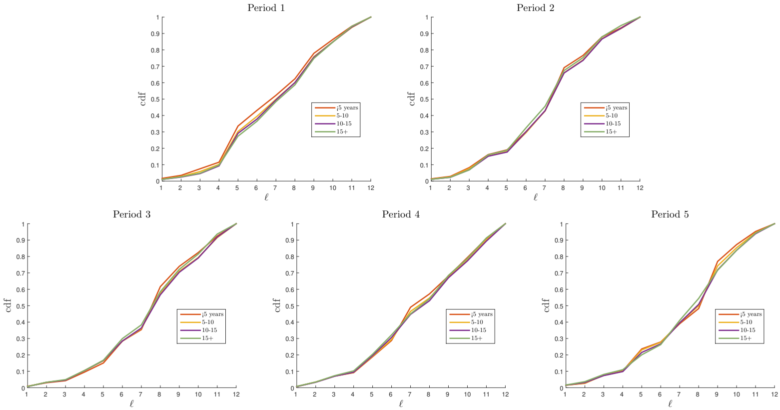
CDF of worker types by tenure

Short tenure gathers more low type workers (indirect effect of job destruction)



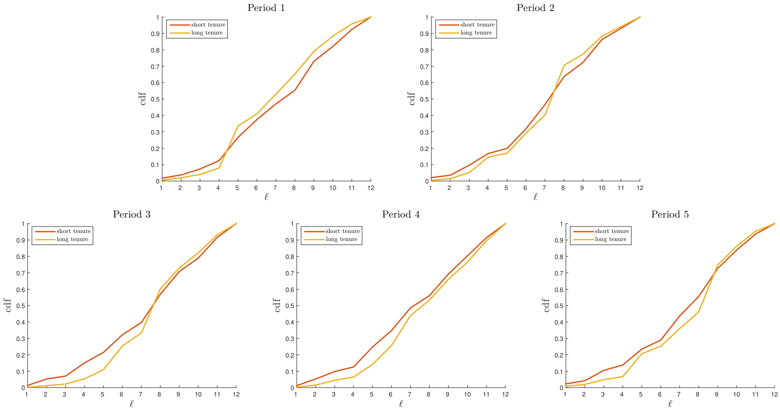
CDF of firm types by age

Nothing!



CDF of firm types by tenure

Short tenure workers more often on low-type firms (particularly during booms).
⇒ **More ladder effect at short tenure**

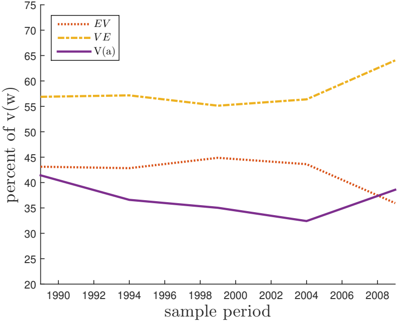


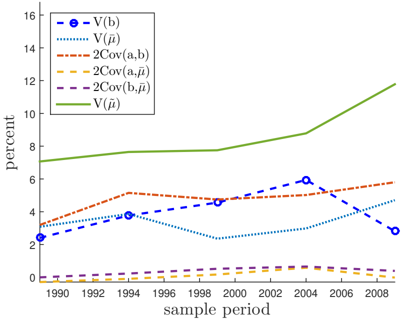
4. Variance decomposition

Variance decomposition

$$\mathrm{Var}(w) = \underbrace{\mathrm{Var}[\mu_{k\ell}(x)]}_{\text{between}} + \underbrace{\mathbb{E}(v)}_{\text{idiosyncratic}}$$

$$\mathrm{Var}[\mu_{k\ell}(x)] = \mathrm{Var}[\overline{\mu}(x)] + \mathrm{Var}[a_k] + \mathrm{Var}[b_\ell] + 2\mathrm{Cov}[a_k, b_\ell] + 2\mathrm{Cov}[\overline{\mu}(x), a_k] + 2\mathrm{Cov}[\overline{\mu}(x), b_\ell] + \mathrm{Var}[\widetilde{\mu}_{k\ell}(x)]$$





Comparison with AKM

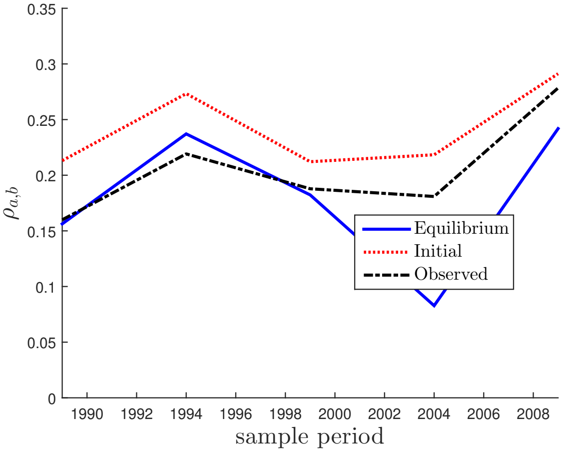
Kind of similar (after adding firm effect and match effect)

		AKM	LPR				
		87-13	89-93	94-98	99-03	04-08	09-13
Residual	$\mathrm{E}\sigma^2$	39.0	43.1	42.8	44.9	43.6	35.9
Person effect	$\mathrm{V}a$	42.9	41.4	36.6	35.0	33.3	38.6
Firm effect	$\mathrm{V}b$	11.6	2.4	3.8	4.6	5.5	2.8
Cross effect	$2\mathrm{Cov}(a,b)$	3.3	3.2	5.2	4.7	5.0	5.8
Match effect	$\mathrm{V}\widetilde{\mu}$		7.1	7.6	7.8	8.2	11.8
Observed heterogeneity	$V\overline{\mu}$	1.8	3.1	3.9	2.4	2.9	4.7
	$2\mathrm{Cov}(a,x)$	0.88	-0.29	-0.09	0.17	0.62	-0.02
	$2\mathrm{Cov}(b,x)$	0.52	0.00	0.23	0.52	0.69	0.39
Sorting	$\mathrm{Corr}(a,b)$	7.4	16.0	21.9	18.8	18.5	27.8

5. Anatomy of sorting

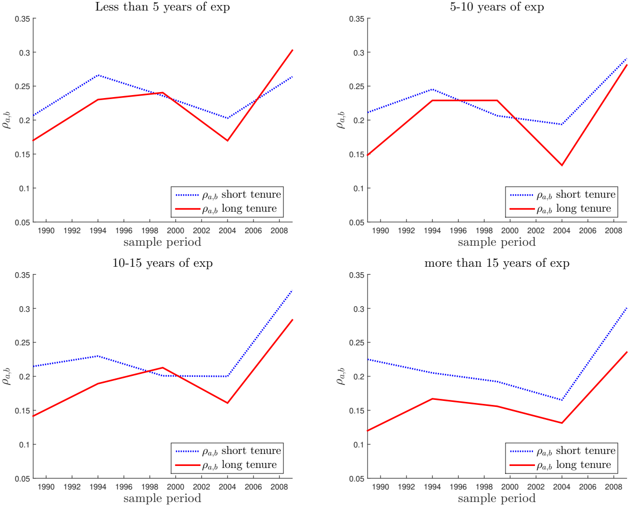
Observed, initial and equilibrium sorting

More sorting is obtained initially



Correlation between a_k and b_ℓ by x

Young and short tenure workers are more sorted.



Correlation between $\sigma_{k\ell}$ and $\mu_{k\ell}$

Short tenure: good matches have smaller residual variance in booms.
Long tenure: positive association in troughs

$\sigma_{k\ell}$	period				
	1	2	3	4	5
experience	tenure < 100 weeks				
<5 years	-0.13	-0.47	-0.59	-0.56	-0.21
5-10 years	-0.03	-0.40	-0.56	-0.52	-0.09
10-15 years	0.02	-0.34	-0.51	-0.47	0.00
> 15 years	0.02	-0.37	-0.56	-0.50	0.02
	tenure > 100 weeks				
<5 years	0.07	-0.07	-0.03	0.13	0.33
5-10 years	0.16	-0.02	-0.03	0.10	0.35
10-15 years	0.17	0.04	-0.01	0.13	0.35
> 15 years	0.13	0.06	0.00	0.09	0.32

Correlation between $\delta_{k\ell}$ and $\mu_{k\ell}$

Short tenure: greater unemployment risk for bad matches, particularly during troughs

$\delta_{k\ell}$	period				
	1	2	3	4	5
experience	tenure < 100 weeks				
<5 years	-0.35	-0.75	-0.72	-0.84	-0.46
5-10 years	-0.42	-0.69	-0.77	-0.78	-0.52
10-15 years	-0.38	-0.63	-0.70	-0.72	-0.44
> 15 years	-0.33	-0.62	-0.70	-0.68	-0.36
	tenure > 100 weeks				
<5 years	-0.02	-0.13	-0.15	-0.08	0.11
5-10 years	0.05	-0.19	-0.18	-0.08	0.08
10-15 years	0.02	-0.13	-0.17	-0.16	0.07
> 15 years	0.05	-0.11	-0.10	-0.12	0.07

Correlation between $\psi_{k\ell}$ and $\mu_{k\ell}$

Age effect: younger workers exit unemployment to good matches more often
No business cycle effects

$\psi_{k\ell}$	period				
	1	2	3	4	5
experience	duration < 26 weeks				
<5 years	0.24	0.27	0.23	0.27	0.32
5-10 years	0.24	0.22	0.22	0.27	0.20
10-15 years	0.16	0.18	0.20	0.25	0.16
> 15 years	0.16	0.17	0.19	0.22	0.17
	duration > 26 weeks				
<5 years	0.05	0.02	-0.10	-0.11	0.03
5-10 years	-0.08	-0.10	-0.21	-0.18	-0.11
10-15 years	-0.19	-0.18	-0.24	-0.14	-0.10
> 15 years	-0.23	-0.26	-0.22	-0.21	-0.05

Correlation between $\gamma_{k\ell}$ and $\mu_{k\ell}$

Short tenure and booms: job mobility better aligned with wage ranking
Long tenure: correlation becomes negative!

$\gamma_{k\ell}$	period				
	1	2	3	4	5
experience	tenure < 100 weeks				
<5 years	0.06	0.19	0.21	0.29	0.05
5-10 years	-0.06	0.20	0.19	0.33	0.02
10-15 years	-0.09	0.14	0.19	0.25	0.00
> 15 years	-0.13	0.12	0.16	0.26	0.00
	tenure > 100 weeks				
<5 years	-0.13	-0.12	-0.21	0.05	-0.24
5-10 years	-0.19	-0.12	-0.18	0.06	-0.35
10-15 years	-0.18	-0.21	-0.20	-0.02	-0.25
> 15 years	-0.20	-0.22	-0.08	0.01	-0.23

6. Synthetic cohorts

- Benchmark
- Counterfactuals

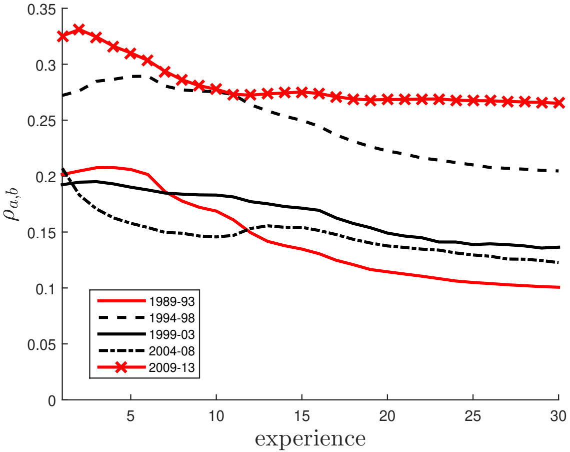
Synthetic cohorts

- Create cohort starting with zero experience and zero tenure for a given time period.
- Initialize by the estimated $m^0(k, \ell)$. Simulate forward 30 years, holding calendar time fixed.
- Illustrates estimated sorting model interaction with tenure and experience.

6.1. Benchmark

Synthetic cohorts - $\rho_{a,b}$ by cohort experience

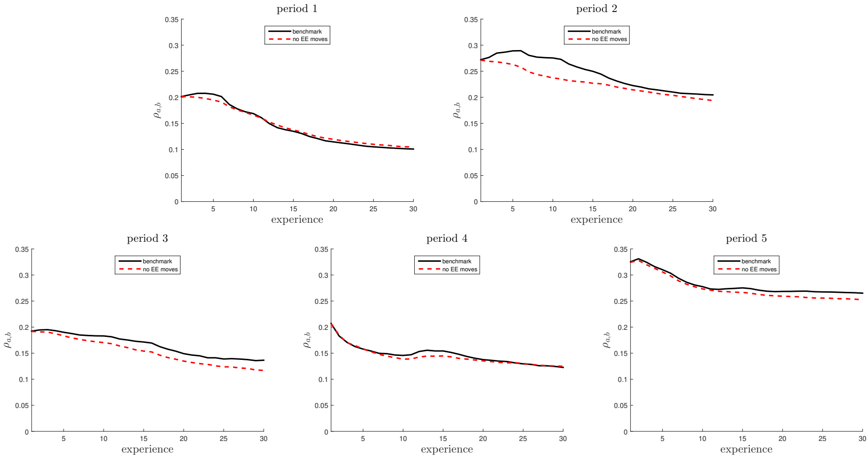
Sorting declines over the life-cycle



6.2. Counterfactuals

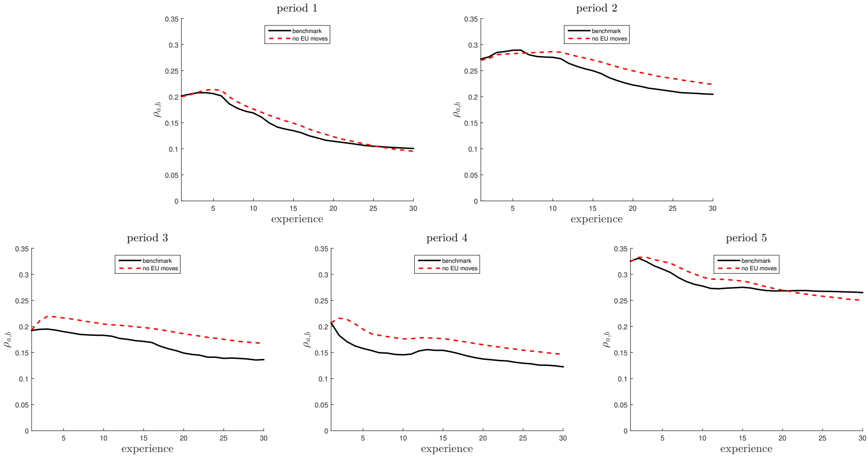
Correlation - no EE moves ($\lambda_k = 0$)

EE moves have (limited) positive impact on sorting over life
Not surprising, given low variance in firm fixed effect contribution



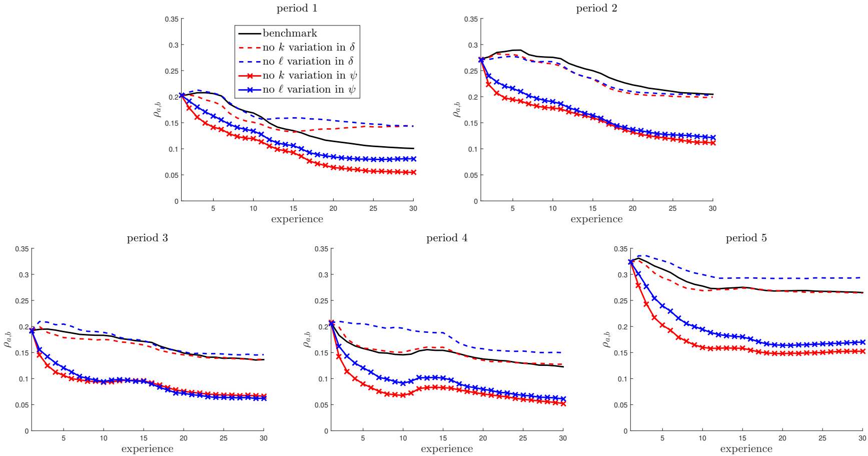
Correlation - no U shocks ($\delta_{k\ell} = 0$)

Job destruction weakens sorting particularly during the booms



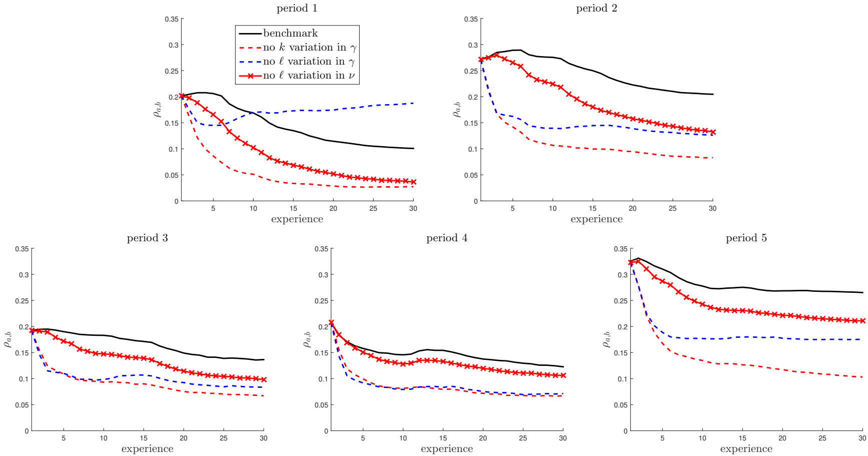
Correlation and U-E transitions

Job destruction mixes and reemployment sorts



Correlation and EE transitions

Similar effect as for job finding rates (ψ).



7. Conclusion

Concluding remarks

- Following discrete mixture approach in BLM, we present a CEM+MM algorithm for the flexible estimation of wage and mobility parameters.
  - Fast estimation of nonlinear mobility parameters from MM algorithm in M-step.
  - C-step improves performance of estimator.
- Main Findings
  - Bigger residual wage variation than is usually found
  - Sizable degree of sorting over time, essentially
  - Most sorting is obtained when young. Early career mobility strengthens sorting. Subsequent mobility tends to undo it.
  - Differences in re-employment risk, mobility out of non-employment, and preferences for firms are the key drivers to sustain sorting.
  - Main channel is that the conditional wage expectation differs by firm only when contrasting smallest firms. Workers who draw small firms tend to leave them fast.