# Sorting in the labor Market

## Part 1: AKM framework

Thibaut Lamadon

U. Chicago

October 24, 2017

# Features in the data that we want to understand:

## Failure of the law of one price: Wage dispersion

- Similar workers are paid differently
  Observable characteristics only explain $\sim 30\%$ of wage dispersion
- Some firms/industry pay permanently higher wages
  Even when controlling for worker quality

## Allocation of workers and its dynamics

- A fraction of the population is actively looking for a job
- Firms have difficulties finding the right candidates
- Worker reallocation to more productive jobs is important for productivity growth ( $\sim 24\%$) but rate is falling

# CPS

Wage equations for full-time employees by sex, 1983[a,b]

| Variable | Male employees | | | Female employees | | |
|---|---|---|---|---|---|---|
| | Mean | *B* | *t* | Mean | *B* | *t* |
| *A. Firm/plant size dummies*[c] | | | | | | |
| F2SP | 0.030 | 0.110 | 3.96 | 0.032 | 0.088 | 3.06 |
| F3SP | 0.025 | 0.092 | 3.04 | 0.027 | 0.127 | 4.06 |
| F4SP | 0.008 | 0.147 | 2.76 | 0.007 | 0.048 | 0.83 |
| F5SP | 0.051 | 0.117 | 5.17 | 0.040 | 0.131 | 4.96 |
| F2LP | 0.115 | 0.087 | 5.32 | 0.116 | 0.075 | 4.41 |
| F3LP | 0.109 | 0.142 | 8.38 | 0.124 | 0.127 | 7.50 |
| F4LP | 0.043 | 0.134 | 5.53 | 0.055 | 0.160 | 7.00 |
| F5LP | 0.353 | 0.245 | 17.90 | 0.316 | 0.232 | 17.00 |
| *B. Worker/job characteristics* | | | | | | |
| Education | 12.915 | 0.063 | 33.45 | 12.684 | 0.064 | 26.77 |
| Ten | 8.205 | 0.020 | 12.09 | 5.537 | 0.028 | 14.17 |
| Ten-2 | 145.516 | −0.040e−2 | −8.01 | 72.606 | −0.058e−2 | −8.05 |
| Exp | 18.452 | 0.025 | 16.02 | 17.772 | 0.012 | 8.22 |
| Exp-2 | 496.391 | −0.043e−2 | −13.10 | 473.881 | −0.027e−2 | −8.35 |
| Married | 0.744 | 0.122 | 10.52 | 0.629 | 0.003 | 0.30 |
| Black | 0.055 | −0.170 | −8.14 | 0.078 | −0.100 | −5.33 |
| SMSA | 0.374 | 0.122 | 11.48 | 0.390 | 0.134 | 13.16 |
| South | 0.280 | −0.048 | −4.64 | 0.292 | −0.047 | −4.29 |
| *C. Industrial affiliation* | | | | | | |
| Agriculture | 0.025 | −0.351 | −11.28 | 0.005 | −0.170 | −2.40 |
| Mining | 0.024 | 0.193 | 6.31 | 0.005 | 0.326 | 4.69 |
| Construction | 0.084 | 0.186 | 9.91 | 0.012 | 0.079 | 1.70 |
| TCU (Utilities) | 0.094 | 0.103 | 6.08 | 0.055 | 0.161 | 6.86 |
| Trade | 0.216 | −0.129 | −9.53 | 0.240 | −0.190 | −12.44 |
| Finance | 0.055 | 0.031 | 1.43 | 0.119 | −0.006 | −0.35 |
| Service | 0.162 | −0.112 | −7.49 | 0.350 | −0.026 | −1.84 |
| *Summary statistics* | | | | | | |
| ln AHE | 2.155 | | | 1.777 | | |
| R-square | 0.4064 | | | 0.3352 | | |
| N | 7833 | | | 5973 | | |

[a] Source: May 1983 CPS.

[b] Dependent variable is ln(average hourly earnings).

[c] F2–F5 correspond to firm size categories 25–99, 100–499, 500–999, 1000+; SP, LP correspond to small plants (1–24) and larger plants (25+), respectively.

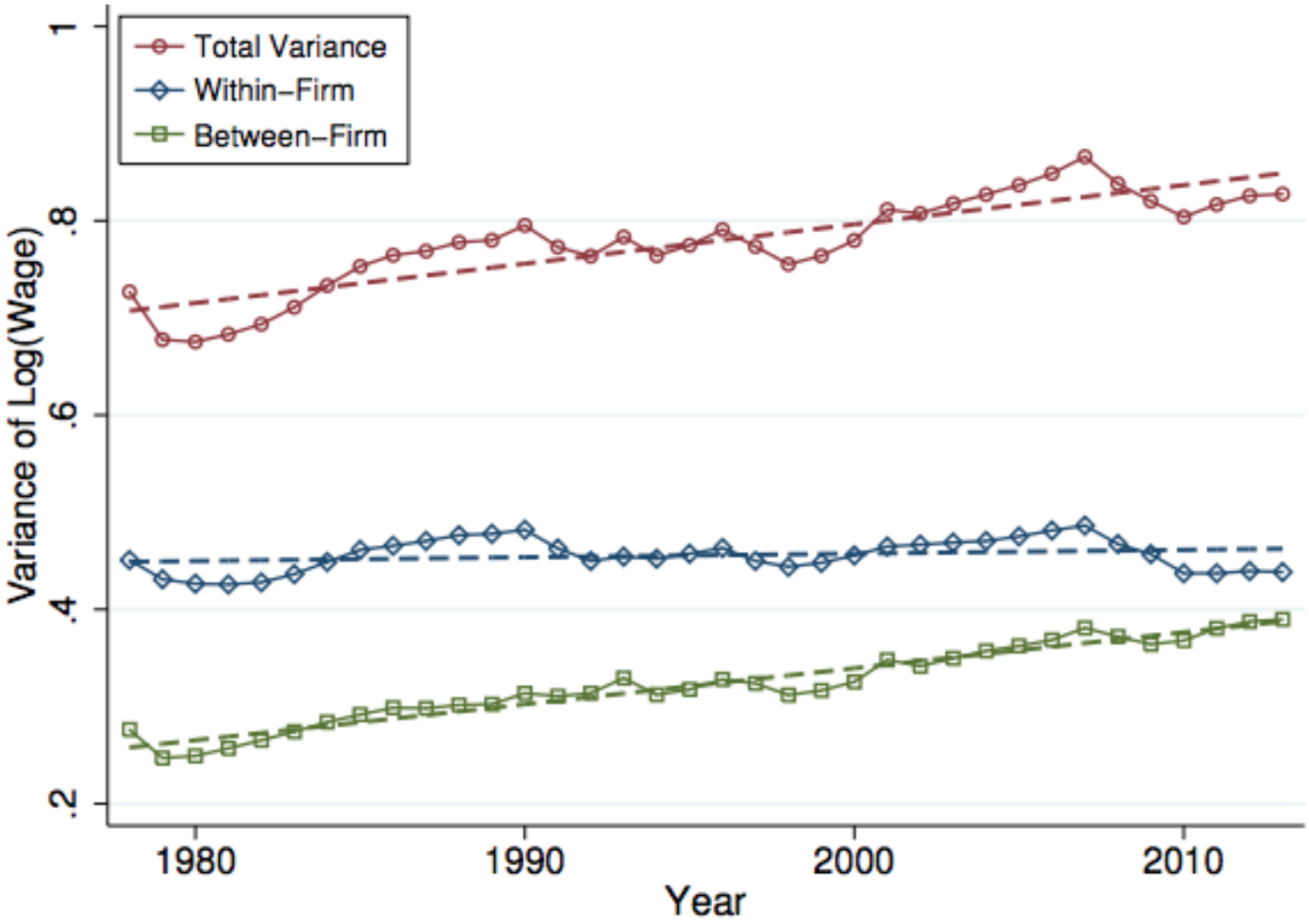
# CPS

Table 4
Wages and related variables by firm size and sex, 1993[a]

| Variable | F1 1–24 | F2 25–99 | F3 100–499 | F4 500–999 | F5 1000+ | Ratio[b] |
|---|---|---|---|---|---|---|
| Females | | | | | | |
| Sample size | 2120 | 1087 | 1081 | 442 | 3167 | |
| Wage | 8.203 | 9.052 | 10.114 | 10.525 | 10.683 | 1.302 |
| Tenure | 5.664 | 6.093 | 6.843 | 7.212 | 8.128 | 1.435 |
| Education | 12.698 | 12.807 | 13.109 | 13.239 | 13.137 | 1.035 |
| White | 91.698 | 88.960 | 88.714 | 87.330 | 85.475 | 0.932 |
| Married | 58.726 | 56.486 | 56.152 | 56.335 | 54.500 | 0.928 |
| Part-time | 39.906 | 24.103 | 21.462 | 19.231 | 23.745 | 595 |
| Union[c] | 1.063 | 4.019 | 7.034 | 11.848 | 13.583 | 12.778 |
| Pension[d] | 14.554 | 28.044 | 48.293 | 50.856 | 61.544 | 4.229 |
| Males | | | | | | |
| Sample size | 2144 | 1302 | 1189 | 451 | 3698 | |
| Wage | 10.289 | 12.381 | 13.459 | 13.528 | 14.951 | 1.452 |
| Tenure | 6.338 | 7.030 | 8.089 | 9.125 | 11.246 | 1.774 |
| Education | 12.515 | 12.786 | 13.193 | 13.181 | 13.494 | 1.078 |
| White | 90.951 | 90.860 | 91.926 | 89.135 | 88.886 | 0.977 |
| Married | 55.364 | 61.290 | 63.751 | 66.962 | 66.820 | 1.207 |
| Part-time | 18.470 | 8.372 | 7.653 | 7.539 | 9.708 | 0.536 |
| Union | 5.005 | 10.925 | 13.832 | 18.307 | 24.784 | 4.952 |
| Pension | 12.748 | 38.591 | 56.495 | 61.575 | 73.604 | 5.774 |

[a] Source: April 1993 Current Population Survey
[b] Ratio = F5/F1.
[c] Union = 1 if either a union member or covered by a union contract.
[d] Pension = 1 if covered by a pension or retirement plan.

# SSA data, firming up inequality, Bloom et Al



# The data and the econometric problem

Addressing these questions empirically is made possible by the availability of very rich micro data:

- Several countries offer access to **administrative data**
  - individual tax records (earnings, capital gains, education, ...)
  - firm tax records (wages and work force, balance sheets, ...)
  - unemployment and government benefit records
- Using this records we can construct a detailed panel:
  - track individuals earnings, participation and benefits
  - track individuals from one firm to another
  - link earnings to firm performance

The main econometric problem is to disentangle the contribution of the worker from the contribution of the firm

- only observed the outcome of workers-firms pairs
- assignment and mobility are endogenous

# The data and the econometric problem

Addressing these questions empirically is made possible by the availability of very rich micro data:

- Several countries offer access to **administrative data**
  - individual tax records (earnings, capital gains, education, ...)
  - firm tax records (wages and work force, balance sheets, ...)
  - unemployment and government benefit records
- Using this records we can construct a detailed panel:
  - track individuals earnings, participation and benefits
  - track individuals from one firm to another
  - link earnings to firm performance

The main econometric problem is to disentangle the contribution of the worker from the contribution of the firm

- only observed the outcome of workers-firms pairs
- assignment and mobility are endogenous

# Research agenda:

Develop **models** and **empirical methods** to:

1. Understand the **allocation of workers to jobs**
   - are workers and jobs assortatively matched ?
   - what is the output loss due to mismatch ?
   - can labor policies improve market efficiency ?
   - how is the allocation changing over time (more/less sorting )?
2. Understand how **wages** are set
   - what are the sources of wage inequality (worker/firm/sorting) ?
   - how are wages linked to productivity ?

# Course agenda

1. Log linear framework of wages for matched-data
   - started with Abowd, Kramarz, and Margolis (1999)
   - $y_{it} = \alpha_i + \psi_{j(i,t)} + \epsilon_{it}$
   - cover results and limitations

2. Models of sorting
   - frictionless (Becker, 1974)
   - matching with search frictions (Shimer and Smith, 2000)
   - identification of model, results on data (Hagedorn, Law, and Manovskii, 2014)

3. Distributional of wages for matched-data
   - based on Bonhomme, Lamadon, and Manresa (2015)
   - identification with and without exogenous mobility
   - estimation on the data for exogenous case
   - performance on structural models

# The log-linear fixed effect framework

# A log linear model for wages

Abowd, Kramarz, and Margolis (1999) introduces the following model:

$$y_{it} = \alpha_i + \psi_{j(i,t)} + x_{it}\beta + \epsilon_{it}$$

- $x_{it}\beta$: observables rewarded equally at all employers
  includes year dummies, age functions, education ...
- $\alpha_i$: unobservables rewarded equally at all employers
  skills ...
- $\psi_j$: pay premium for all employed at firm $j$
- $\epsilon_{it}$: residuals

- Estimates can be used to derive interesting variance decomposition as well as sorting patterns

# More precisely

- Consider the following **potential wage equation**

$$Y^*_{ijt} = X_{ijt}\beta + \alpha_i + \psi_j + \epsilon_{ijt}$$

- denote $D_{ijt} = 1$ when worker $i$ works at firm $j$ at time $t$
- stacking variables in $\tilde{A}, \tilde{P}, \tilde{X}$ we get that

$$\mathbb{E}\Big[Y^*|D, X, \alpha, \psi\Big] = \tilde{X}\beta + \tilde{A}\alpha + \tilde{P}\psi$$

- however, we do no observe $Y^*$ but only the matched in the population. Let's call $S$ the projection matrix constructed from $D$, in practice we use

$$\mathbb{E}\Big[Y|D, X, \alpha, \psi\Big] = X\beta + A\alpha + P\psi$$

where $A = S\tilde{A}$...

# More precisely

i) Exogenous mobility

$$\mathbb{E}[\epsilon | D, X, \alpha, \psi] = 0$$

- individual movement is conditional on types only
- rules out offer sampling, selection on match specific components
- this is conditional on the whole network

iii) Firms have to be in the same connected set

- this is the rank condition
- firms that are not part of the same connected set can't be compared
- the identification comes from the movers

# Direct Estimation

- the model is linear
  - construct regressors with dummy for each worker and firm
- in practice i) get firm fixed effect by looking at movers
  - $y_{it'} - y_{it} = \psi_{j(i,t')} - \psi_{j(i,t)} + \epsilon_{it'} - \epsilon_{it}$
  - solve on movers only
- ii) recover worker fixed effects by
  - $\hat{\alpha}_i = \frac{1}{n_i} \sum_t \left( y_{it} - \hat{\psi}_{j(i,t')} \right)$
  - do this for the full connected sample

# Zig-Zag Estimation

- Solving the linear system on movers can be very expensive
  - IRS data has 50 millions firms
- The least square problem is given by

$$\min \sum_i \sum_t \left( y_{it} - x_{it}\beta - \alpha_i - \psi_{j(i,t)} \right)^2$$

- Guimaraes, Portugal, et al. (2010) proposed the following:
  1. update $\beta$ given $(\alpha_i, \psi_j)$
  2. update $\alpha_i$ given $(\beta, \psi_j)$
  3. update $\psi_i$ given $(\alpha_i, \beta)$
  4. repeat
- each step is very efficient, mostly within averages
- makes estimation possible on very large sample (IRS 250M)

# Older results, collected by Rafael De Melo in JMP

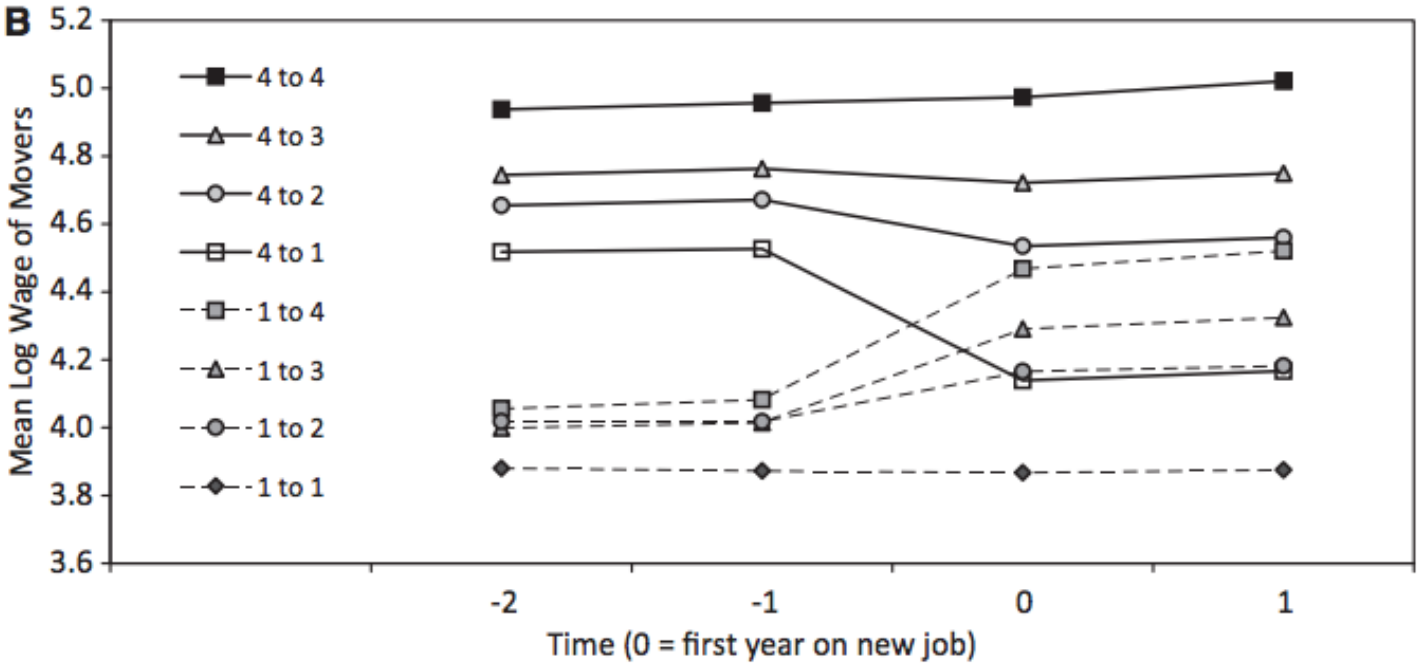
| Country | US 1[a] | US 2 | FR | GE | IT | DE[b] | BR |
|---|---|---|---|---|---|---|---|
| $Var\,(x\beta)$ | 0.03 | 0.14 | 0.02 | — | 0.01 | — | 0.02 |
| $Var\,(\theta)$ | 0.29 | 0.23 | 0.21 | 0.05 | 0.05 | 0.08 | 0.40 |
| $Var\,(\psi)$ | 0.08 | 0.053 | 0.08 | 0.013 | 0.01 | 0.00 | 0.18 |
| $\frac{Var(\psi)}{Var(\theta+\psi)}$ | 0.22 | 0.19 | 0.32 | 0.22 | 0.23 | 0.03 | 0.31 |
| $Corr\,(\theta,\psi)$ | −0.01 | −0.03 | −0.28 | −0.19 | 0.04 | 0.00 | 0.04[f] |
| $Corr\left(\theta,\widetilde{\theta}\right)$ | — | — | — | — | 0.17[c] | 0.40[d] | 0.52 |
| $R^2$ | 0.89 | 0.9 | 0.84 | — | — | 0.85 | 0.93 |
| Sample Statistics | | | | | | | |
| Years | 90-99 | 84-93 | 76-87 | 93-97 | 81-97 | 94-03 | 95-05 |
| Nobs | $37.7M$ | $4.3M$ | $5.3M$ | $4.8M$ | — | $6.9M$ | $16.0M$ |
| Nworkers | $5.2M$ | $293K$ | $1.2M$ | $1.8M$ | $1.7M$ | $563K$ | $2.0M$ |
| Nfirms | $476K$ | $80K$ | $500K$ | 1821 | $421K$ | $53.6K$ | $137K$ |
| % 1st Group[e] | — | 99.1% | 88.3% | 94.9% | 99.5% | — | 98.6% |

(a) "US1" from Woodcock [41], which covers two non-identified states, and includes all workers who were employed in 1997. "US2" and "Fr" from Abowd et al [2]. The US data covers 1/10 of workers in the state of Washington, whereas the French data covers 1/25 of all workers. "GE" from Andrews et al [4] and uses data from around 2000 establishments in West Germany. "IT" from Iranzo et al [22], which covers 1200 plants with at least 50 workers. "DE" from Bagger and Lentz [5], which covers covers all Danish population. "BR" refers to our own calculations.
(b) This study uses a random effects estimator under the assumption that the two components of heterogeneity are orthogonal.
(c) Iranzo et al [22] compute the index of segregation proposed by Kremer and Maskin [24], using worker fixed effects from the AKM regression as their measure of skill. When firms are large (as in their sample) that measure is very similar to our worker co-worker measure. However, they use Pearson correlations instead of rank correlations.
(d) This number was provided by the authors, and may not come from the same sample described on the table. Also, that was computed using the fixed effects method, not random effects.
(e) This denotes the fraction of the sample in the largest connected group.
(f) We use rank correlations.

# Results & Caveats

- Results in the literature pre 2010:
  - firm heterogeneity explains $20\%$ to $30\%$ of explained variance
  - correlation between types is zero or negative
  - this suggests no or negative sorting in the labor market
- Possible pitfalls:
  - additivity is not the correct specification
  - presence of bias due to small $T$ or small $N$
  - endogenous mobility?
- let's look at additivity and biases

# Linearity: Card and Kline QJE plot



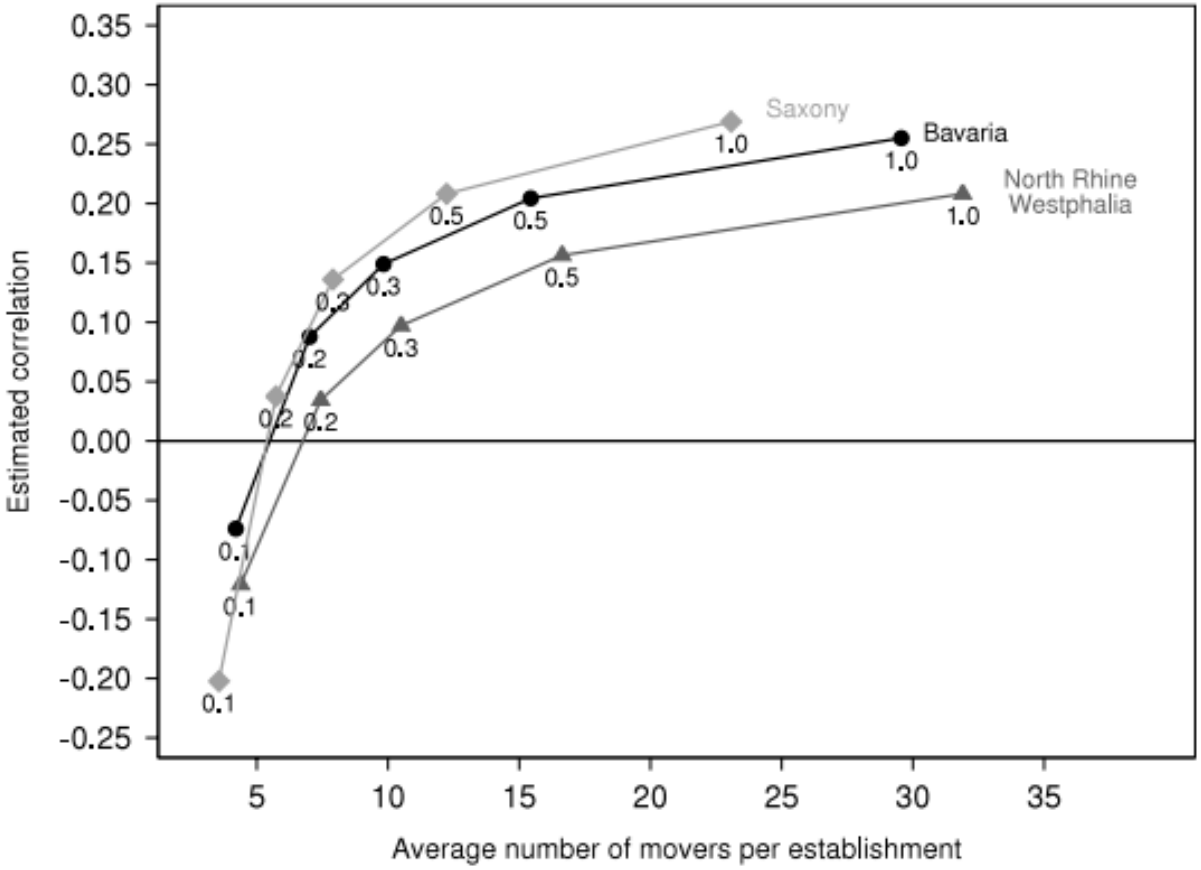

- wage gains and losses appear to be symmetric
- "suggests" linearity
- we will come back to this plot in the last section

# Limited mobility bias

- remember that $\hat{\alpha}_i = \frac{1}{n_i} \sum_t \left( y_{it} - \hat{\psi}_{j(i,t')} \right)$
- if there are only a few movers, noise in the construction of $\hat{\psi}$ enters **negatively** in $\hat{\alpha}_i$
- this can bias up $Var(\psi_j)$ and negatively $cov(\alpha_i, \psi_{j(i,t)})$
- Andrews, Gill, Schank, and Upward (2012) documents this possibility
  - use German data
  - keep set of establishment fixed
  - varies number of movers

# Limited mobility bias

| $p$ | Bavaria $J = 65{,}032$ | | North-Rhine Westphalia $J = 84{,}564$ | | Saxony $J = 19{,}877$ | |
|---|---|---|---|---|---|---|
| | $N^*$ | $M/J$ | $N^*$ | $M/J$ | $N^*$ | $M/J$ |
| 0.1 | 1,779,562 | 4.2 | 2,309,319 | 4.4 | 436,766 | 3.6 |
| 0.2 | 3,393,479 | 7.0 | 4,409,560 | 7.4 | 820,059 | 5.7 |
| 0.3 | 5,003,038 | 9.8 | 6,519,154 | 10.5 | 1,205,597 | 7.9 |
| 0.5 | 8,214,938 | 15.4 | 10,735,633 | 16.6 | 1,977,795 | 12.2 |
| 1.0 | 16,278,473 | 29.6 | 21,270,334 | 31.9 | 3,904,445 | 23.1 |



# Limited mobility bias

- remember that $\hat{\alpha}_i = \frac{1}{n_i} \sum_t \left( y_{it} - \hat{\psi}_{j(i,t')} \right)$
- if there are only a few movers, noise in the construction of $\hat{\psi}$ enters **negatively** in $\hat{\alpha}_i$
- this can bias up $Var(\psi_j)$ and negatively $cov(\alpha_i, \psi_{j(i,t)})$
- Andrews, Gill, Schank, and Upward (2012) documents this possibility
  - use German data
  - keep set of establishment fixed
  - varies number of movers
- Card, Heining, and Kline (2013) uses a very large data: $15$M
  - find strong sorting ( at least in late years)
  - maybe bias is not an important issue if data is big enough

# Card, Heining, and Kline (2013)

Decomposition of the Rise in Wage Inequality

| | Interval 1 (1985–1991) | | Interval 4 (2002–2009) | | Change from interval 1 to 4 | |
|---|---|---|---|---|---|---|
| | (1) Var. component | (2) Share of total | (3) Var. component | (4) Share of total | (5) Var. component | (6) Share of total |
| Total variance of log wages | 0.137 | 100.0 | 0.249 | 100.0 | 0.112 | 100 |
| Components of variance: | | | | | | |
| Variance of person effect | 0.084 | 61.3 | 0.127 | 51.2 | 0.043 | 39 |
| Variance of establ. effect | 0.025 | 18.5 | 0.053 | 21.2 | 0.027 | 25 |
| Variance of Xb | 0.015 | 10.7 | 0.007 | 2.8 | –0.008 | –7 |
| Variance of residual | 0.011 | 8.2 | 0.015 | 5.9 | 0.003 | 3 |
| 2cov(person, establ.) | 0.003 | 2.3 | 0.041 | 16.4 | 0.038 | 34 |
| 2cov(Xb, person + establ.) | –0.001 | –1.0 | 0.006 | 2.4 | 0.007 | 7 |
| Counterfactuals for variance of log wages* | | | | | | |
| 1. No rise in correl. of person/estab. effects | 0.137 | | 0.213 | | 0.077 | 69 |
| 2. No rise in var. of estab. effect | 0.137 | | 0.209 | | 0.072 | 64 |
| 3. Both 1 and 2 | 0.137 | | 0.184 | | 0.047 | 42 |

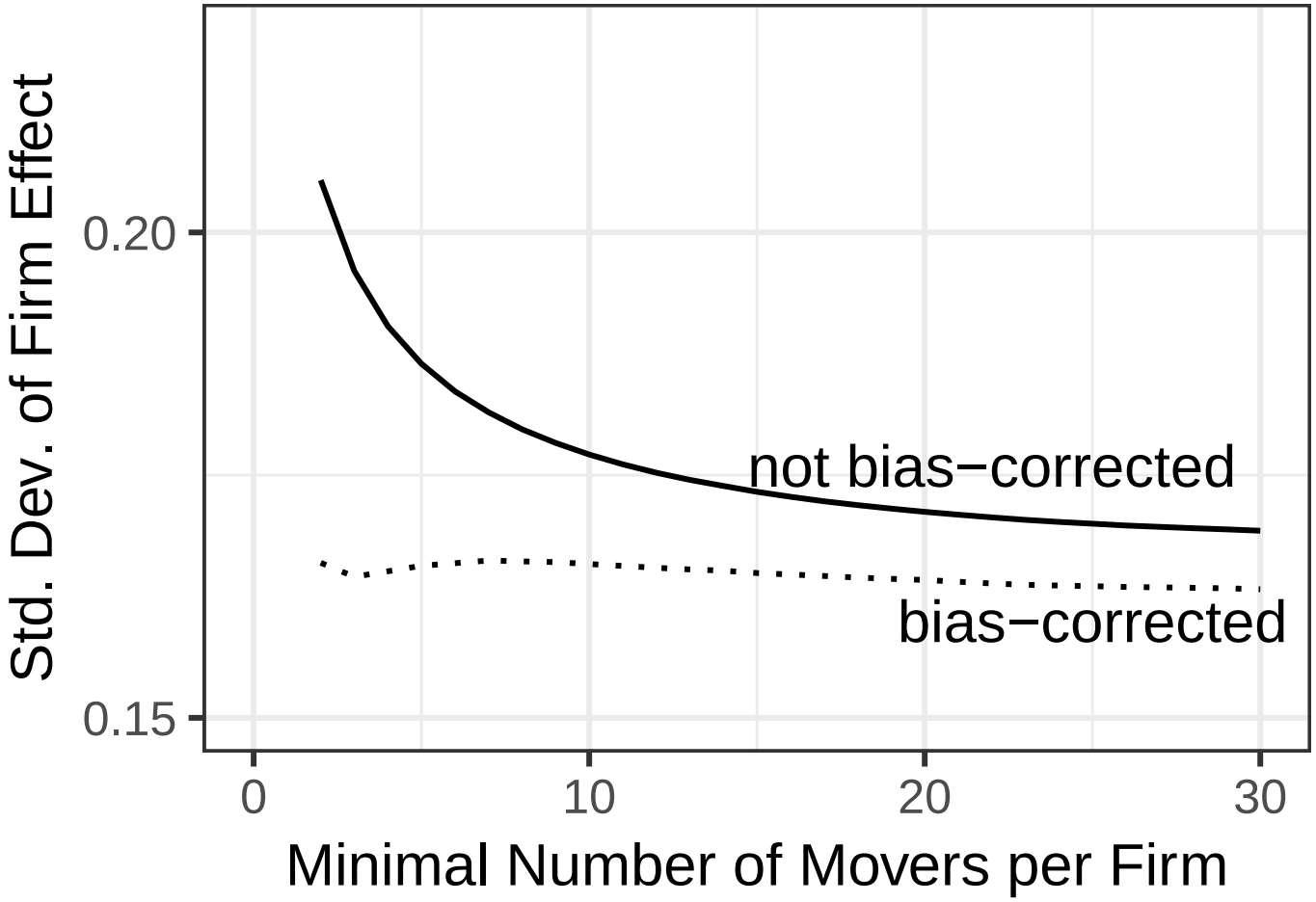
- they find that establishment heterogeneity and sorting are the drivers of increase in inequality
- sorting does affect inequality over time

# Limited mobility bias

- Estimate of firm effect $\hat{\psi}_j = \psi_j + u_j$
- Then $Var(\hat{\psi}_j) \simeq Var(\psi_j) + \frac{\sigma^2}{n_m}$
- Dhaene and Jochmans (2015) propose a Jacknife method to reduce incidental parameter bias in panel data settings
- Bonhomme, Lamadon, Manresa proposes to use it on movers:
  - don't split N, split the movers
  - $Var_{\text{split1}}(\hat{\psi}_j) \simeq Var(\psi_j) + 2 \cdot \frac{\sigma^2}{n_m}$
- The procedure is then:
  1. split movers in 2 sub-samples
  2. compute AKM on all data and in each sub-sample
  3. correct param. estimates:

$$\theta^{BR} = 2\theta - \frac{\theta_{s1} + \theta_{s2}}{2}$$

# Lamadon, Mogstad and Setzler WP



# Lamadon, Mogstad and Setzler WP

| | Card, et al. | | Ours | | Ours, Bias-corrected | |
|---|---|---|---|---|---|---|
| **Impose Flat Earnings Profile:** | **Age 40** | **Age 50** | **Age 40** | **Age 50** | **Age 40** | **Age 50** |
| **Panel A.** | | | **Levels** | | | |
| Total SD ($\log W$) | | | 0.69 | 0.69 | 0.69 | 0.69 |
| Person Effects SD ($x$) | 0.42 | 0.41 | 0.56 | 0.56 | 0.55 | 0.55 |
| Firm Effects SD ($\psi$) | 0.25 | 0.25 | 0.21 | 0.21 | 0.18 | 0.18 |
| Covariates SD ($Xb$) | 0.07 | 0.10 | 0.14 | 0.15 | 0.14 | 0.14 |
| Correlation: $x$ and $\psi$ | 0.17 | 0.16 | 0.13 | 0.13 | 0.27 | 0.27 |
| Correlation: $x$ and $Xb$ | 0.19 | 0.19 | -.00 | -.02 | -.01 | -.02 |
| Correlation: $\psi$ and $Xb$ | 0.11 | 0.14 | 0.04 | 0.05 | 0.05 | 0.06 |
| **Panel B.** | | | **Percentages** | | | |
| $Var(x + Xb)$ | 63% | 63% | 70% | 69% | 67% | 67% |
| $Var(x)$ | 58% | 58% | 66% | 65% | 63% | 63% |
| $Var(Xb)$ | 2% | 2% | 4% | 5% | 4% | 5% |
| $2Cov(x, Xb)$ | 3% | 3% | -0% | -1% | -0% | -1% |
| $Var(\psi)$ | 20% | 20% | 10% | 10% | 7% | 7% |
| $2Cov(\psi, x + Xb)$ | 12% | 12% | 7% | 7% | 12% | 12% |
| $2Cov(\psi, x)$ | 11% | 10% | 7% | 7% | 11% | 11% |
| $2Cov(\psi, Xb)$ | 1% | 2% | 1% | 1% | 1% | 1% |
| Residual | 5% | 5% | 14% | 14% | 14% | 15% |

# Conclusion

- The log linear model is a very tractable way to approach the problem
- Potential caveats are:
  - mobility is more complicated
  - additivity in the wage function is incorrect
  - limited mobility bias, which can be dramatic in some samples
- The framework is applied to other economic questions:
  - Health Care Utilization: patient health condition versus geographic location: Finkelstein, Gentzkow, Williams (2015, QJE forth)
  - Intergenerational Mobility: child ability vs neighborhood: Chetty Hendren (2015)
- We now turn our attention to theoretical foundation of sorting!

# Table of content

# References


Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67(2), 251–333.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2012): "High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias," *Econ. Lett.*, 117(3), 824–827.

Becker, G. S. (1974): "A theory of marriage," in *Economics of the family: Marriage, children, and human capital*, pp. 299–351. University of Chicago Press.

Bonhomme, S., T. Lamadon, and E. Manresa (2015): "A Distributional Framework for Matched Employer Employee Data," *Working Paper*.

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality*," *Q. J. Econ.*, 128(3), 967–1015.

Guimaraes, P., P. Portugal, et al. (2010): "A simple feasible procedure to fit models with high-dimensional fixed effects," *Stata Journal*, 10(4), 628.

Hagedorn, M., T. H. Law, and I. Manovskii (2014): "Identifying