Workplace Heterogeneity and the Rise of West German Wage Inequality

David Card, Jorg Heining, Patrick Kline

Quarterly Journal of Economics, August 2013

Outline

Slichter (1950): a 1940 wage survey from Boston

	Average Hourly Earnings in All Plants (cents)	Average Hourly Earnings in Lowest Plant (cents)	Average Hourly Earnings in Highest Plant (cents)	Spread between High and Low Plants (cents)
Common labor	57.9	44.8	74.1	29.3
Janitor	55.3	41.0	70.5	29.5
Watchman	59.6	45.2	74.0	28.8
Producing and processing laborers	64.2	44.8	100.7	55.9
Producing and processing operators	72.0	57.8	88.6	30.8
Receiving and shipping clerks	68.0	50.0	89.6	39.6
Machinists	87.5	70.0	105.0	35.0
Steamfitter	86.4	70.0	105.0	35.0
Electrician	88.0	67.9	105.0	37.1
Carpenter	81.5	65.0	99.5	34.5
Sheet metal workers	85.4	77.8	90.5	12.7
Millwright	86.1	82.5	95.5	13.0
Maintenance helper	67.1	50.7	82.0	31.3
Female producing laborers	45.1	33.8	63.4	29.6
Female producing operators	47.9	37.7	58.3	20.6
Firemen	78.4	63.0	90.8	27.8

"Hourly earnings do not represent the price of labor."

Slichter (1950): firms set wages

1. Positive correlation with wages of skilled co-workers
2. Negative correlation with % female
3. Positive correlation with industry value-added/worker-hour
4. Positive correlation with sales/worker-hour
5. Negative correlation with payroll/sales
6. Positive correlation with profit margin
7. Stable over time (high correlation of industry wage rank)

"*The results of this study give strong support to the proposition that managerial policy is important in firms setting wages.*"

Krueger and Summers (1988): noncompetitive rents

- Was Slichter (1950) right that firms set wages?
- Use panel data and include worker FE to account for sorting between industries based on unmeasured ability.
- Industry wage-premia are similar to cross-sectional estimates.
- Industry wage-premia are not a compensating differential.
- People don't quit high wage jobs.

Workers in high-wage industries receive noncompetitive rents.

Variance of industry FE

- Target parameter is size weighted variance of industry effects:

$$\theta_\psi \equiv Var(\psi_j) = \sum_{j=1}^{J} s_j(\psi_j - \bar{\psi})^2$$

where s_j is firm j's employment share and $\bar{\psi} = \sum_{j=1}^{J} s_j \psi_j$.

- Use OLS estimates $\hat{\psi}_j$ to compute "plug-in" estimates of variance components:

$$\begin{aligned}\hat{\theta}_\psi &= \sum_{j=1}^{J} s_j(\hat{\psi}_j - \hat{\bar{\psi}})^2 \\ &= \sum_{j=1}^{J} s_j(\hat{\psi}_j)^2 - (\hat{\bar{\psi}})^2\end{aligned}$$

Bias in the square of an unbiased estimate

- OLS estimates $\hat{\psi}_j$ are unbiased:

$$\mathbb{E}\left[\hat{\psi}_j\right] = \psi_j$$

- But the square of an unbiased estimator is upward biased:

$$\begin{aligned}
\mathbb{E}\left[\left(\hat{\psi}_j\right)^2\right] &= \mathbb{E}\left[\left(\hat{\psi}_j - \psi_j + \psi_j\right)^2\right] \\
&= \mathbb{E}\left[\left(\hat{\psi}_j - \psi_j\right)^2\right] + 2\mathbb{E}\left[\hat{\psi}_j - \psi_j\right]\psi_j + \psi_j^2 \\
&= \psi_j^2 + \underbrace{\mathbb{V}\left[\hat{\psi}_j\right]}_{\text{bias}}
\end{aligned}$$

- Bias is the variance of estimated industry FE.

Bias in plug-in estimator of variance of industry FE

- Similarly, the estimated variance of industry FE is biased:

$$\begin{aligned}\mathbb{E}\left[\hat{\theta}_\psi\right] &= \sum_{j=1}^{J} s_j \mathbb{E}\left[\hat{\psi}_j{}^2\right] - \mathbb{E}\left[\hat{\bar{\psi}}^2\right] \\ &= \sum_{j=1}^{J} s_j \left\{\psi_j^2 + \mathbb{V}[\,\hat{\psi}_j\,]\right\} - \bar{\psi}^2 - \mathbb{V}\left[\hat{\bar{\psi}}\right]\end{aligned}$$

- When J is large, the last term is negligible:

$$\mathbb{E}\left[\hat{\theta}_\psi\right] \approx \theta_\psi + \underbrace{\sum_{j=1}^{J} s_j \mathbb{V}\left[\hat{\psi}_j\right]}_{\text{bias}}$$

- Bias is weighted variance of estimated industry FE.

Substantial cross-sectional variability in industry wage-premia

TABLE I

ESTIMATED WAGE DIFFERENTIALS FOR ONE-DIGIT INDUSTRIES—MAY CPS[a]
(Standard Errors in Parentheses)

Industry	(1) 1974	(2) 1979	(3) 1984	(4) 1984 Total Compensation
Construction	.195	.126	.108	.091
	(.021)	(.031)	(.034)	(.035)
Manufacturing	.055	.044	.091	.131
	(.020)	(.029)	(.032)	(.032)
Transportation & Public Utilities	.111	.081	.145	.203
	(.021)	(.031)	(.034)	(.034)
Wholesale & Retail Trade	−.128	−.082	−.111	−.136
	(.020)	(.030)	(.033)	(.033)
Finance, Insurance and Real Estate	.047	−.010	.055	.069
	(.022)	(.035)	(.034)	(.034)
Services	−.070	−.055	−.078	−.111
	(.021)	(.030)	(.032)	(.032)
Mining	.179	.229	.222	.231
	(.035)	(.058)	(.075)	(.075)
Weighted Adjusted Standard Deviation of Differentials[b]	.097**	.069**	.094**	.126**
Sample Size	29,945	8,978	11,512	11,512

[a] Other explanatory variables are education and its square, 6 age dummies, 8 occupation dummies, 3 region dummies, sex dummy, race dummy, central city dummy, union member dummy, ever married dummy, veteran status, marriage × sex interaction, education × sex interaction, education squared × sex interaction, 6 age × sex interactions, and a constant. Each column was estimated from a separate cross-sectional regression.

[b] Weights are employment shares for each year.

** F test that industry wage differentials jointly equal 0 rejects at the .000001 level.

Estimates including worker FE are similar

TABLE IV

THE EFFECTS OF UNMEASURED LABOR QUALITY[a]

Industry	(1) Fixed Effects Unadjusted for Measurement Error	(2) Fixed Effects Adjusted for Measurement Error I[b]	(3) Fixed Effects Adjusted for Measurement Error II[c]	(4) Levels
Construction	.063	.098	.174	.174
	(.033)	(.060)	(.060)	(.024)
Manufacturing	.028	.055	.107	.064
	(.031)	(.058)	(.058)	(.022)
Transportation and	.019	.060	.049	.114
Public Utilities	(.035)	(.059)	(.059)	(.024)
Wholesale and	−.042	−.068	−.125	−.133
Retail Trade	(.031)	(.056)	(.056)	(.023)
Finance, Insurance	.027	.017	.018	.035
and Real Estate	(.036)	(.061)	(.061)	(.025)
Services	−.040	−.088	−.128	−.079
	(.032)	(.056)	(.057)	(.023)
Mining	.067	.122	.142	.156
	(.004)	(.057)	(.058)	(.040)

[a] Data set is three matched May CPS's pooled together: 1974–1975, 1977–1978, and 1979–1980. Sample size is 18,122. Levels are 1974, 1977, and 1979 data pooled. Results of the 1975, 1978, and 1980 sample are qualitatively the same. Controls for fixed effects regressions are change in education and its square, change in occupation, 3 region dummies, change in union membership, experience squared, change in marital status, year dummies, and a constant. Controls for level regressions are the same as Table I plus year dummies.

[b] Adjustment I assumes 3.4 per cent error rate and that misclassifications are proportional to industry size. See Appendix for description.

[c] Adjustment II assumes average error rate is 3.4 per cent and misclassifications are allocated according to employer-employee mismatches. See Appendix for description.

No evidence of compensating differentials

TABLE VI

ANALYSIS OF INDUSTRY WAGE DIFFERENTIALS WITH AND WITHOUT CONTROLS FOR WORKING CONDITIONS—QES 1977[a]

Industry	Coefficient (SE) (1)	Coefficient (SE) (2)
Construction	.113	.100
	(.098)	(.100)
Manufacturing	.050	.046
	(.086)	(.087)
Transportation	.113	.124
	(.095)	(.096)
Wholesale & Retail Trade	−.056	−.061
	(.090)	(.091)
Finance, Insurance and Real Estate	.071	.053
	(.104)	(.105)
Services	−.107	−.104
	(.090)	(.091)
Mining	.233	.308
	(.205)	(.220)
10 Working Condition Variables[b]	no	yes
Weighted Adjusted Standard Deviation of 2-Digit Industry Premiums	.113*	.118*
R^2	.496	.519

[a] Other explanatory variables are education and its square, derived experience and its square, sex, race, 3 region dummies, tenure with employer and its square, union status, and 8 occupation dummies. Sample size is 1,033.

[b] Working condition variables are weekly hours, variables indicating dangerous or unhealthy conditions on the job and whether the danger/threat is serious, commuting time, second and third shift dummies, two dummies indicating extent of choice of overtime, and two dummies indicating whether the physical working conditions are pleasant.

* F test that industry wage differentials jointly equal 0 is rejected at .00005 level.

Workers don't quit high-wage jobs

TABLE IX

THE EFFECT OF INDUSTRY WAGE DIFFERENTIALS ON JOB TENURE AND QUITS

Independent Variables	Dependent Variable[a] (1) Tenure	(2) Quit[b]
Industry wage premium	2.198 (.676)	−.073 (.135)
Union (1 = yes)	3.179 (.157)	−.164 (.037)
Other variables	Age dummies (6), Age * Sex (6), Education, Education Squared * Sex, Region Dummies (3), Race Dummy, Sex Dummy, Central City Dummy, Firm Size Dummies (4), Plant Size Dummies (4), Marriage Dummy, Marriage * Sex, Veteran Status Dummy	Education, Education Squared, Region Dummies (3), Race Dummy, Sex Dummy, SMSA Dummy, (Age—Education—5) and its square
Sample Size	8,978	633
R^2	.40	.20

[a] Mean (SD) of Tenure is 5.70 (7.61); Mean (SD) of Quit is .26 (.44).
[b] Quit equation was estimated with a linear probability model.

Abowd, Kramarz and Margolis (1999): AKM

- Use matched employer-employee data to study firm FE
- High-dimensional FE specification becomes:

$$y_{it} = \alpha_i + \psi_{J(i,t)} + x_{it}'\beta + \epsilon_{it}$$

- Key finding in constrast to Krueger and Summers (1988): 90% of industry wage-premia attributable to person effects.
- However, can't invert design matrix with millions of FE.
- Instead, they use approximate solution method to estimate FE.

Revisiting Abowd, Kramarz and Margolis (1999)

1. **Abowd, Creecy and Kramarz (2002)**:
    - Approximate FE very weakly correlated with exact FE in French data ⇒ original AKM results invalid!
    - Exact results find firm FE explain 55% of variation in wages in France and 45% in Washington state.
    - But $Cov(\alpha_i, \psi_{J(i,t)}) < 0 \Rightarrow$ negative assortative matching?
2. **Abowd, Lengermann, and McKinney (2003)**:
    - Use a 100% sample extract from LEHD of 7 US states instead of small subsamples.
    - Firm FE explain 20% of variation in wages (less inflation due to sampling error as we explained before).
    - Assortative matching becomes positive! (less limited-mobility bias because larger connected set as we'll discuss later)

Card, Heining and Kline (2013): AKM revival

- Study changes in German wage structure.
- Earlier work by Dustmann, Ludsteck and Schoenberg (2009) documented increase in German wage dispersion.
- Interpreted within traditional SDI framework
- Typical view: S+D influence the price of skill, I is barrier to price adjustment.
- What about the importance of firms?

Outline

Data

TABLE I

SUMMARY STATISTICS FOR SAMPLES OF FULL-TIME MEN AND WOMEN

	(1) Number observations	Log real wage, unallocated (2) Mean	(3) Std. dev.	(4) Percent censored	Log real wage, allocated (5) Mean	(6) Std. dev.
Panel A. Full-time men						
1985	11,980,159	4.221	0.387	10.63	4.247	0.429
1990	13,289,988	4.312	0.398	11.92	4.342	0.445
1995	13,101,809	4.340	0.415	9.78	4.361	0.447
2000	12,930,046	4.327	0.464	10.31	4.352	0.502
2005	11,857,526	4.310	0.519	9.36	4.336	0.562
2009	12,104,223	4.277	0.535	10.00	4.308	0.586
Panel B. Full-time women						
1985	6,068,863	3.836	0.462	1.52	3.840	0.470
1990	7,051,617	3.942	0.476	2.01	3.947	0.486
1995	7,030,596	4.026	0.483	1.95	4.030	0.491
2000	7,009,075	4.019	0.532	2.47	4.026	0.545
2005	6,343,006	3.999	0.573	2.36	4.006	0.588
2009	6,566,429	3.979	0.587	2.80	3.988	0.606

Notes. Samples includes employees in West Germany age 20–60, working full-time in nonmarginal jobs. Real wage is based on average daily earnings at the full-time job with highest total earnings during the calendar year, adjusted for inflation using the Consumer Price Index. Unallocated wage data in columns (2) and (3) are based on raw daily wage data, which are censored at social security maximum for the corresponding year. Allocated wage data in columns (5) and (6) include stochastic allocation of censored observations based on a Tobit model.

Trends in Wage Inequality

Rise in wage inequality

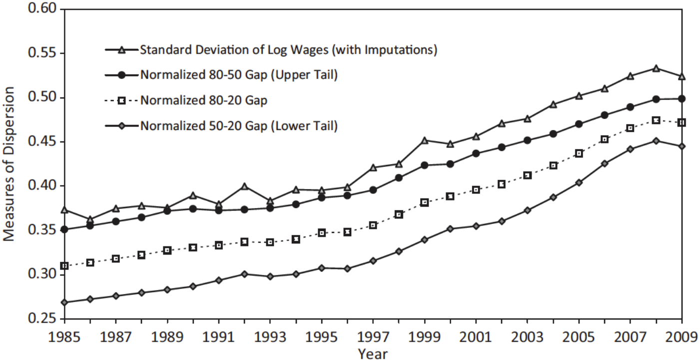
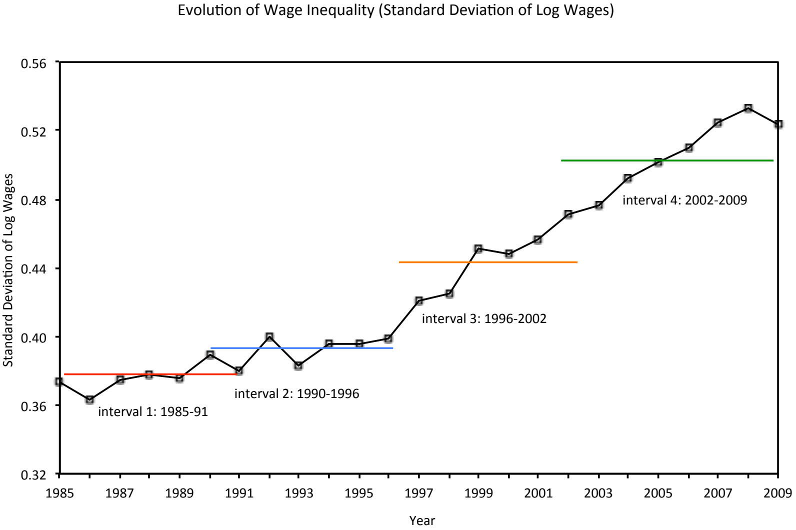


FIGURE II

Trends in Wage Inequality for Full-Time Male Workers

This figure shows measures of dispersion in real daily wage for full-time male workers. Normalized percentile gaps are differences in percentiles divided by corresponding differences in percentiles of standard normal variate.

Overlapping intervals



Focus on full-time jobs held by men between age 20-60

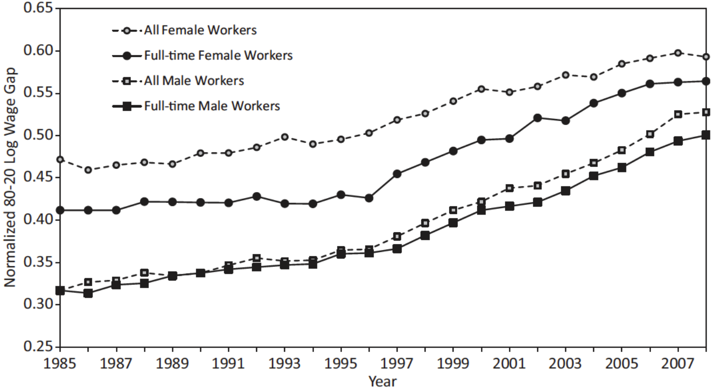


FIGURE III

Wage Inequality Trends for Alternative Samples of Workers

Based on tabulations of SIAB. Measured wage is average daily wage in job with highest total earnings in the year. Wage gap is the difference between the 80th percentile of log real wages and the 20th percentile, divided by 80–20 gap for a standard normal variate.

Growth in wage inequality primarily between firms

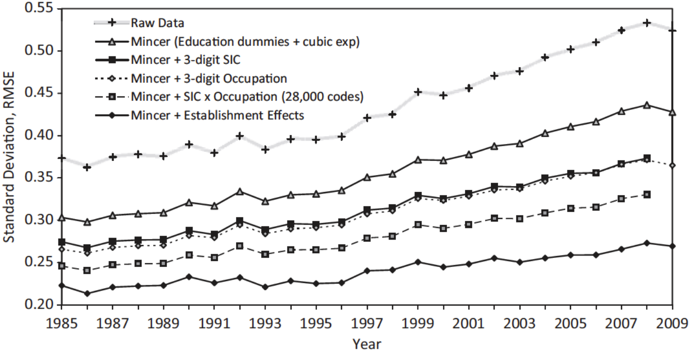


FIGURE IV

Raw and Residual Standard Deviations from Alternative Wage Models

See note to Figure II. Figure shows measures of dispersion in actual and residual real daily wage for full-time male workers. Residual wage is residual from linear regression model. "Mincer" refers to model with dummies for education categories and cubic in experience, fit separately in each year. Other models add controls as indicated.

Effect of Job Changes on Wages

Effect of Job Changes on Wages

- If wage variation across firms is due to workers sorting based on unobserved ability, then workers changing firms do not experience systematic changes in wages.
- But if different firms pay different wage premia to all their workers, then workers who join high-wage (low-wage) firms experience a wage gain (loss).
- Examine effect of moving to another firm on individual wages between 1985-1991 and 2002-2009.
- Origin workplace and destination workplace defined by average coworker wages.

Wage dynamics of job changes

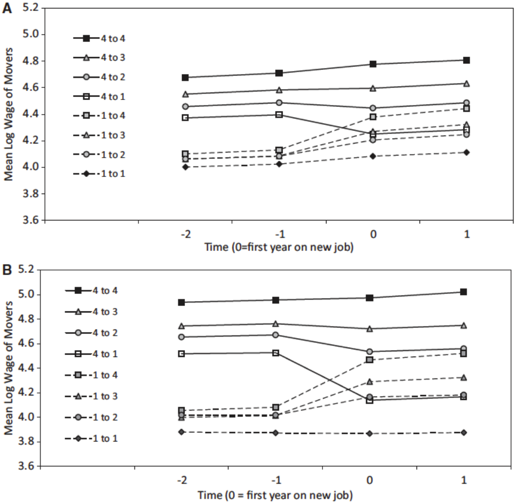


FIGURE V

Mean Wages of Job Changers Classified by Quartile of Mean Wage of Coworkers at Origin and Destination Establishment (A) 1985–1991, (B) 2002–2009

Outline

AKM estimator

- Log daily real wage y_{it} of individual i in year t is given by:

$$y_{it} = \alpha_i + \psi_{J(i,t)} + x_{it}'\beta + r_{it} \tag{1}$$

- In matrix notation:

$$y = D\alpha + F\psi + X\beta + r = Z'\xi + r \tag{2}$$

- OLS solves the standard normal equations:

$$Z'Z\xi = Z'y \tag{3}$$

- Use matlab packages to find the "largest connected set" and solve normal equations using preconditioned conjugate gradient routine.

Largest connected set

TABLE II
SUMMARY STATISTICS FOR OVERALL SAMPLE AND INDIVIDUALS IN LARGEST CONNECTED SET

	All full-time men, age 20–60				Individuals in largest connected set			
			Log real daily wage				Log real daily wage	
Interval	(1) Number person/yr. obs.	(2) Number individuals	(3) Mean	(4) Std. dev.	(5) Number person/yr. obs.	(6) Number individuals	(7) Mean	(8) Std. dev.
1985–1991	86,230,097	17,021,779	4.344	0.379	84,185,730	16,295,106	4.351	0.370
Ratio: largest connected/all					97.6	95.7	100.2	97.7
1990–1996	90,742,309	17,885,361	4.391	0.392	88,662,398	17,223,290	4.398	0.384
Ratio: largest connected/all					97.7	96.3	100.2	97.9
1996–2002	85,853,626	17,094,254	4.397	0.439	83,699,582	16,384,815	4.405	0.432
Ratio: largest connected/all					97.5	95.8	100.2	98.3
2002–2009	93,037,963	16,553,835	4.387	0.505	90,615,841	15,834,602	4.397	0.499
Ratio: largest connected/all					97.4	95.7	100.2	98.8
Change from first to last interval			0.043	0.126			0.045	0.128

Notes. Sample consists of full-time male workers ages 20–60 employed in nonmarginal jobs and not currently in training. Daily wage is imputed for censored observations using a Tobit model. "Connected set" refers to group of firms connected by worker mobility over the sample interval (for details, see Abowd, Creecy, and Kramarz 2002).

Exogeneity assumptions

- OLS requires the following independence conditions to hold:

$$\forall i : E[d^{i\prime} r] = 0; \forall j : E[f^{j\prime} r] = 0; \forall k : E[x^{k\prime} r] = 0 \qquad (4)$$

- $\forall k : E[x^{k\prime} r] = 0$ is standard assumption that error term is independent of time-varying covariates.
- $\forall i : E[d^{i\prime} r] = 0$ holds if all components of the error term have mean zero for each worker across years.
- $\forall j : E[f^{j\prime} r] = 0$ assumes that there is no sorting of workers to firms based on time-varying unobserved components of wages.

Specifying the error term

- Assume r_{it} consists of three separate random effects:

$$r_{it} = \eta_{iJ(i,t)} + \zeta_{it} + \epsilon_{it}$$

  with
  - $\eta_{iJ(i,t)}$ an idiosyncratic match component
  - ζ_{it} drift in a portable component
  - ϵ_{it} a mean-reversing transitory component.
- We assume that $\forall k : E[x^{k\prime} r] = 0$ and $\forall i : E[d^{i\prime} r] = 0$.
- Condition $\forall j : E[f^{j\prime} r] = 0$ requires that sorting of workers to firms based on time-varying unobserved component of wages is independent of firm FE.

Exogenous assignment of workers to firms

- Sufficient condition is that assignment of workers to firms is independent of r:

$$P(J(i,t) = j|r) = P(J(i,t) = j) = G_{jt}(\alpha_i, \psi_1, ..., \psi_J) \quad (5)$$

  with G_{jt} the employment probability function summing to 1 for every worker in every period.

- Systematic patterns of job mobility based on unobserved person-effects and firm-effects allowed!

- The paper discusses 3 violations.

V1: Selection on match component ($Cov(\psi_j, \eta_{ij}) \neq 0$)

- Possible if lots of scope for comparative advantage (a la Roy (1952)) and worker has bargaining power.
- Transitions in both directions should usually be associated with wage increases:

$$\mathbb{E}\left[y_{i,t} - y_{i,t-1} | 1 \to 4\right] = -\psi_1 + \psi_4 + \mathbb{E}\left[\eta_{i,4} - \eta_{i,1} | 1 \to 4\right]$$
$$\mathbb{E}\left[y_{i,t} - y_{i,t-1} | 4 \to 1\right] = \psi_1 - \psi_4 + \mathbb{E}\left[\eta_{i,1} - \eta_{i,4} | 4 \to 1\right]$$

- And unrestricted match effects model should fit much better than worker FE + firm FE.

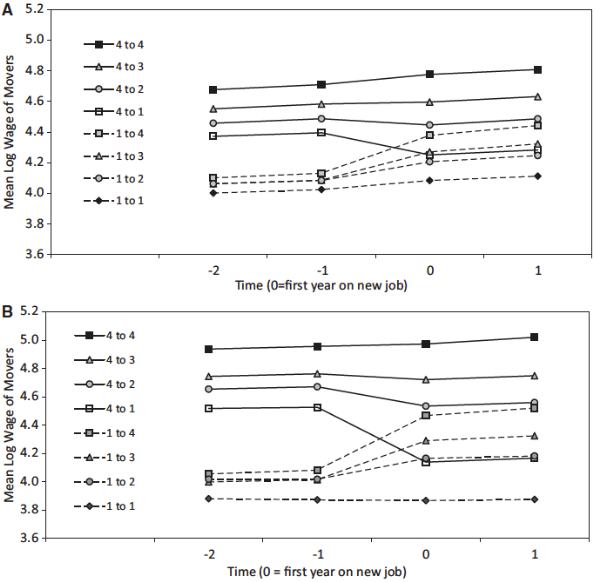


FIGURE V

Mean Wages of Job Changers Classified by Quartile of Mean Wage of Coworkers at Origin and Destination Establishment (A) 1985–1991, (B) 2002–2009

V2: Selection on drift component ($Cov(\psi_j, \zeta_{it}) \neq 0$)

1. Possible if firms learn rapidly about workers and learning is associated with job-to-job mobility (Gibbons & Katz, 1992).
2. But, learning takes years (Lange, 2007).
   - Ought to see an increasing trend in event study before upward transitions and decreasing trend before downward transitions.
3. Another violation would be offer shopping in job ladder models (e.g. Postel-Vinay and Robin, 2002)
   - Difficulty explaining symmetry in event study.

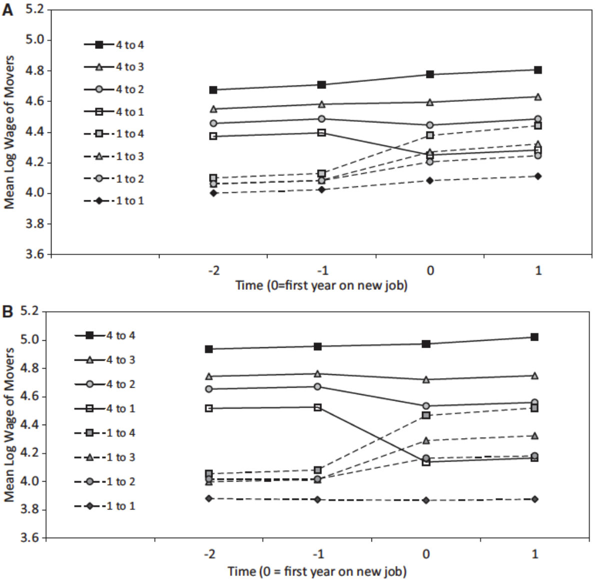


FIGURE V

Mean Wages of Job Changers Classified by Quartile of Mean Wage of Coworkers at Origin and Destination Establishment (A) 1985–1991, (B) 2002–2009

V3: Selection on transitory component ($Cov(\psi_j, \epsilon_{it}) \neq 0$)

- Possible if firm has a bad year, wages fall, and people leave the following year.
- Understate the establishment effect at origin and overstates it at destination.
- But should see an Ashenfelter (1978) style dip in event study.
- Also: shocks at each firm should eventually average out to zero as T grows large.

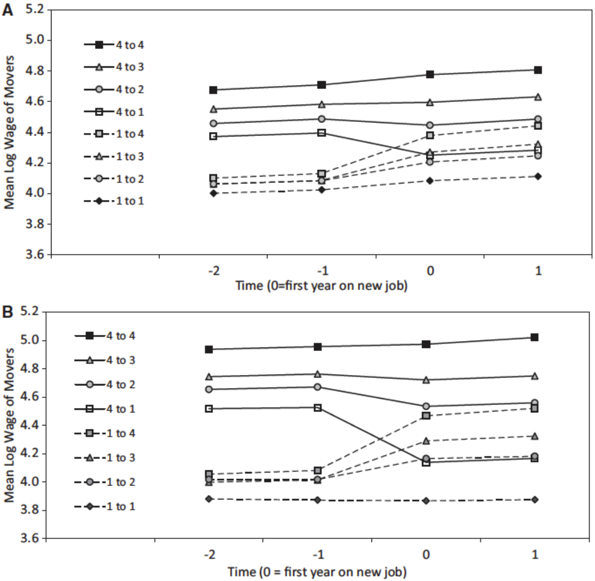


FIGURE V

Mean Wages of Job Changers Classified by Quartile of Mean Wage of Coworkers at Origin and Destination Establishment (A) 1985–1991, (B) 2002–2009

Outline

Introduction

Data

Econometric Model

Results

Decomposing between-group wage differentials

Institutional change as explanation

Conclusion

AKM Results

Wage dynamics of job changes with firm FE

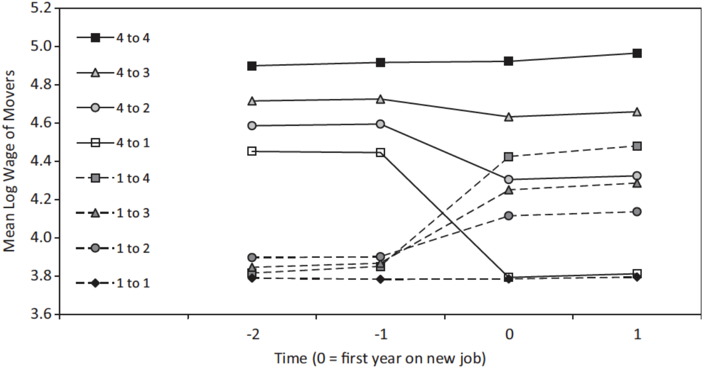


FIGURE VII

Mean Wages of Movers Classified by Quartile of Establishment Effects for Origin and Destination Firms, 2002–2009

Figure shows mean wages of male workers observed in 2002–2009 who change jobs in the interval and held the preceding job for two or more years, and the new job for two or more years. "Job" refers to main job in year, excluding part-time jobs. Each job is classified into quartiles based on estimated establishment effect from AKM model presented in Table III column 4.

AKM estimates

TABLE III

Estimation Results for AKM Model, Fit by Interval

	(1) Interval 1 1985–1991	(2) Interval 2 1990–1996	(3) Interval 3 1996–2002	(4) Interval 4 2002–2009
Person and establishment parameters				
Number person effects	16,295,106	17,223,290	16,384,815	15,834,602
Number establishment effects	1,221,098	1,357,824	1,476,705	1,504,095
Summary of parameter estimates				
Std. dev. of person effects (across person-year obs.)	0.289	0.304	0.327	0.357
Std. dev. of establ. Effects (across person-year obs.)	0.159	0.172	0.194	0.230
Std. dev. of Xb (across person-year obs.)	0.121	0.088	0.093	0.084
Correlation of person/establ. Effects (across person-year obs.)	0.034	0.097	0.169	0.249
Correlation of person effects/Xb (across person-year obs.)	−0.051	−0.102	−0.063	0.029
Correlation of establ. effects/Xb (across person-year obs.)	0.057	0.039	0.050	0.112
RMSE of AKM residual	0.119	0.121	0.130	0.135
Adjusted R-squared	0.896	0.901	0.909	0.927
Comparison match model				
RMSE of match model	0.103	0.105	0.108	0.112
Adjusted R^2	0.922	0.925	0.937	0.949
Std. dev. of match effect*	0.060	0.060	0.072	0.075
Addendum				
Std. dev. log wages	0.370	0.384	0.432	0.499
Sample size	84,185,730	88,662,398	83,699,582	90,615,841

Notes. Results from OLS estimation of equation (1). See notes to Table II for sample composition. Xb includes year dummies interacted with education dummies, and quadratic and cubic terms in age interacted with education dummies (total of 39 parameters in intervals 1–3, 44 in interval 4). Match model includes Xb and separate dummy for each job (person-establishment pair).

*Standard deviation of match effect estimated as square root of difference in mean squared errors between AKM model and match effect model.

Standard deviation of worker FE effects has increased

TABLE III

Estimation Results for AKM Model, Fit by Interval

	(1) Interval 1 1985–1991	(2) Interval 2 1990–1996	(3) Interval 3 1996–2002	(4) Interval 4 2002–2009
Person and establishment parameters				
Number person effects	16,295,106	17,223,290	16,384,815	15,834,602
Number establishment effects	1,221,098	1,357,824	1,476,705	1,504,095
Summary of parameter estimates				
Std. dev. of person effects (across person-year obs.)	0.289	0.304	0.327	0.357
Std. dev. of establ. Effects (across person-year obs.)	0.159	0.172	0.194	0.230
Std. dev. of Xb (across person-year obs.)	0.121	0.088	0.093	0.084
Correlation of person/establ. Effects (across person-year obs.)	0.034	0.097	0.169	0.249
Correlation of person effects/Xb (across person-year obs.)	−0.051	−0.102	−0.063	0.029
Correlation of establ. effects/Xb (across person-year obs.)	0.057	0.039	0.050	0.112
RMSE of AKM residual	0.119	0.121	0.130	0.135
Adjusted R-squared	0.896	0.901	0.909	0.927
Comparison match model				
RMSE of match model	0.103	0.105	0.108	0.112
Adjusted R^2	0.922	0.925	0.937	0.949
Std. dev. of match effect*	0.060	0.060	0.072	0.075
Addendum				
Std. dev. log wages	0.370	0.384	0.432	0.499
Sample size	84,185,730	88,662,398	83,699,582	90,615,841

Notes. Results from OLS estimation of equation (1). See notes to Table II for sample composition. Xb includes year dummies interacted with education dummies, and quadratic and cubic terms in age interacted with education dummies (total of 39 parameters in intervals 1–3, 44 in interval 4). Match model includes Xb and separate dummy for each job (person-establishment pair).

*Standard deviation of match effect estimated as square root of difference in mean squared errors between AKM model and match effect model.

Standard deviation of firm FE has increased

TABLE III

ESTIMATION RESULTS FOR AKM MODEL, FIT BY INTERVAL

	(1) Interval 1 1985–1991	(2) Interval 2 1990–1996	(3) Interval 3 1996–2002	(4) Interval 4 2002–2009
Person and establishment parameters				
Number person effects	16,295,106	17,223,290	16,384,815	15,834,602
Number establishment effects	1,221,098	1,357,824	1,476,705	1,504,095
Summary of parameter estimates				
Std. dev. of person effects (across person-year obs.)	0.289	0.304	0.327	0.357
Std. dev. of establ. Effects (across person-year obs.)	0.159	0.172	0.194	0.230
Std. dev. of Xb (across person-year obs.)	0.121	0.088	0.093	0.084
Correlation of person/establ. Effects (across person-year obs.)	0.034	0.097	0.169	0.249
Correlation of person effects/Xb (across person-year obs.)	−0.051	−0.102	−0.063	0.029
Correlation of establ. effects/Xb (across person-year obs.)	0.057	0.039	0.050	0.112
RMSE of AKM residual	0.119	0.121	0.130	0.135
Adjusted R-squared	0.896	0.901	0.909	0.927
Comparison match model				
RMSE of match model	0.103	0.105	0.108	0.112
Adjusted R^2	0.922	0.925	0.937	0.949
Std. dev. of match effect*	0.060	0.060	0.072	0.075
Addendum				
Std. dev. log wages	0.370	0.384	0.432	0.499
Sample size	84,185,730	88,662,398	83,699,582	90,615,841

Notes. Results from OLS estimation of equation (1). See notes to Table II for sample composition. Xb includes year dummies interacted with education dummies, and quadratic and cubic terms in age interacted with education dummies (total of 39 parameters in intervals 1–3, 44 in interval 4). Match model includes Xb and separate dummy for each job (person-establishment pair).

*Standard deviation of match effect estimated as square root of difference in mean squared errors between AKM model and match effect model.

Positive assortative matching has increased

TABLE III

ESTIMATION RESULTS FOR AKM MODEL, FIT BY INTERVAL

	(1) Interval 1 1985–1991	(2) Interval 2 1990–1996	(3) Interval 3 1996–2002	(4) Interval 4 2002–2009
Person and establishment parameters				
Number person effects	16,295,106	17,223,290	16,384,815	15,834,602
Number establishment effects	1,221,098	1,357,824	1,476,705	1,504,095
Summary of parameter estimates				
Std. dev. of person effects (across person-year obs.)	0.289	0.304	0.327	0.357
Std. dev. of establ. Effects (across person-year obs.)	0.159	0.172	0.194	0.230
Std. dev. of Xb (across person-year obs.)	0.121	0.088	0.093	0.084
Correlation of person/establ. Effects (across person-year obs.)	0.034	0.097	0.169	0.249
Correlation of person effects/Xb (across person-year obs.)	−0.051	−0.102	−0.063	0.029
Correlation of establ. effects/Xb (across person-year obs.)	0.057	0.039	0.050	0.112
RMSE of AKM residual	0.119	0.121	0.130	0.135
Adjusted R-squared	0.896	0.901	0.909	0.927
Comparison match model				
RMSE of match model	0.103	0.105	0.108	0.112
Adjusted R^2	0.922	0.925	0.937	0.949
Std. dev. of match effect*	0.060	0.060	0.072	0.075
Addendum				
Std. dev. log wages	0.370	0.384	0.432	0.499
Sample size	84,185,730	88,662,398	83,699,582	90,615,841

Notes. Results from OLS estimation of equation (1). See notes to Table II for sample composition. Xb includes year dummies interacted with education dummies, and quadratic and cubic terms in age interacted with education dummies (total of 39 parameters in intervals 1–3, 44 in interval 4). Match model includes Xb and separate dummy for each job (person-establishment pair).

*Standard deviation of match effect estimated as square root of difference in mean squared errors between AKM model and match effect model.

Positive assortative matching has increased

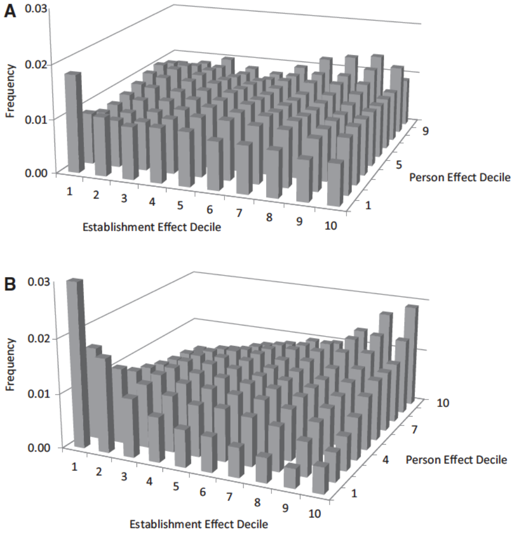


FIGURE VIII

Joint Distribution of Person and Establishment Effects (A) 1985–1991, (B) 2002–2009

Variance decomposition

Variance decomposition

- Variance of wages can be decomposed as:

$$\begin{aligned} Var(y_{it}) = {} & Var(\alpha_i) + Var(\psi_{J(i,t)}) + Var(x_{it}'\beta) \\ & + 2Cov(\alpha_i, \psi_{J(i,t)}) + 2Cov(\psi_{J(i,t)}, x_{it}'\beta) \\ & + 2Cov(\alpha_i, x_{it}'\beta) + Var(r_{it}) \end{aligned} \tag{6}$$

- $Var(\hat{\alpha}_i)$ and $Var(\hat{\psi}_{J(i,t)})$ are upward biased
- $Cov(\hat{\alpha}_i, \hat{\psi}_{J(i,t)})$ are downward biased due to limited-mobility bias in small largest connected sets (Andrews et al. 2008)
- No bias corrections assuming no trend in biases over intervals

TABLE IV

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

Notes. See notes to Table II for sample composition. Calculations based on estimated AKM models summarized in Table III. Entry in column (5) is change in variance component from interval 1 to interval 4. Entry in column (6) is ratio of the change in the variance component to the total change in variance of wages reported in first row of table (as a percentage).

*Counterfactual 1 computes the counterfactual rise in variance assuming the correlation between the person and establishment effects remains at its interval 1 value—that is, imposing the restriction that $Cov_4(\text{person, establ.}) = \rho_1 \, Var_4(\text{person})^{1/2} \times Var_4(\text{establ.})^{1/2}$ where the subscript 4 refers to the interval 4 value of the statistic and ρ_1 is the correlation between the person and establishment effects in interval 1. Counterfactual 2 assumes that the variance of establishment effects remains at its interval 1 level. Counterfactual 3 imposes both restrictions.

TABLE IV

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

Notes. See notes to Table II for sample composition. Calculations based on estimated AKM models summarized in Table III. Entry in column (5) is change in variance component from interval 1 to interval 4. Entry in column (6) is ratio of the change in the variance component to the total change in variance of wages reported in first row of table (as a percentage).

*Counterfactual 1 computes the counterfactual rise in variance assuming the correlation between the person and establishment effects remains at its interval 1 value—that is, imposing the restriction that $Cov_4(\text{person, establ.}) = \rho_1 Var_4(\text{person})^{1/2} \times Var_4(\text{establ.})^{1/2}$ where the subscript 4 refers to the interval 4 value of the statistic and ρ_1 is the correlation between the person and establishment effects in interval 1. Counterfactual 2 assumes that the variance of establishment effects remains at its interval 1 level. Counterfactual 3 imposes both restrictions.

TABLE IV

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

Notes. See notes to Table II for sample composition. Calculations based on estimated AKM models summarized in Table III. Entry in column (5) is change in variance component from interval 1 to interval 4. Entry in column (6) is ratio of the change in the variance component to the total change in variance of wages reported in first row of table (as a percentage).

*Counterfactual 1 computes the counterfactual rise in variance assuming the correlation between the person and establishment effects remains at its interval 1 value—that is, imposing the restriction that $Cov_4(\text{person, establ.}) = \rho_1 \, Var_4(\text{person})^{1/2} \times Var_4(\text{establ.})^{1/2}$ where the subscript 4 refers to the interval 4 value of the statistic and ρ_1 is the correlation between the person and establishment effects in interval 1. Counterfactual 2 assumes that the variance of establishment effects remains at its interval 1 level. Counterfactual 3 imposes both restrictions.

TABLE IV

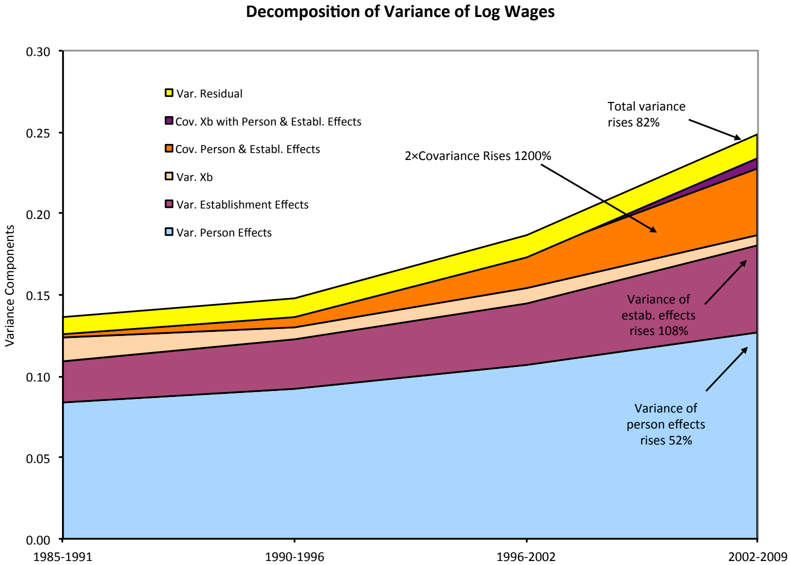
DECOMPOSITION OF THE RISE IN WAGE INEQUALITY

	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

Notes. See notes to Table II for sample composition. Calculations based on estimated AKM models summarized in Table III. Entry in column (5) is change in variance component from interval 1 to interval 4. Entry in column (6) is ratio of the change in the variance component to the total change in variance of wages reported in first row of table (as a percentage).

*Counterfactual 1 computes the counterfactual rise in variance assuming the correlation between the person and establishment effects remains at its interval 1 value—that is, imposing the restriction that $Cov_4(\text{person, establ.}) = \rho_1 \, Var_4(\text{person})^{1/2} \times Var_4(\text{establ.})^{1/2}$ where the subscript 4 refers to the interval 4 value of the statistic and ρ_1 is the correlation between the person and establishment effects in interval 1. Counterfactual 2 assumes that the variance of establishment effects remains at its interval 1 level. Counterfactual 3 imposes both restrictions.

Rising variance of FE and sorting in rising variance of wages



Outline

Education

Comparing changes over time in group averages

- From (1) and (4), the mean wage for workers in group g is:

$$\mathbb{E}_g\left[y_{it}\right] = \mathbb{E}_g\left[\alpha_i\right] + \mathbb{E}_g\left[\psi_{J(i,t)}\right] + \mathbb{E}_g\left[x_{it}'\beta\right] \qquad (7)$$

where $\mathbb{E}_g\left[.\right] \equiv \mathbb{E}\left[.\middle| G_i = g\right]$.

- The change over time in the mean wage differential between groups g_1 and g_2 is given by:

$$\left\{\mathbb{E}_{g_1}\left[y_{i,t}\right] - \mathbb{E}_{g_1}\left[y_{i,t-1}\right]\right\} - \left\{\mathbb{E}_{g_2}\left[y_{i,t}\right] - \mathbb{E}_{g_2}\left[y_{i,t-1}\right]\right\}$$

Increasing return to schooling due to firm FE and sorting

TABLE V

DECOMPOSITION OF CHANGES IN RELATIVE WAGES BY EDUCATION LEVEL, 1985–1991 VERSUS 2002–2009

	(1) Change in mean log wage relative to apprentices	(2) Change in mean person effect	(3) Change in mean establishment effect	(4) Remainder
Highest education qualification				
1. Missing/none	–14.6	1.8	–12.2	–4.2
2. Lower secondary school or less (no vocational training)	–10.5	–0.1	–6.3	–4.1
4. Abitur with or without vocational training*	10.1	0.0	2.6	7.5
5. University or more	5.7	1.5	3.9	0.3

Notes. Wage changes are measured between intervals 1 (1985–1991) and 4 (2002–2009). Remainder (column (4)) represents changing relative contribution of Xb component.

**Abitur* refers to Allgemeine Hochschulreife, a certificate of completion of advanced level high school.

Occupation and Industry

Decomposing the variance of group averages

- From (7), the variance in mean wages across groups is:

$$\begin{aligned} Var(\mathbb{E}_g[y_{it}]) = {} & Var(\mathbb{E}_g[\alpha_i]) \\ & + Var(\mathbb{E}_g\left[\psi_{J(i,t)}\right]) \\ & + Var(\mathbb{E}_g\left[x'_{it}\beta\right]) \\ & + 2Cov(\mathbb{E}_g[\alpha_i]), \mathbb{E}_g\left[\psi_{J(i,t)}\right]) \\ & + 2Cov(\mathbb{E}_g[\alpha_i]), \mathbb{E}_g\left[x'_{it}\beta\right]) \\ & + 2Cov(\mathbb{E}_g\left[\psi_{J(i,t)}\right], \mathbb{E}_g\left[x'_{it}\beta\right]) \end{aligned} \tag{8}$$

- For g occupation or industry, compare this decomposition in different time intervals.

Firms and sorting important in changing variance of premia

TABLE VI

Contribution of Person and Establishment Effects to Wage Variation across Occupations and Industries

	(1) Interval 1 1985–1991	(2) Interval 2 1990–1996	(3) Interval 3 1996–2002	(4) Interval 4 2002–2009	Change in variance (Int. 1 to Int. 4)* (5)	(6) Share
Panel A: Between occupations (342 three-digit occupations)						
Std. dev. of mean log wages	0.233	0.243	0.263	0.289	0.029	100
Std. dev. of mean person effects	0.186	0.203	0.198	0.207	0.008	28
Std. dev. of mean estbl. effects	0.101	0.104	0.124	0.135	0.008	28
Correlation of mean person effects and establ. effects	0.110	0.171	0.238	0.291	0.012	42
Panel B: Between industries (96 two-digit industries)						
Std. dev. of mean log wages	0.173	0.184	0.203	0.224	0.020	100
Std. dev. of mean person effects	0.103	0.114	0.128	0.140	0.009	44
Std. dev. of mean estbl. effects	0.104	0.110	0.108	0.121	0.004	19
Correlation of mean person effects and establ. effects	0.242	0.301	0.422	0.403	0.008	42

Notes. Decompositions based on estimated AKM models summarized in Table III. Occupation is based on main job in each year; establishments are assigned one industry per interval, using consistently-coded two-digit industry.

*Entry in column (5) represents change in variance or covariance component. Entry in column (6) is the share of the total change in variance explained. Shares do not add to 100% because Xb component and its covariances are omitted.

Outline

Std. dev. of firm FE increased due to de-unionisation

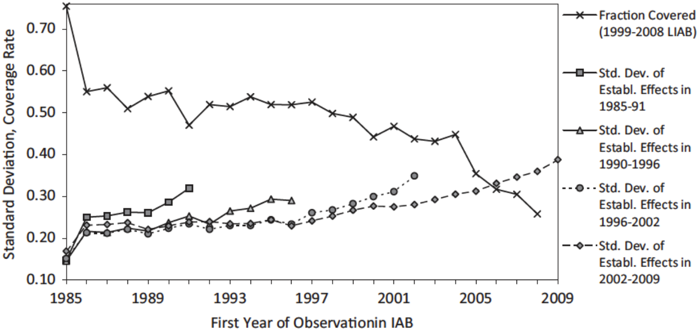


FIGURE IX

Standard Deviation of Establishment Effects and Fraction Covered by Collective Agreements, by Birth Year of Establishment

Figure shows standard deviation of estimated establishment effects in a given observation interval (1985–1991, 1990–1996, 1996–2002, or 2002–2009) for establishments that are present in that interval and first appeared in the IEB data in the “birth year” indicated on the horizontal axis. Figure also shows fraction of establishments in a given birth year surveyed in the 1999–2008 LIAB that are covered by collective agreements.

Std. dev. of firm FE increased due to de-unionisation

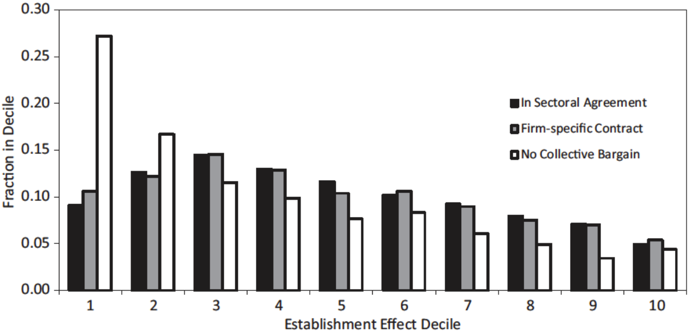


FIGURE X

Distribution of Establishment Effects by Collective Bargaining Status, Based on Establishment Effects for 1996–2002 and Bargaining Status in 2000 Wave of LIAB

Figure shows distribution of collective bargaining coverage status (no collective bargain, covered by firm-specific agreement, or covered by sectoral agreement) for 7,080 establishments in 2000 wave of LIAB that can be linked to IEB data. Establishments are classified into deciles of their estimated establishment effects from AKM model fit to 1996–2002 data.

Outline

Conclusion

- This paper studies trends in the dispersion of firm-specific wage premiums and and sorting.
- Estimate an AKM model using more efficient techniques.
- Rising wage inequality due to rising dispersion in worker- and firm-specific components, and increased sorting.
- Rising dispersion in worker FE explains about 40%, firm FE 25%, and increased sorting about 35% of rising wage inequality in West-Germany between 1985 and 2009.
- This paper has lead to a revival of AKM.

Further readings

1. Card, D., Cardoso, A., and Kline, P. (2016). "Bargaining, sorting, and the gender wage gap: quantifying the impact of firms on the relative pay of women". *Quarterly Journal of Economics*. 131 (2).
2. Goldschmidt, D., and Schmieder, J. (2017). "The rise of domestic outsourcing and the evolution of the German wage structure". *Quarterly Journal of Economics*. 132 (3).
3. Card, D., Cardoso, A., Heining, J., and Kline, P. (2018). "Firms and labor market inequality: evidence and some theory". *Journal of Labor Economics*. 36 (S1).
4. Song, J., Price. D.J., Guvenen, F., Bloom, N., and T. Von Wachter (2019). "Firming up inequality." *Quarterly Journal of Economics*. 134(1).
5. Bonhomme, S., Holzeu, K., Lamadon, T., Manresa, E., Mogstad, M., and B. Setzler (2020). "How much should we trust estimates of firm effects and worker sorting?". *NBER Working Paper* 27368. June 2020.

Coding

- Packages have been written for Stata:
  - **felsdvreg**
  - **regxfe**
  - **a2group** + **a2reg**
- Packages and simulation in R: https://floswald.github.io/ScPo-Labor/lab-akm.html
- Card, Heining, and Kline (2013) use matlab - see Pat Kline's webpage for code.
- Bonhomme et al. (2020) use Python - see Magne Mogstad's webpage for code.