Industry Wage Differentials: A Firm-Based Approach

David Card
UC Berkeley

Jesse Rothstein
UC Berkeley

Moises Yi
Census Bureau

March 2023

Disclaimer: All results in derived from non-public Census data have been reviewed to ensure that no confidential information is disclosed (DRB clearance number CBDRB-FY22-422). Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau.

Industry wage differentials

- An old question in labor economics: Do wages differ systematically across industries?
  - Would the same worker earn more in the finance industry than in the hospitality industry?
  - Or do observed wage differences reflect sorting of different types of workers into different industries?

- Classical, competitive labor economics models assume the “law of one price.”

- A classic literature, dating back at least to Krueger and Summers (1988), explores systematic pay differences across industries, largely using survey data.

Approaches to estimating industry wage differentials

- The simplest method: Compare average wages across industries.
- Slightly less simple: Regress wages or earnings on industry indicators, with controls for worker education, age, etc.

  → *Cross-sectional* analysis

Cross-sectional estimates:

Regress wages or earnings on industry indicators, controlling for education and other factors that might influence earnings.

TABLE I

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

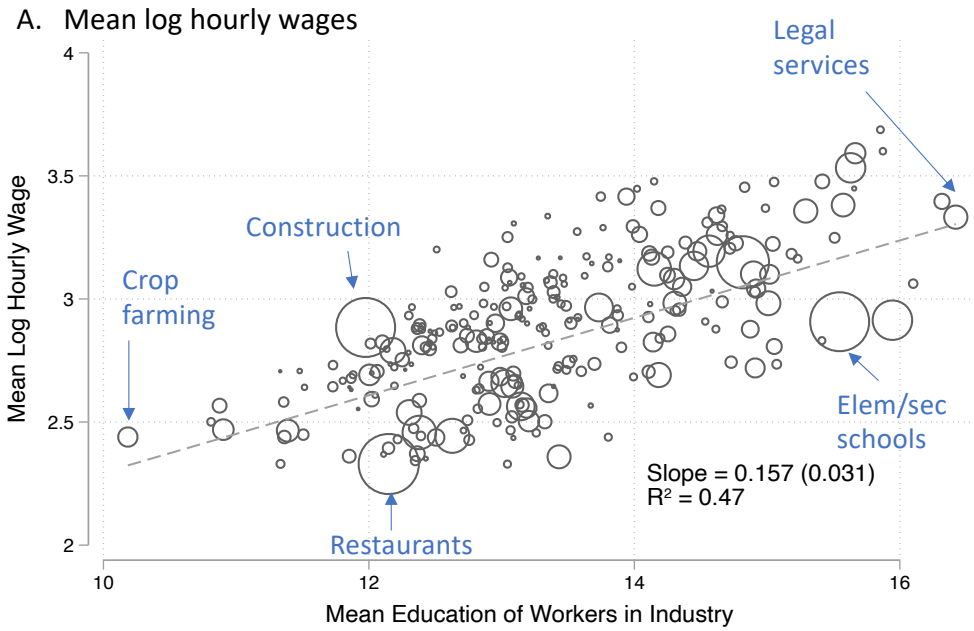
Industry	(1) 1974	(2) 1979	(3) 1984	(4) 1984 Total Compensation
Construction	.195	.126	.108	.091
	(.021)	(.031)	(.034)	(.035)
Manufacturing	.055	.044	.091	.131
	(.020)	(.029)	(.032)	(.032)
Transportation & Public Utilities	.111	.081	.145	.203
	(.021)	(.031)	(.034)	(.034)
Wholesale & Retail Trade	−.128	−.082	−.111	−.136
	(.020)	(.030)	(.033)	(.033)
Finance, Insurance and	.047	−.010	.055	.069
Real Estate	(.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.

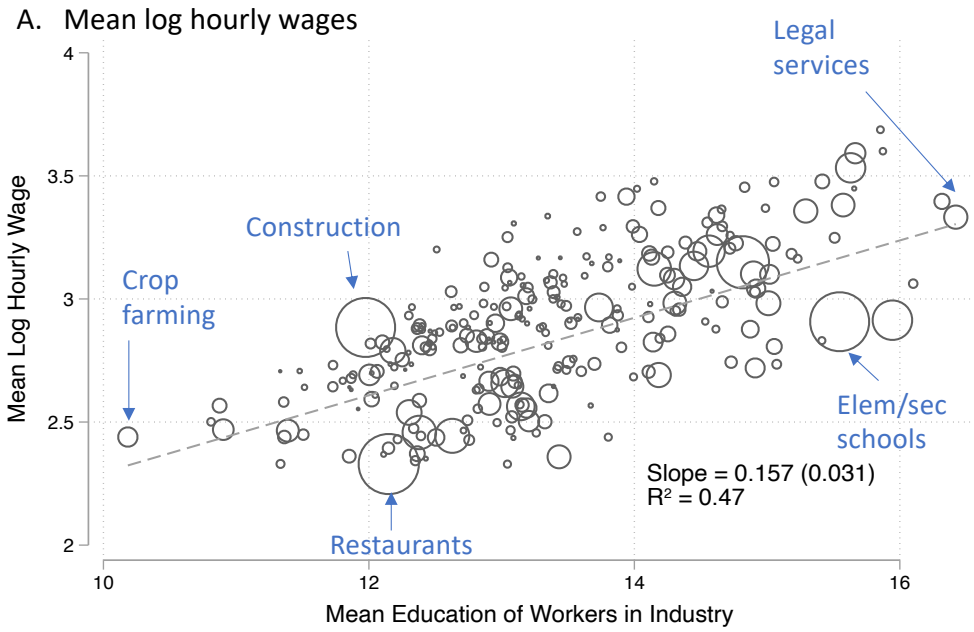
[b] Weights are employment shares for each year.

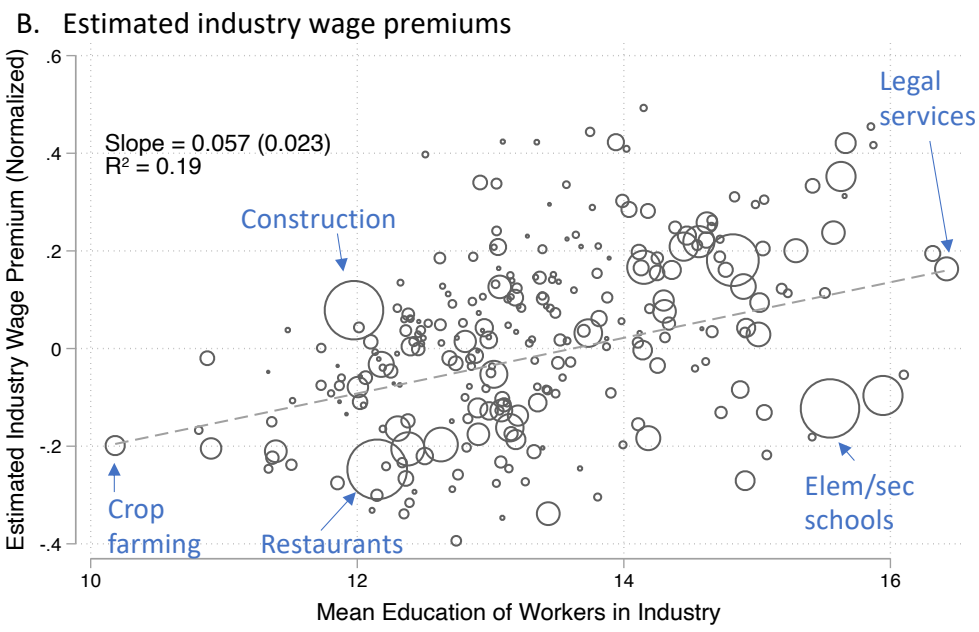
** F test that industry wage differentials jointly equal 0 rejects at the .000001 level.

Cross-sectional industry (4-digit) wage differentials estimated from American Community Survey data, pooled 2010-2018.



Cross-sectional industry (4-digit) wage differentials estimated from American Community Survey data, pooled 2010-2018.





The “movers” design:

Track workers as they move from one industry to another.

Are earnings systematically higher in some industries for the same workers?

Implemented as a regression with worker fixed effects – constant additive factors capturing observed and unobserved components of workers’ permanent skill.

The "movers" design:

Track workers as they move from one industry to another.

Are earnings systematically higher in some industries for the same workers?

Implemented as a regression with worker fixed effects – constant additive factors capturing observed and unobserved components of workers' permanent skill.

Similar results as cross-sectional method.

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.

The LEHD as an opportunity

- U.S. evidence to date largely relies on data from repeated surveys, with small samples and substantial measurement error.
- LEHD provides enormous samples and good industry measures.
  - Industry comes from the establishment, which is imputed for workers at multi-establishment firms, but in most cases the industry can be imputed with high confidence.
- Our project: Use “movers” in the LEHD to construct better measures of industry wage differentials.
  - Large samples mean we can estimate effects for 4-digit NAICS industries.
  - Control flexibly for location to avoid confounding industry and location effects.

Establishment wage premiums

- Workers are nested in establishments which are nested in industries.
- An enormous literature beginning with Abowd, Kramarz, and Margolies (1999), estimates *establishment* wage premiums using a similar movers design.
  - *log wages = permanent worker skill* < as in industry movers design >
    + establishment effects < common to all workers >
  - Assumption: A worker who moves from Walmart to Target sees the same gain as the loss seen by a worker moving from Target to Walmart.
  - Model works surprisingly well across many settings.
  - Substantial variation across establishments or firms.
  - More profitable firms pay higher premiums.
  - Little further evidence about the structure of these premiums.
  - Few estimates to date from U.S. data - AKM was set in France.

Industries or establishments?

- What is the industry wage premium when establishment premia vary?
- A natural definition: The average establishment premium across all establishments in an industry.
- If a randomly selected worker in industry A is moved to a randomly selected establishment in industry B, how much will earnings change?
- This is is subtly different from what the movers design estimates.

Hierarchy bias: Intuition

- A central question in the AKM literature: Do high-wage-premium establishments employ higher-skill workers?
  - Methodological challenges, but evidence suggests yes.
  - This suggests a matching process of worker skill to establishment pay policy.
- Given matching, we might expect workers to move among firms with similar wage policies.
  - Moves from finance to hospitality likely are from below-average finance firms to above-average hospitality firms.
  - Movers design estimates contrasts among firms where movers work, not among average firms in each industry.
  - Result is that industry differences are attenuated.

Sample construction

- We use data on all 50 states, 2010-2018Q2.
- To approximate full-time, full-quarter earnings, we exclude:
  - Quarters with earnings below $3800
  - Worker-quarters with more than one employer
  - First and last quarters (*transitional* quarters) of job spells.
- We further limit to:
  - Ages 22-62
  - At least 8 quarters of observed employment
  - (In some analyses) matched to ACS 2001-2017 with education information
- We use the first imputation of the worker's establishment. Results are not systematically different if sample is limited to observations without uncertainty about the CZ and industry.

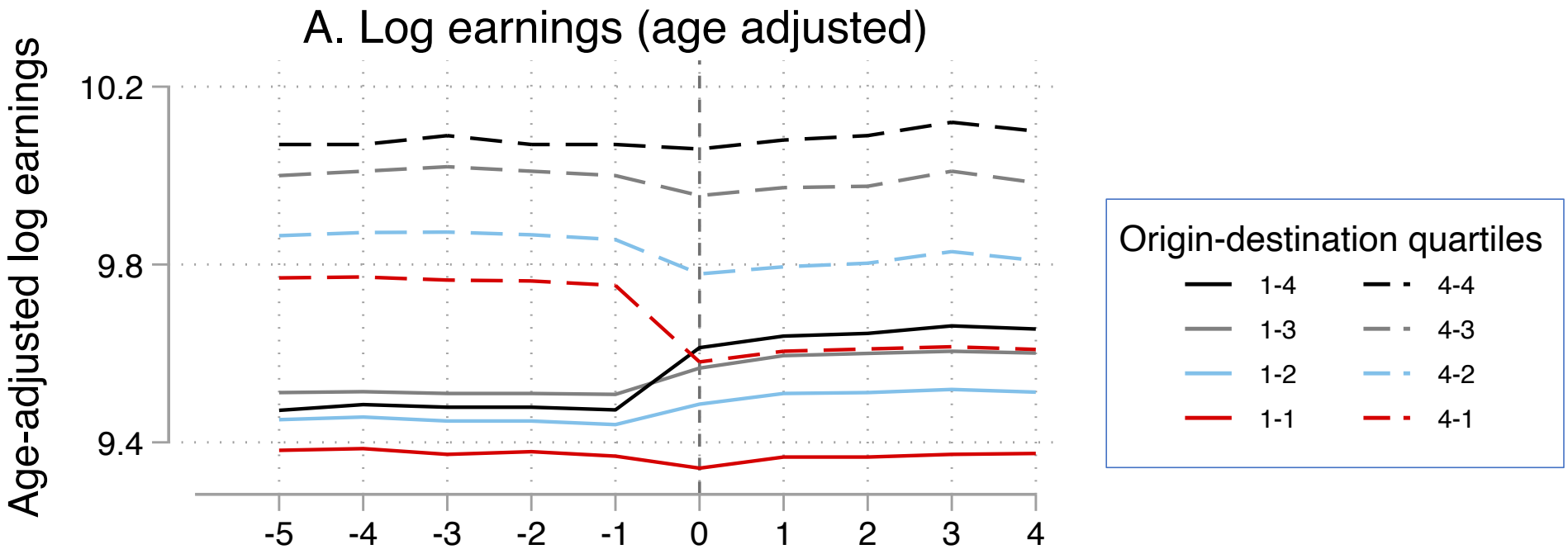
Summary statistics

	Industry stayers	Industry switchers
	(2)	(3)
Quarterly earnings	16,050	14,630
	(19,710)	(14,860)
Age	44	40
	(11)	(10)
Female	0.48	0.46
Foreign born	0.16	0.16
Number of CZs in which observed		
1	0.82	0.72
2	0.14	0.22
3+	0.03	0.05
Number of industry switches (within CZs)		
0	1.00	0.00
1	0.00	0.68
2+	0.00	0.32
Quarters observed	27.2	23.7
	(7.8)	(7.0)
Number of person-quarter observations (millions)	1,544	960.4
Number of unique people (millions)	65.7	46.1

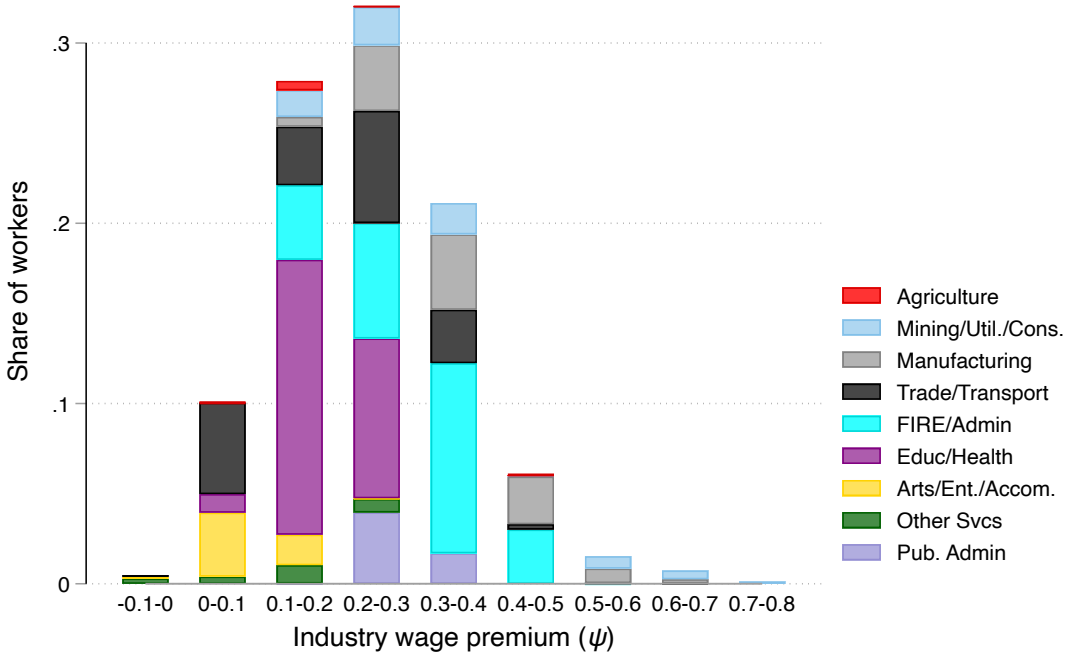
Estimation procedure

- Estimate AKM model:
  - *log quarterly earnings = person effect*
    + establishment premium
    + controls for calendar quarter, age quadratic
  - We normalize the average establishment premium in the restaurant industry to zero.
  - We estimate the model separately by commuting zone – thus, geographic differences (in restaurant pay) are removed.
- Average the estimated establishment effects to the industry level to obtain industry wage premium.
- Note: We can measure the establishment's position in the industry hierarchy – the difference between the establishment effect and the industry average.

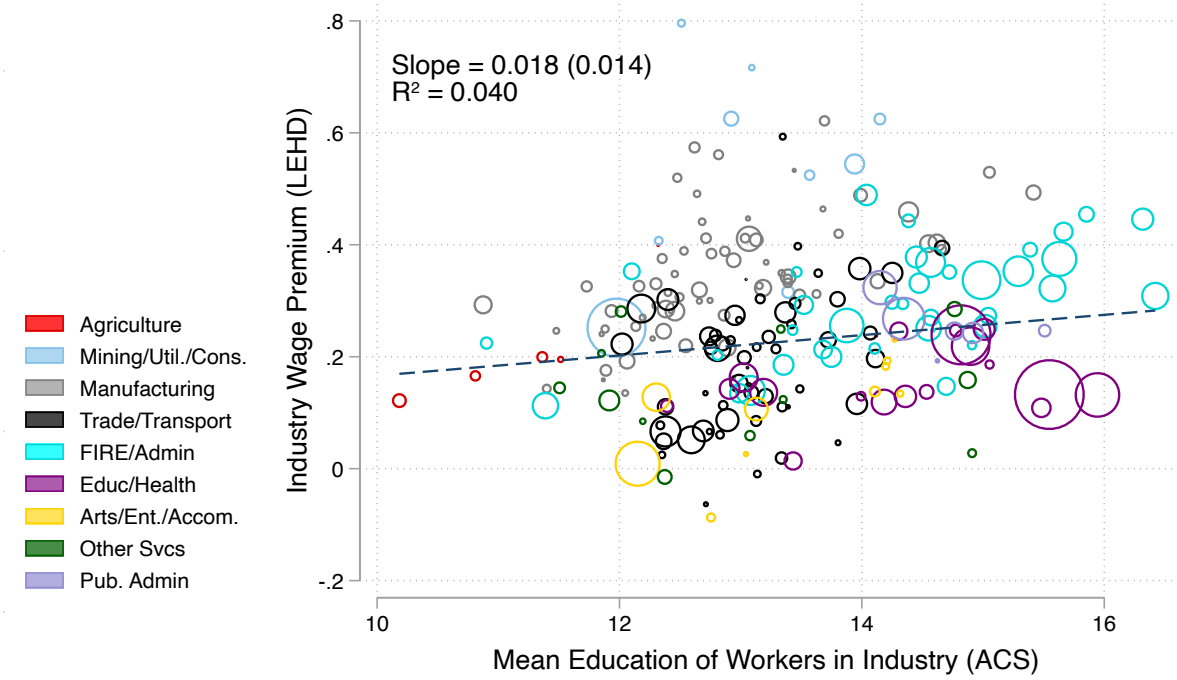
Event study of between-industry movers



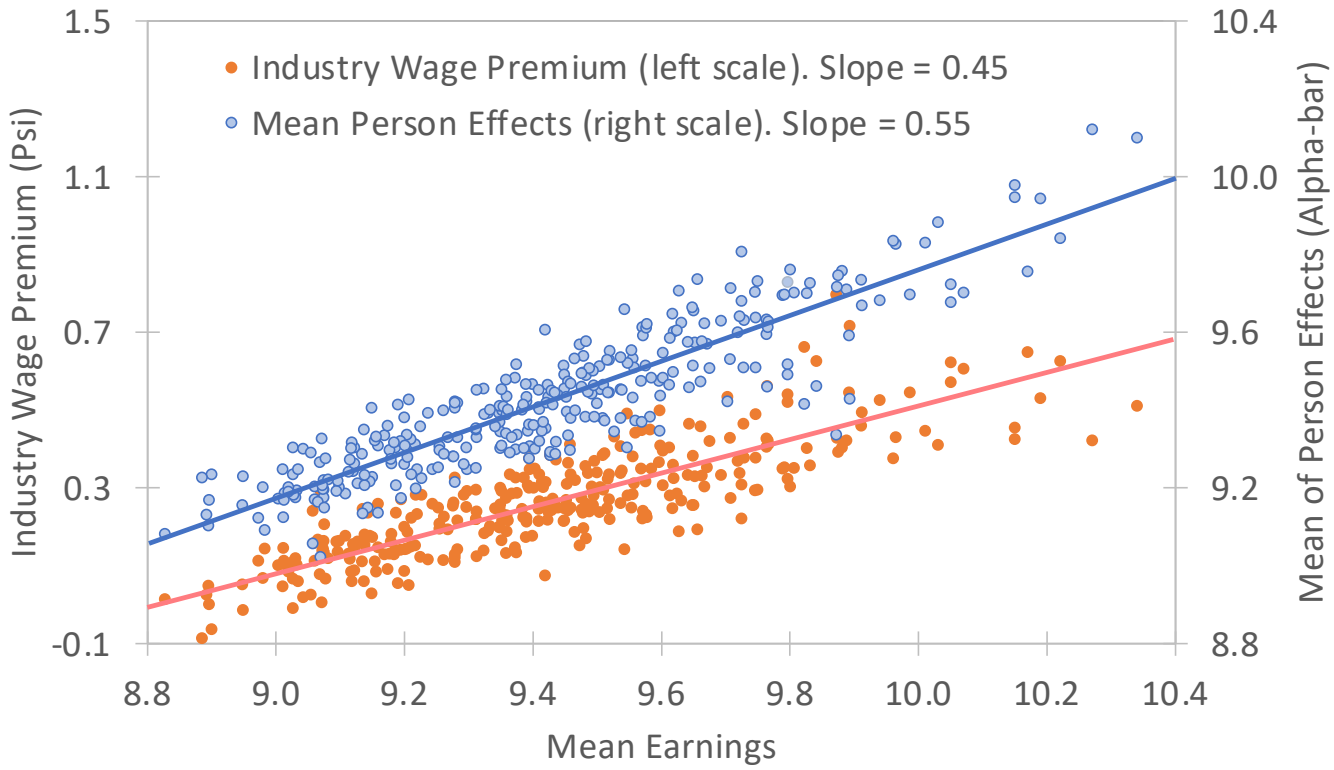
Distribution of 4-digit industry wage premiums



Distribution of 4-digit industry wage premiums

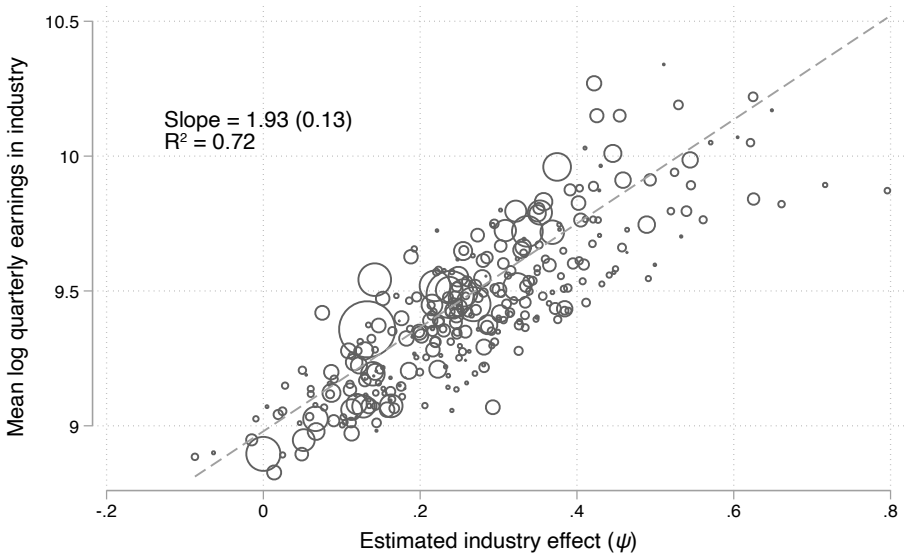


Worker skill explains 55% of industry wage differences. Premia explain 45%.

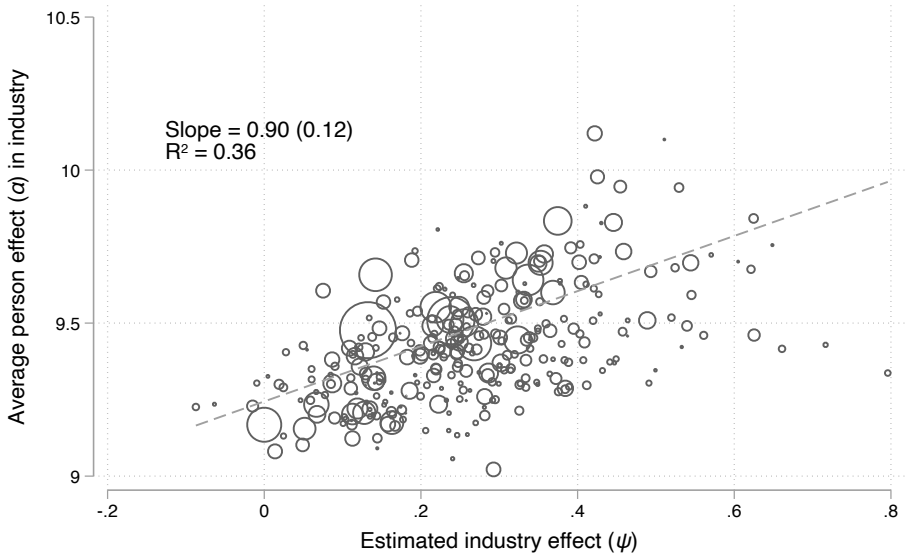


Worker sorting and industry premia

Mean earnings ($\bar{y}_j$) vs. industry premia (ψ_j)

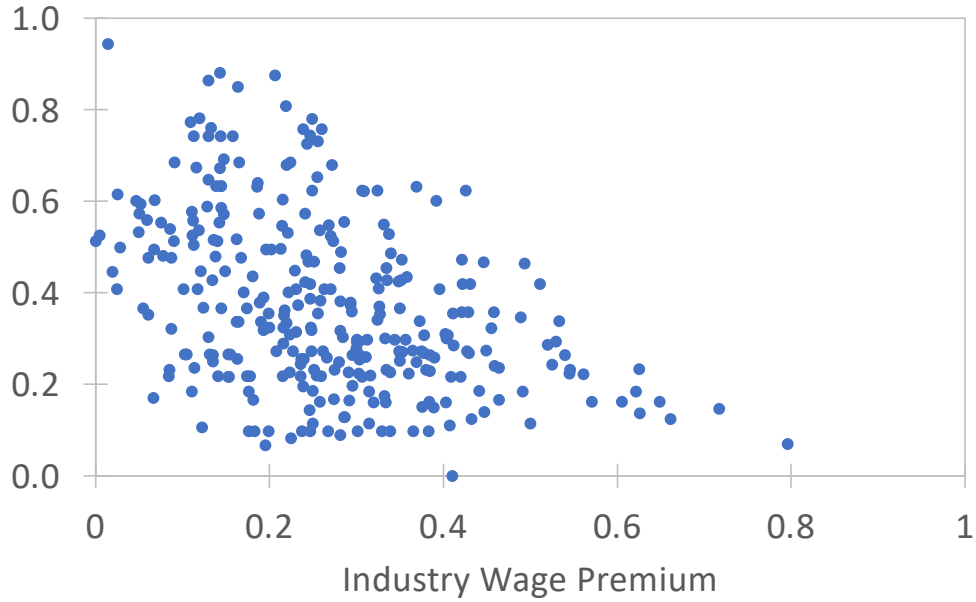


Mean person effects ($\bar{\alpha}_j$) vs. industry premia (ψ_j)

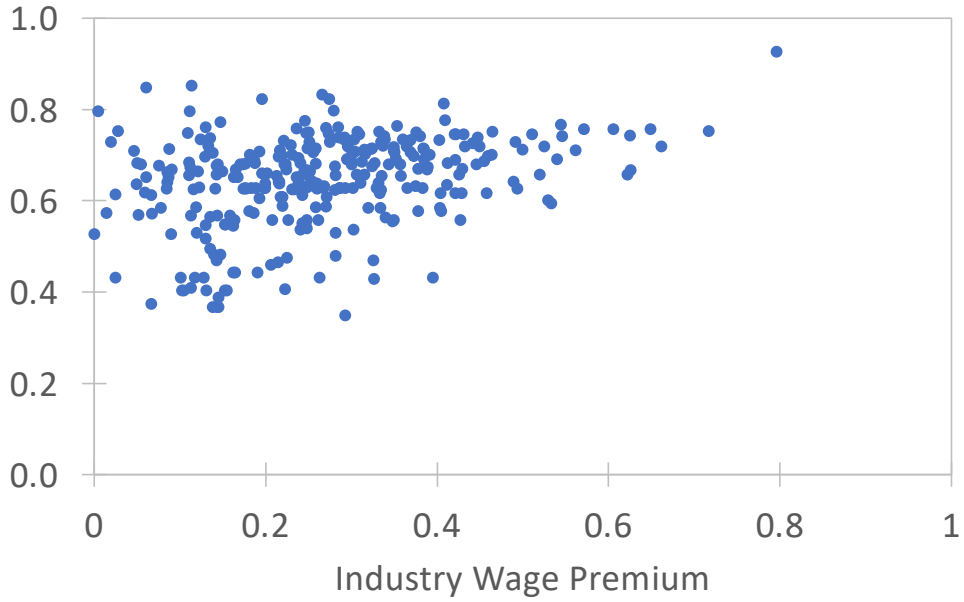


Worker characteristics and industry premia

a. Share of Female Workers

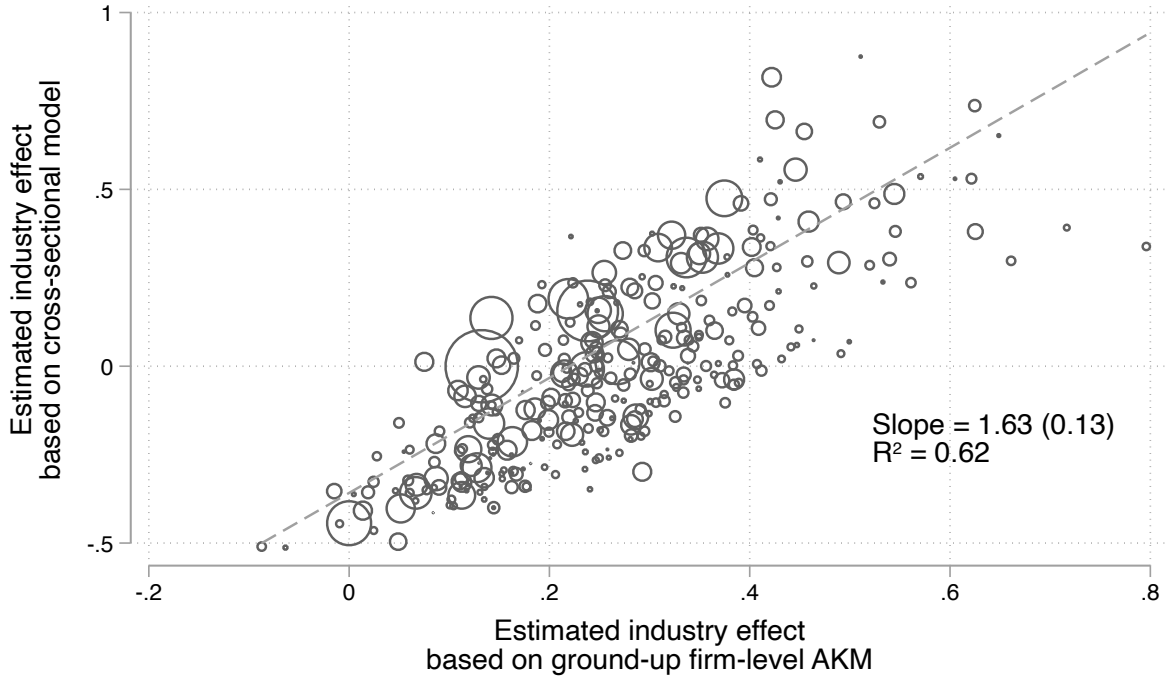


b. Share of White Non-Hispanic Workers



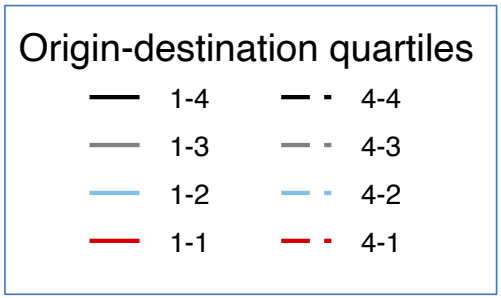
Cross-sectional estimates dramatically overstate ground-up, AKM-based industry premia

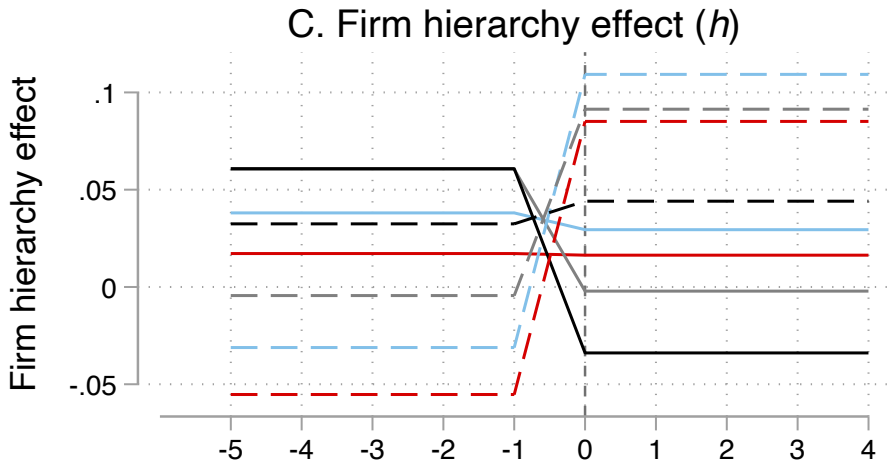
Figure 8. Comparing cross-sectional and AKM-based estimates of industry premia



Between-industry moves are selective in terms of the origin and destination firms.

- Recall that we can measure the deviation of the establishment premium from the industry mean.
- How does this change when people switch industries?
- Hypothesis: People moving from low-premium to high-premium industries should tend to come from good firms in the former and wind up at bad firms in the latter.
- This is what we see!

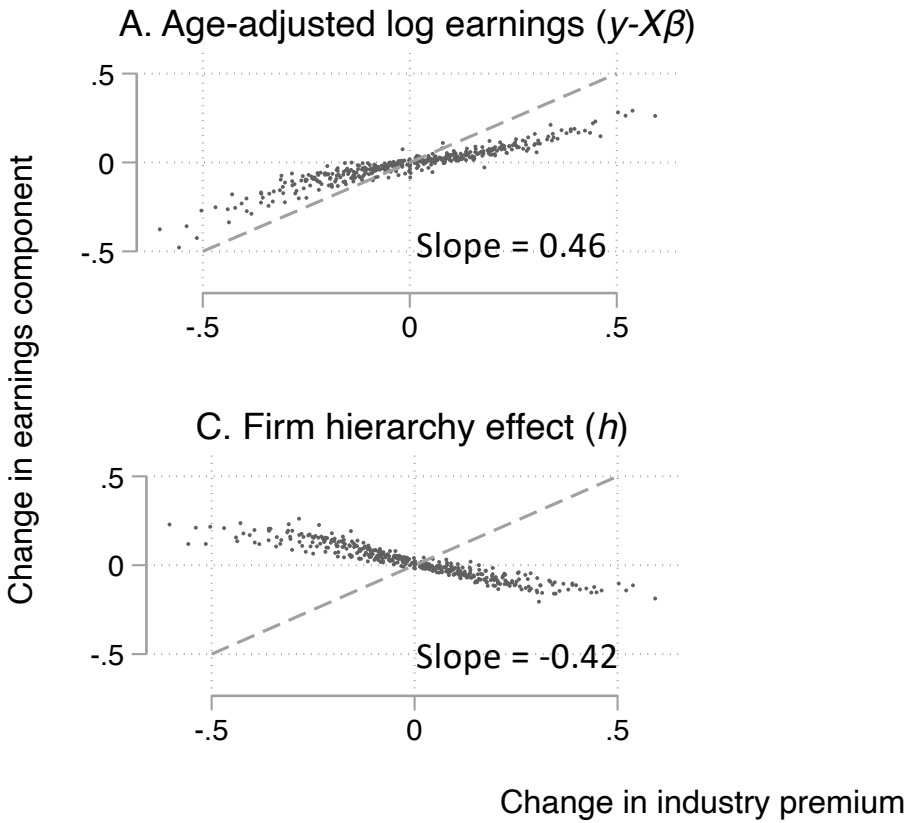




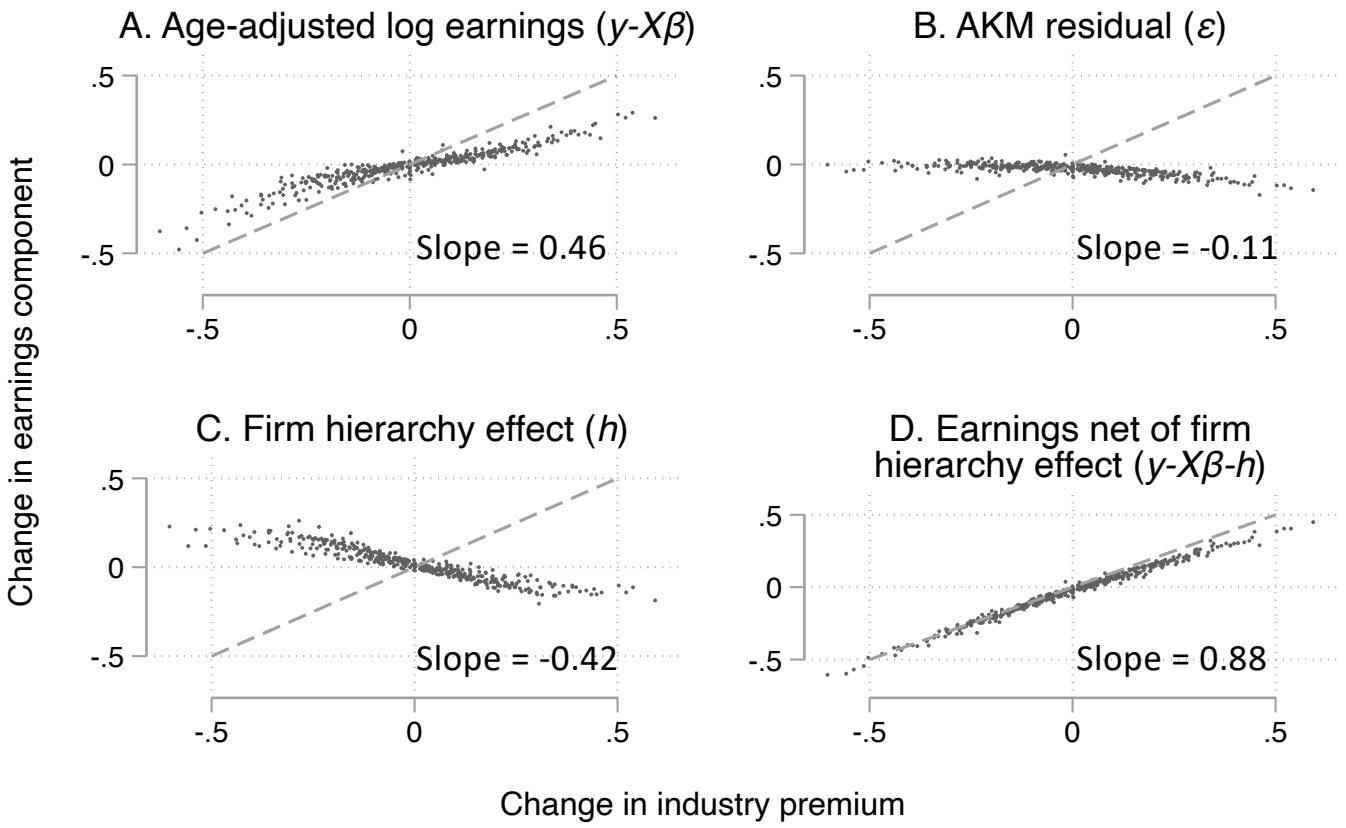
Hierarchy effects revisited

- When people move from a low-wage industry to a high-wage industry, they tend to go from an above-average firm in the origin industry to a below-average firm in the destination industry.
- As a result, earnings rise less than the difference in industry effects.
- The pattern is reversed for downward moves – earnings decline less than the industry effects imply.
- This “hierarchy” change is an omitted variable in the industry movers design, and attenuates estimated industry differences.

Earnings don’t change as expected due to hierarchy term



Once hierarchy term is removed, short-run changes are close to predictions



Movers estimates *understate* ground-up industry premia

Table 4. Comparisons of industry effects from alternative models

	Preferred model	**Cross-sectional models**			**Movers models**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Alternative model controls for:							
Time-varying controls		X	X	X	X	X	X
Time invariant controls			X	X			
CZ FEs				X		X	
Industry-by-CZ FEs							X
Individual FEs					X	X	X
Standard deviation of industry effects	0.122	0.271	0.254	0.240	0.079	0.079	0.082
Regression of alternative model estimates	1.00	1.86	1.63	1.61	0.62	0.62	0.66
on preferred model estimates (N=311)		(0.12)	(0.13)	(0.12)	(0.02)	(0.02)	(0.01)
R^2 (adj)		0.707	0.614	0.672	0.929	0.924	0.954

Workers climb the hierarchy (a bit) with experience

Table 3. Worker experience and the industry hierarchy effect

	Young workers		Older workers	
	(1)	**(2)**	**(3)**	**(4)**
Number of quarters in industry/10	0.012	0.010	0.007	0.006
	(0.002)	(0.001)	(0.001)	(0.001)
(Number of quarters in industry/10)^2	-0.0033	-0.0029	-0.0016	-0.0016
	(0.0006)	(0.0003)	(0.0003)	(0.0002)
Controls for worker, CZ, industry, time FEs	N	Y	N	Y
N (millions of person-quarter observations)	89.8	89.8	421.8	421.8
R2 (adj.)	0.0004	0.7340	0.0002	0.8370
Experience (in quarters) at which slope=0	18.1	17.2	21.8	18.3
Cumulative effect of 5 years of experience	0.011	0.008	0.008	0.005

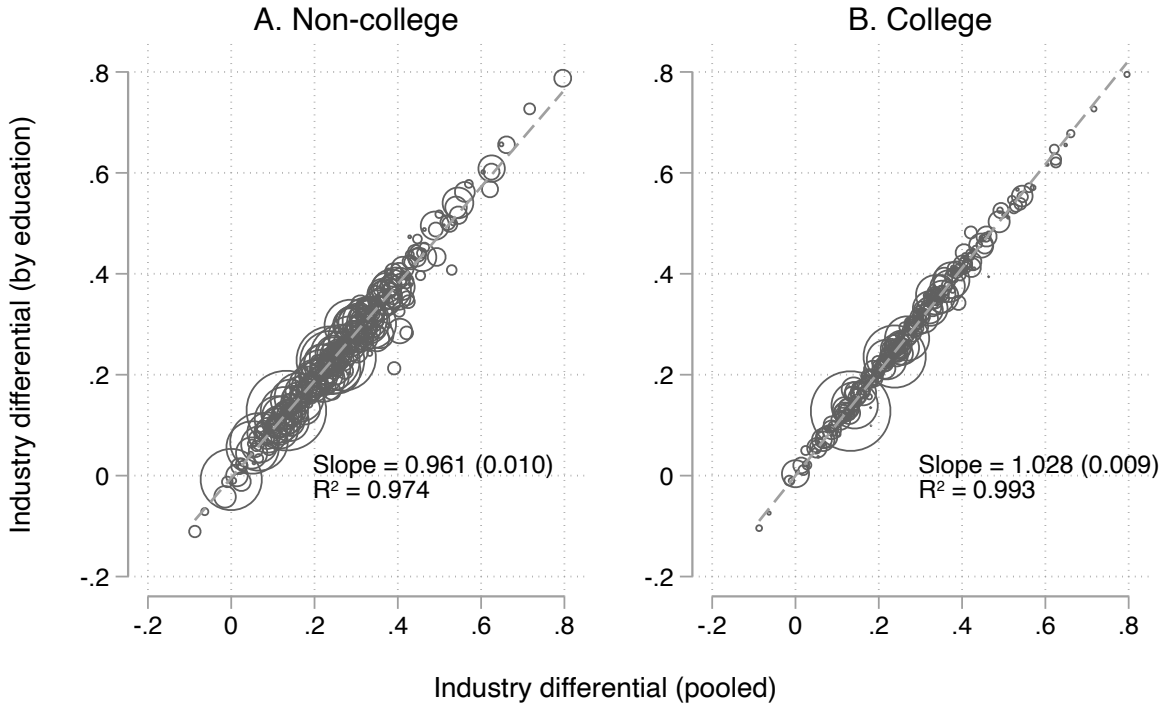
Do workers sort by education within industry?

We construct two alternative measures of the industry premium, by averaging separately over the establishments where college & non-college workers work.

These give very similar estimates.

→ College workers don't systematically sort to higher-premium establishments within industries.

Figure 9. Pooled vs. separate estimates of industry premiums by education



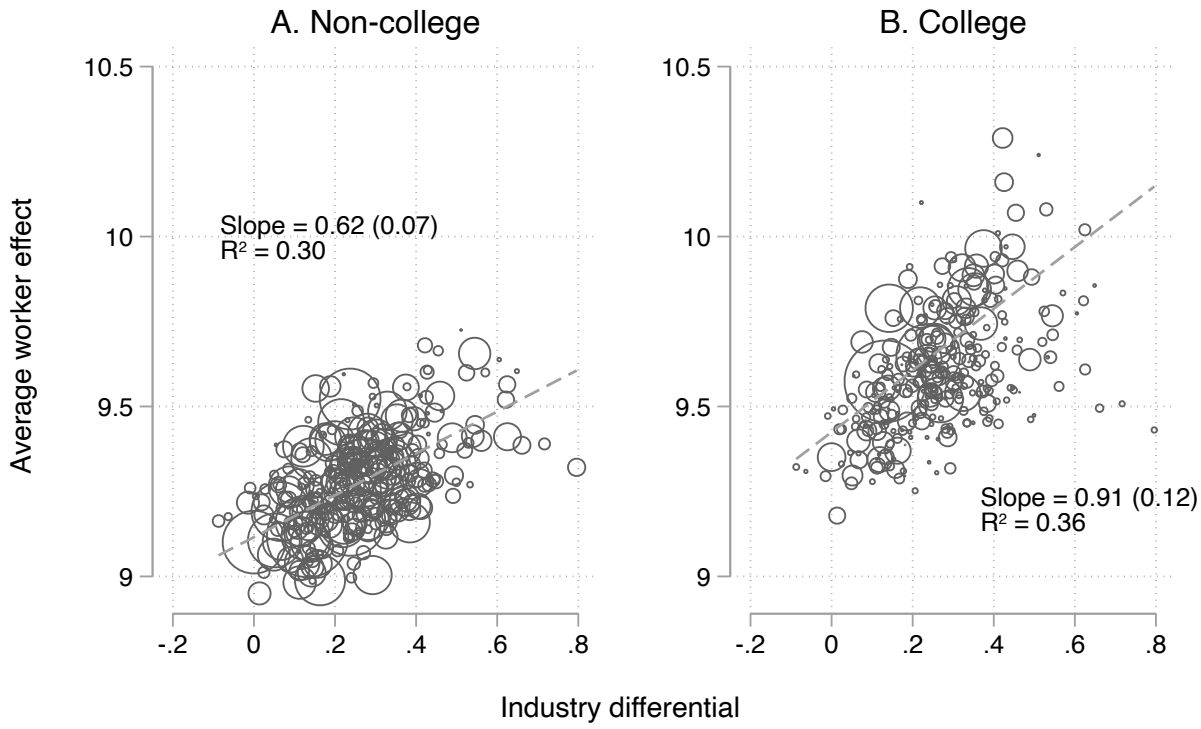
Do workers sort by education and skill across industries?

We do see that college workers in each industry have higher individual pay components than non-college workers.

This is especially true in high-premium industries.

→ College workers are more systematically sorted across industries than non-college workers.

Figure 10. Average worker effects by education and industry



What is the role of geography?

- In a companion paper (Card, Rothstein, and Yi 2022), we study geographic differences in earnings.
- Here, we can construct separate industry premiums for each CZ. In which CZs are the premiums larger than average, and which smaller?
- Summary:
  - No difference in industry premiums by CZ size.
  - CZs with more employment in high-premium industries have higher premiums for those industries.
  - CZs with more employment in high-skill industries also have higher premiums for high-pay industries.
  - No relationship of CZ premiums to CZ unionization rate or minimum wage.

Conclusion

- Modern firm-based methods indicate substantial variation in firm effects across industries, not explained by worker sorting.
  - Standard deviation of industry wage premia is 0.12.
  - Higher premia in resource-based industries; lower in hospitality, education, health.
  - Premia very similar for college & non-college workers.
- Comparison to earlier methods:
  - Cross-sectional estimates overstate premia due to worker sorting.
  - Movers estimates understate premia due to hierarchy term.
- Clear evidence against law of one price in labor markets.
- We couldn't have learned this without the LEHD!
  - Value in further exploration of structure of firm wage differentials.