Firm Wage Effects

741 Macroeconomics
Topic 1

Masao Fukui

2025 fall

Course Logistics

- Lecture:
  - MonWed, 8:30-9:45am in SSW 315
- Instructor:
  - Masao Fukui (mfukui@bu.edu)
  - Office hours: MonTue 4:15-5:45pm in Room 400 (my office)
- Grades: 100% problem sets
  - The first problem is already posted, due Nov 14
- Optional: research proposal (I will read seriously and give feedback)

Is Labor Market Competitive?

- Competitive labor market $\Rightarrow$ firms take wages as given
- This implies…
  1. Identical workers are paid the same wages
  2. Firms do not have wage-setting power
- The competitive labor market paradigm was dominant for many decades

The Course Objective

- In the data, (seemingly) identical workers paid differently depending on employers
  - static: different employers pay differently
  - dynamic: same employer pays differently over time
- Use theoretical models to understand the role of firms in wage determination
  - Why do different firms pay different wages?
  - Why does the same firm pay different wages to the worker?
  - Do firms have wage-setting power?
  - How does the firm distribution translate into wage distribution?

Computation

- Problem sets put emphasis on coding for two reasons:
  1. Coding skills are extremely important:

     Hard to write qualitative papers now, quantification is almost always necessary
  2. It forces you to understand the model:

     If you can't write code to solve the model, you don't understand the model

- Exploit new tools and technologies (it's your comparative advantage!):
  1. Frontier computational methodogies (will cover them if time permits)
  2. AI tools (Codex CLI, Cursor, GitHub Copilot, etc)
  - But don't confuse the goals with the means
  - Your goal is to learn, not to copy&paste AI outputs to get good grades

AKM Model

– Abowd Kramarz & Margolis (1999)

AKM Model

- Consider the following statistical model by Abowd, Kramarz, and Margolis (1999):

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

  - w_{it}: log wage of worker i at time t
  - $j(i,t)$: firm employing worker i at time t
  - ψ_j: wage premium of firm j

- Assume $\mathbb{E}[\epsilon_{it} | j(i,s) = j] = 0$ for all i, t, s, j. This embeds:
  1. Worker's mobility decisions are not driven by time-varying wage fluctuations
  2. log wages are additively separable between worker- and firm-components

- Then, worker's movements across firms identify ψ_j (up to a constant):

$$\mathbb{E}[w_{it'} - w_{it} | j(i,t') = j, j(i,t) = k] = \psi_j - \psi_k$$

How Important is Firm FE?

- Null hypothesis: $\psi_j = \psi$ for all j under the competitive labor market
- This has been firmly rejected by many studies for various countries
- But how important is the firm FE in overall wage inequality?
- Variance decomposition:

$$\text{Var}(w_{it}) = \text{Var}(\alpha_i) + \text{Var}(\psi_j) + \text{Var}(\epsilon_{it}) + \text{Cov}(\alpha_i, \psi_j)$$

- Song, Price, Guvenen, Bloom, and Wachter (2019) implementation:
  - Use tax data covering the universe of workers & firms in the US 1978-2013

Firm Wage and Wage Inequality (US)

		Interval 1 (1980–1986)		Interval 2 (1987–1993)		Interval 3 (1994–2000)		Interval 4 (2001–2007)		Interval 5 (2007–2013)		Change from 1 to 5	
		Comp. (1)	Share (2)	Comp. (3)	Share (4)	Comp. (5)	Share (6)	Comp. (7)	Share (8)	Comp. (9)	Share (10)	Comp. (11)	Share (12)
Total variance	Var(y)	0.708	—	0.776	—	0.828	—	0.884	—	0.924	—	0.216	—
Components of variance	Var(WFE)	0.330	46.6	0.375	48.3	0.422	51.0	0.452	51.2	0.476	51.5	0.146	67.6
	Var(FFE)	0.084	11.9	0.075	9.7	0.067	8.1	0.075	8.5	0.081	8.7	−0.003	−1.6
	Var(Xb)	0.055	7.8	0.065	8.4	0.079	9.5	0.061	6.9	0.059	6.4	0.004	1.8
	Var(ϵ)	0.154	21.7	0.148	19.1	0.146	17.6	0.149	16.8	0.136	14.7	−0.018	−8.2
	2*Cov(WFE, FFE)	0.033	4.7	0.057	7.3	0.076	9.2	0.094	10.6	0.108	11.7	0.075	34.8

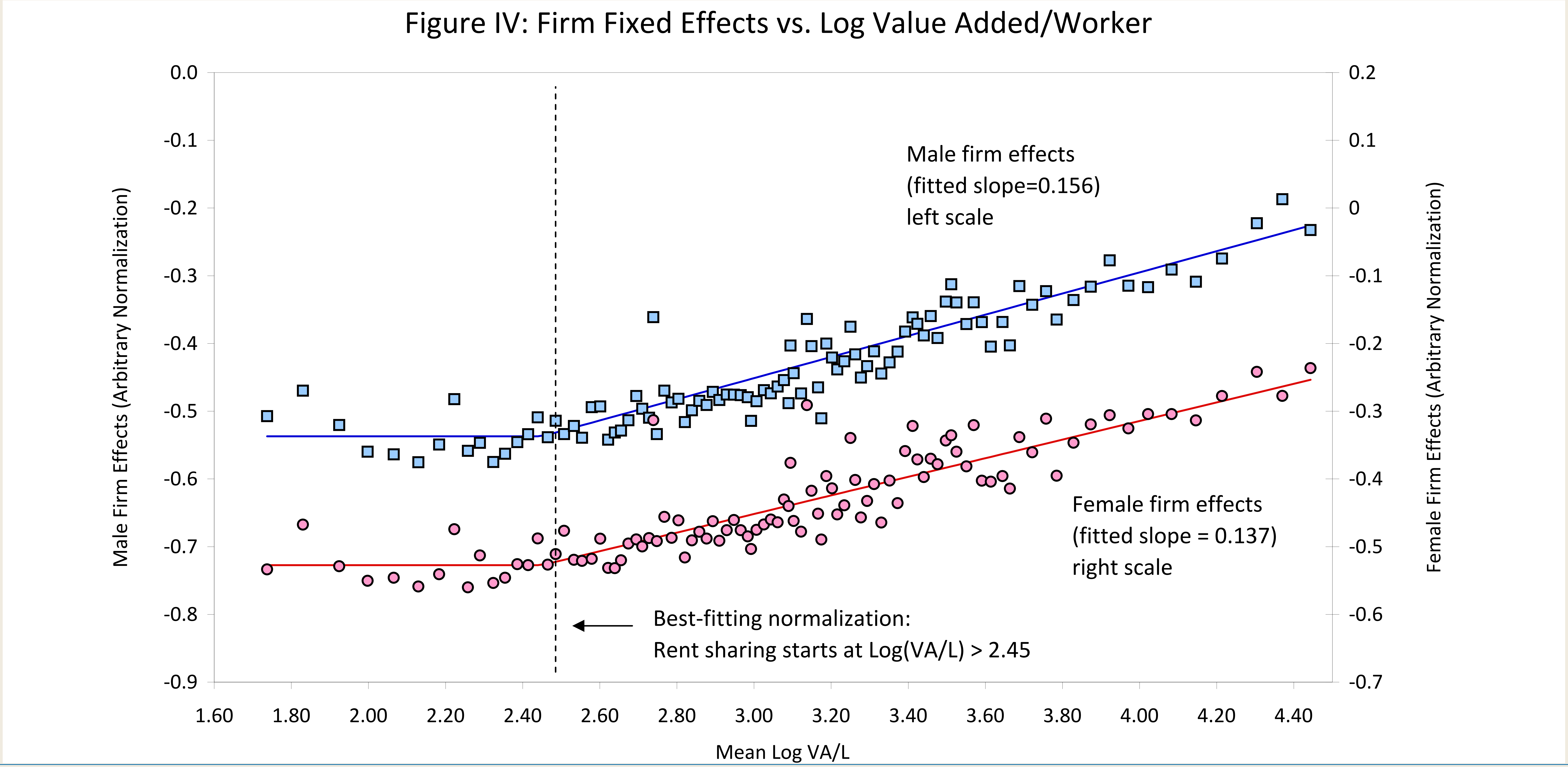
Source: Song, Price, Guvenen, Bloom, and Wachter (2019)

- Firm FE accounts for 8-12% of wage inequality, stable over time
- Rising inequality due to
  1. A rise in the variance of worker FE
  2. A rise in the cavariance between worker FE and firm FE

What Drives the Firm Wage Effect?

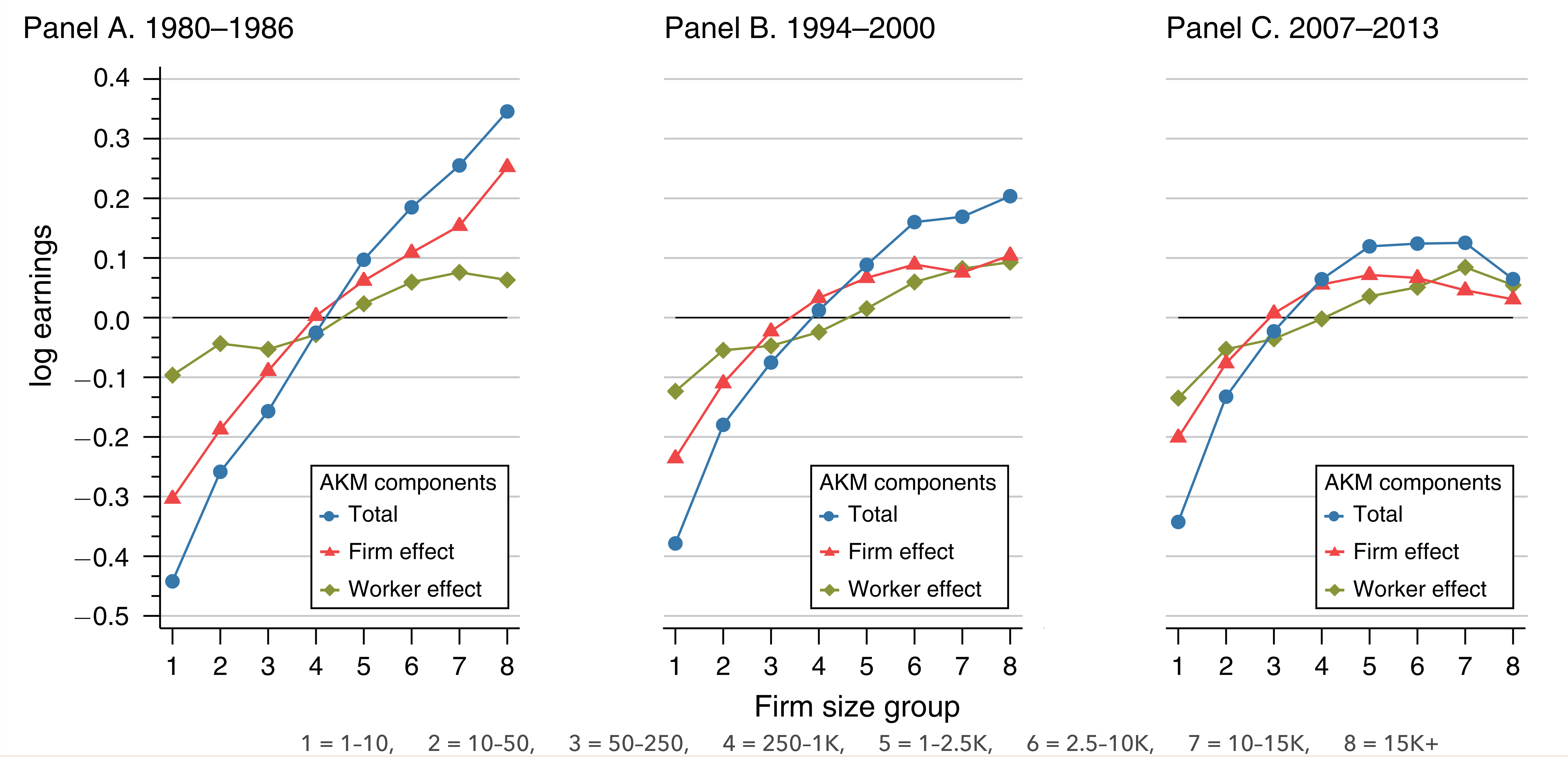
- Why do some firms pay more than others to (seemingly) identical workers?
- Variation in ψ_j is often found to be systematic
- High-wage firms tend to
  - have higher value added (e.g., Card, Cardoso & Kline, 2016)
  - be larger (e.g., Bloom, Guvenen, Smith, Song & von Wachter, 2018)
  - be preferred by workers (e.g., Sorkin, 2018)
  - provide better amenities (e.g., Sockin, 2024)
  - outsource workers in food services, cleaning, security, and logistics occupations (Goldschmidt and Schmieder, 2017)

Higher Firm Wage, Higher Value Added (Portugal)



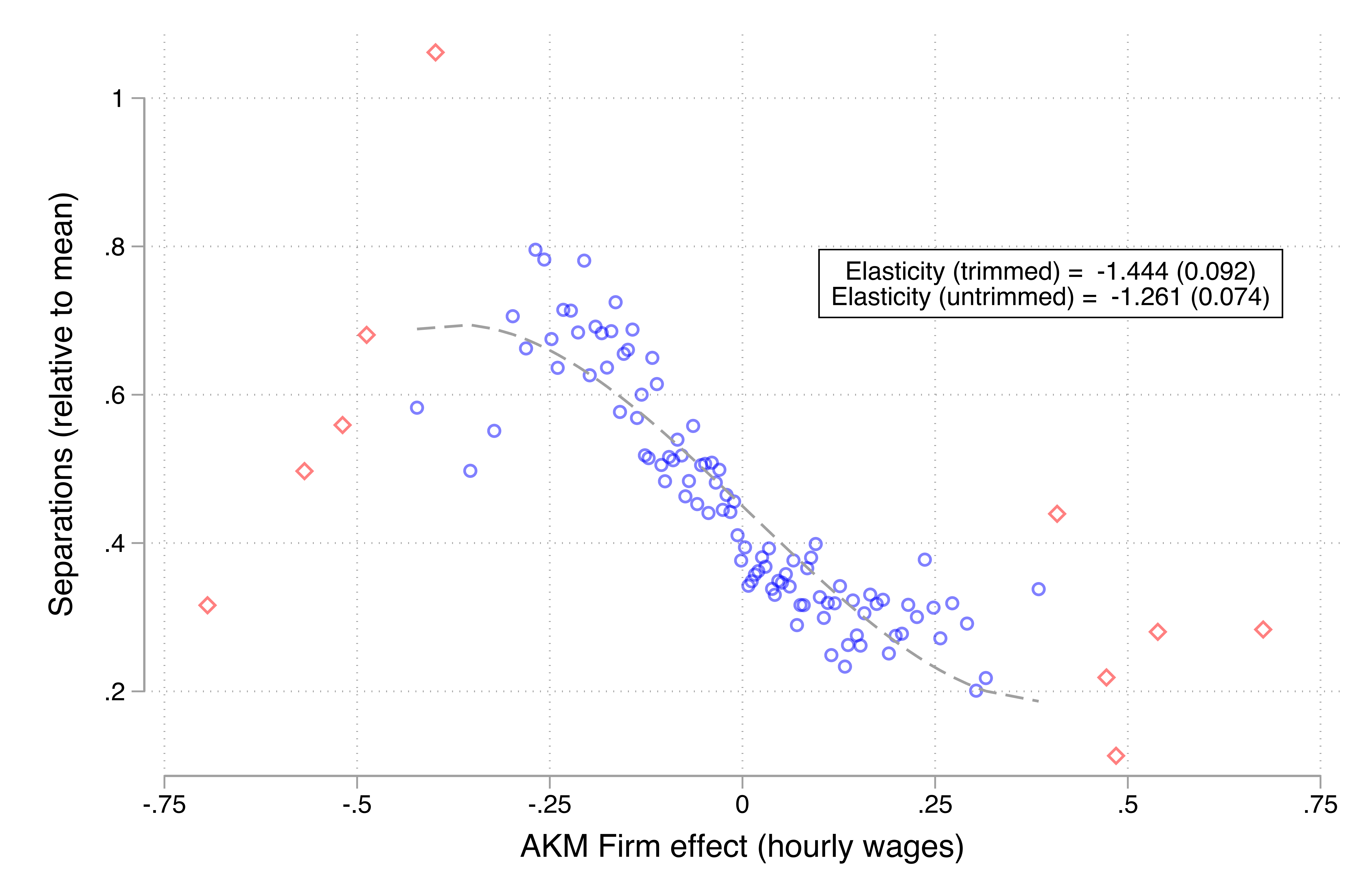
Source: Card, Cardoso, & Kline (2016).

Higher Firm Wage, Larger Firm Size? (US)



Source: Bloom, Guvenen, Smith, Song & von Wachter (2018).

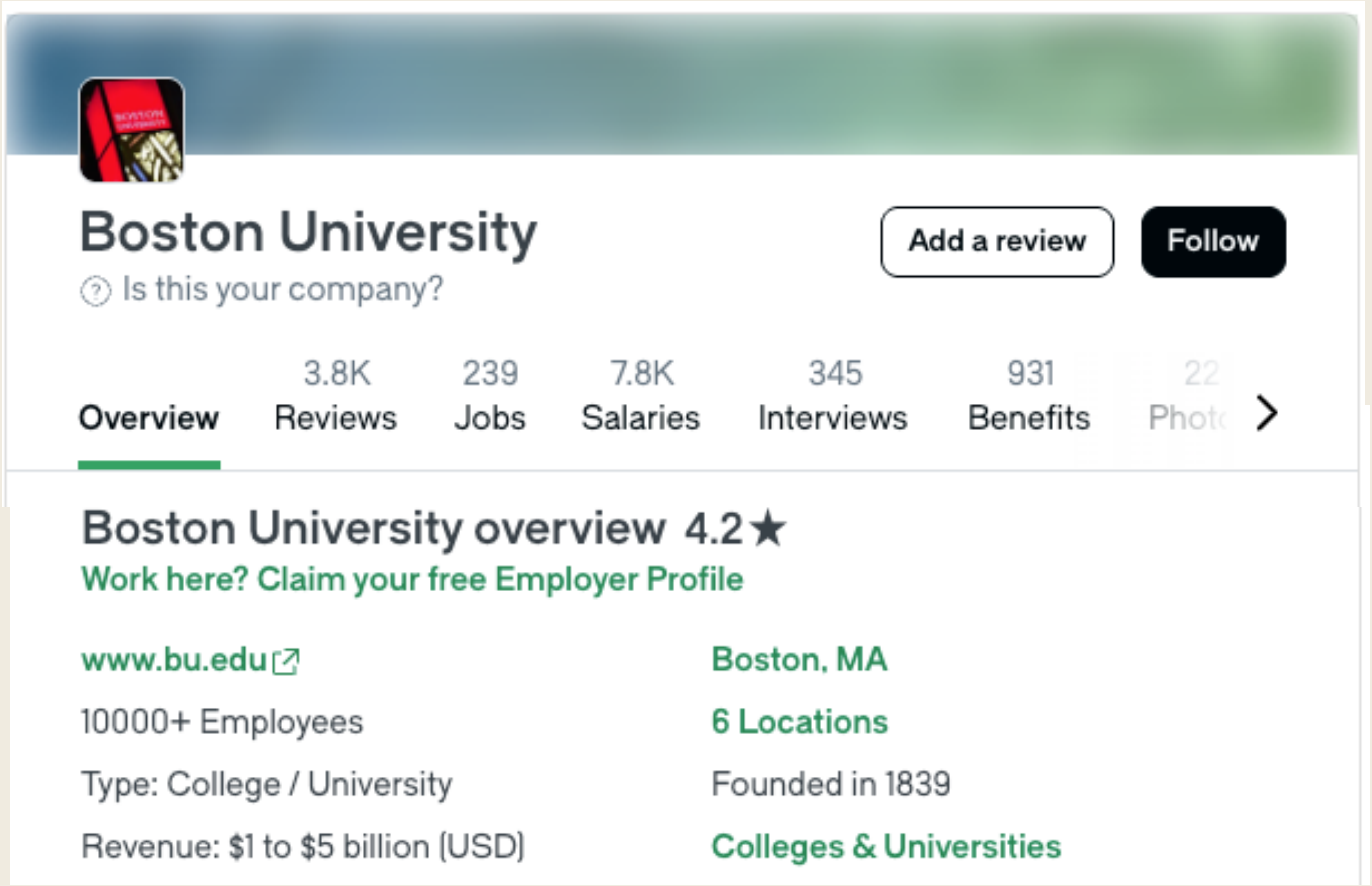
Higher Firm Wage, Lower Separation (US)



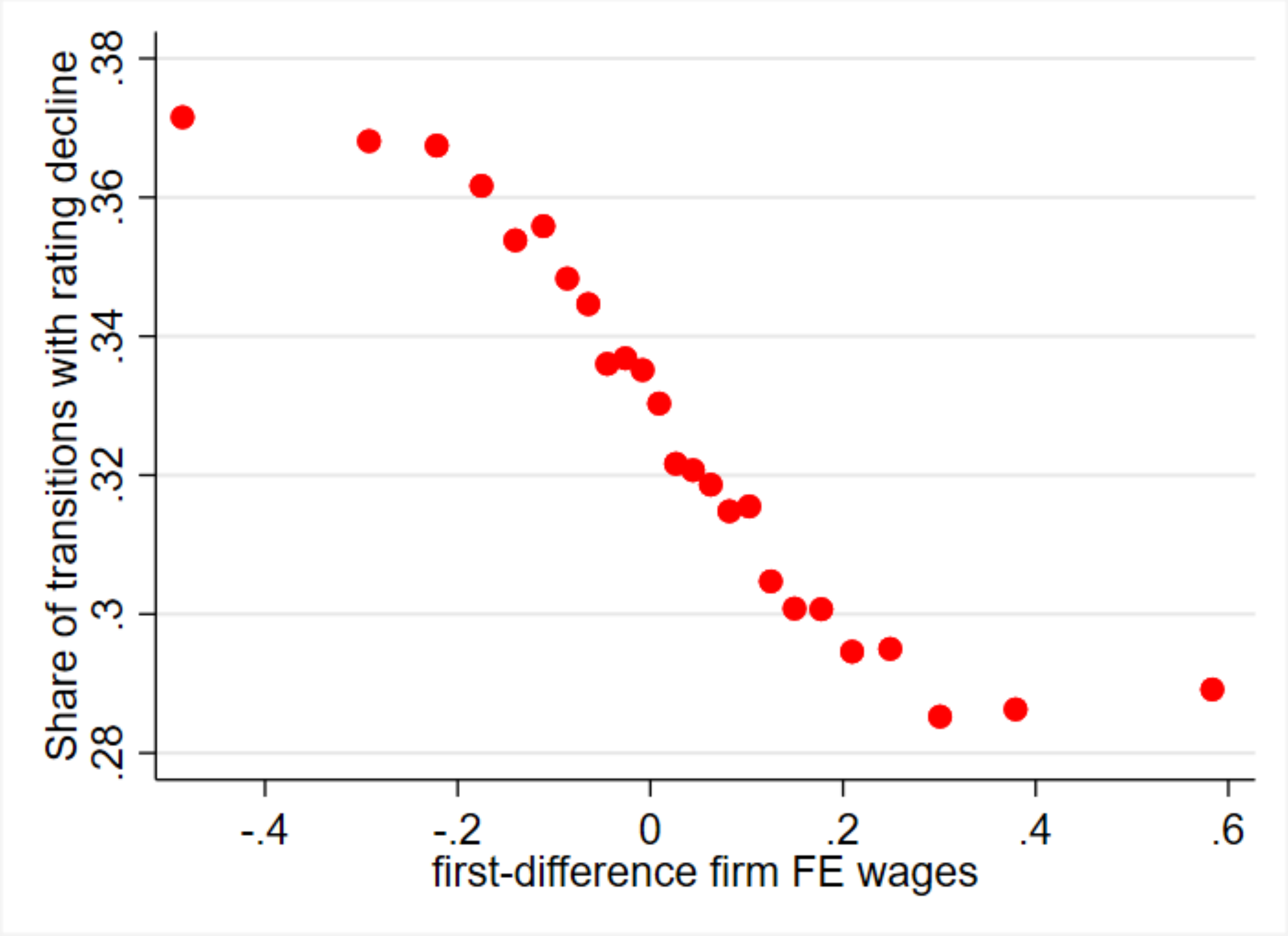
Source: Bassier, Dube, & Naidu (2016).

Higher Firm Wage, Higher Job Satisfaction (US)

Glassdoor



Probability of rating decline when workers change jobs



Firm Wage Effects Drive Wage Dynamics

- So far, firm wage effects seem important for cross-sectional inequality
- Firm wage effects appear to be an important driver of earnings dynamics (Schmieder, von Wachter, and Heining, 2017)

Careers after Mass Layoffs (Germany)

Earnings loss

Careers after Mass Layoffs (Germany)

Panel B. Annual earnings in euros: event study

Panel D. log daily wage: event study

Panel F. Annual days worked: event study

Earnings loss = Wage loss + Employment loss

Careers after Mass Layoffs (Germany)

Panel B. Annual earnings in euros: event study

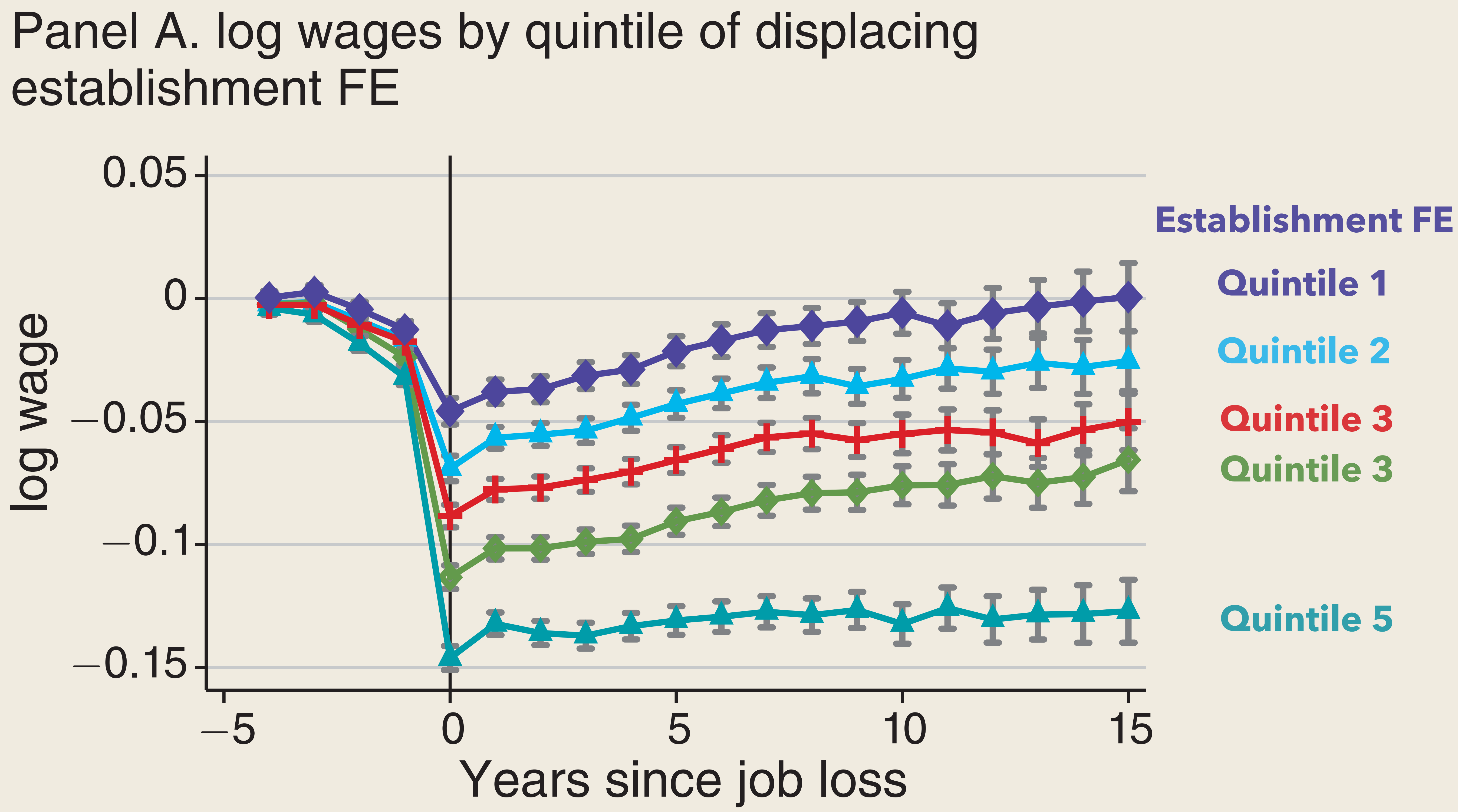
Panel D. log daily wage: event study

Panel F. Annual days worked: event study

Earnings loss = Wage loss + Employment loss

- It is well known that displaced workers suffer large and persistent earnings losses
- What drives the wage loss?

Losses in Firm Wage Effects



- A sharp and permanent drop in firm (establishment) FE after displacement
- Larger wage losses if displaced by a firm with higher FE
- Firm wage effects account for 70-90% of the wage losses from displacement

Beyond Additive Separability: Theory

– Bonhomme, Lamadon, and Manresa (2019)

Beyond Additive Separability?

- AKM imposes (log-)additive separability:

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

- This is a strong assumption
- It implies that low- and high-wage workers equally benefit from high-wage firms (in proportional terms)
- Rules out complementarity
- Can we relax?

BLM Model

- Consider a strict generalization of AKM:

$$w_{it} = b_{j(i,t)}\,\alpha_i + \psi_{j(i,t)} + \epsilon_{it}$$

- Flexible interaction term
  - AKM is a special case with $b_j = 1$ for all j

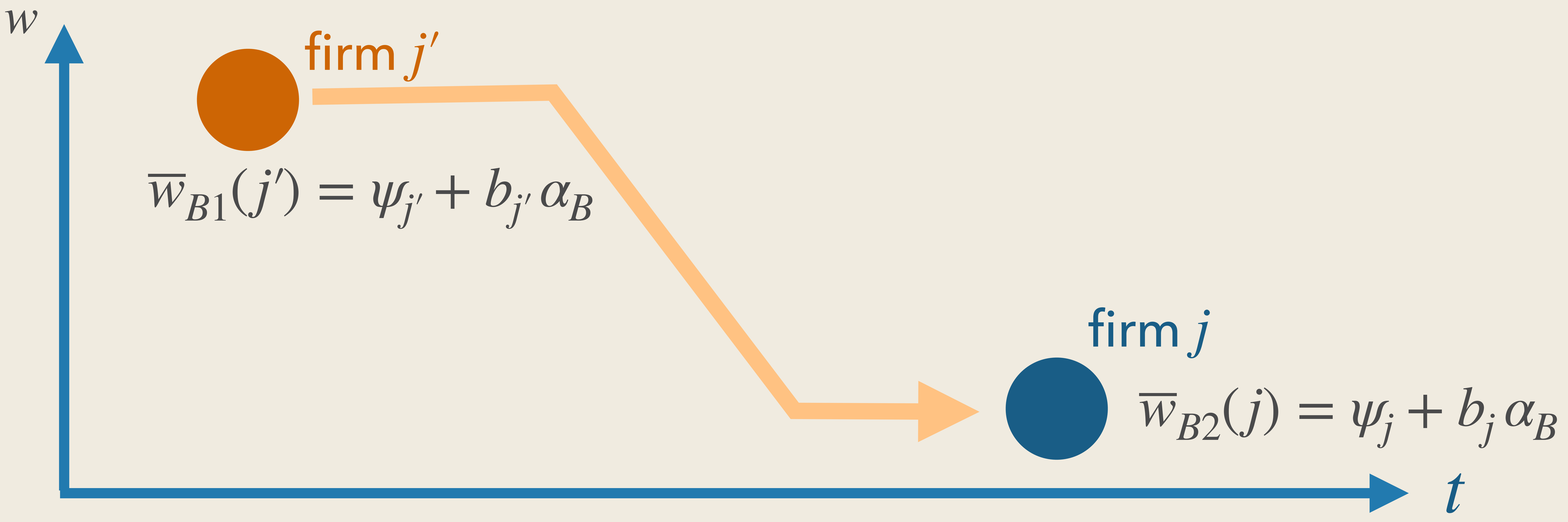
- Can we identify b_j?

Identification Assumptions

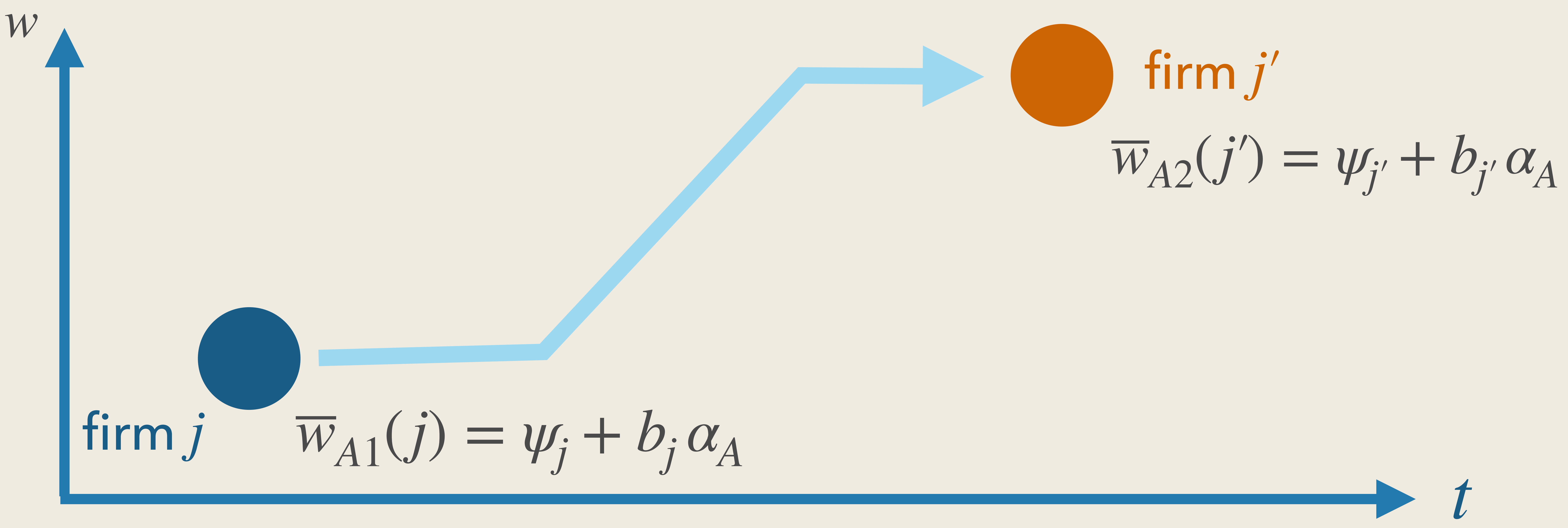
$$w_{it} = \psi_{j(i,t)} + b_{j(i,t)}\,\alpha_i + \epsilon_{it}$$

- Two periods, $t = 1,2$
- Assumptions:
  1. we observe
     - a continuum of workers moving from firm j to j'
     - a continuum of workers moving from firm j' to j
  2. $\mathbb{E}[\epsilon_{it} \,|\, \alpha_i, j(i,1) = j, j(i,2) = j']$ holds for all i, j, j'

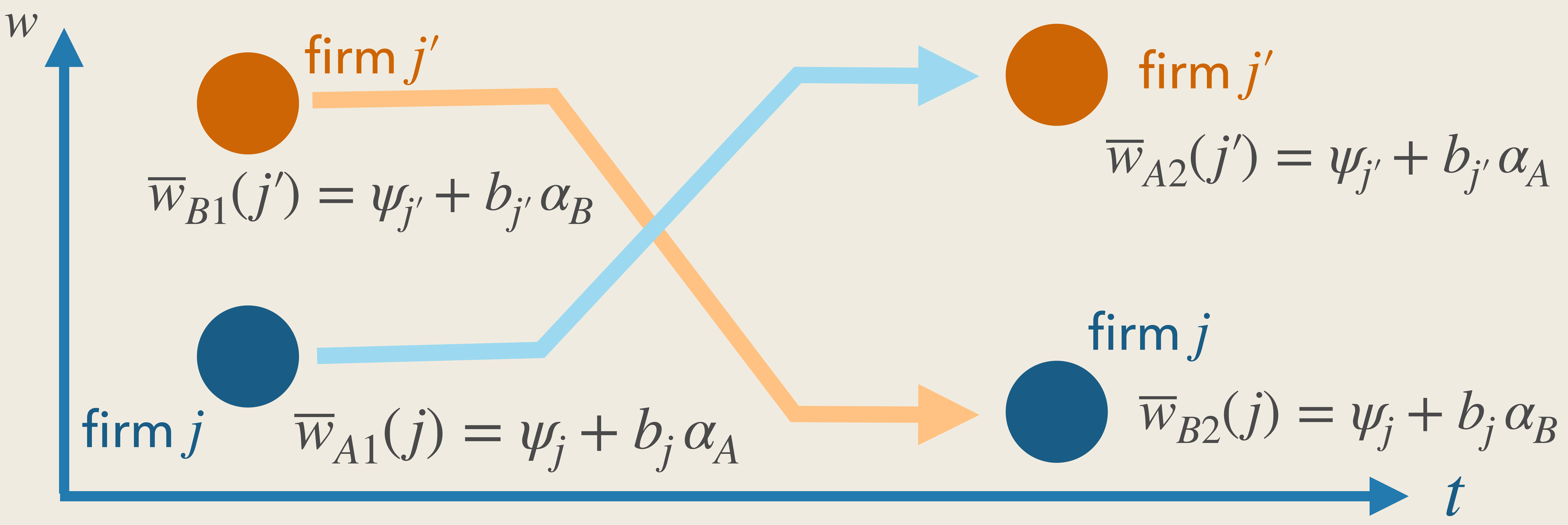
Identification Result



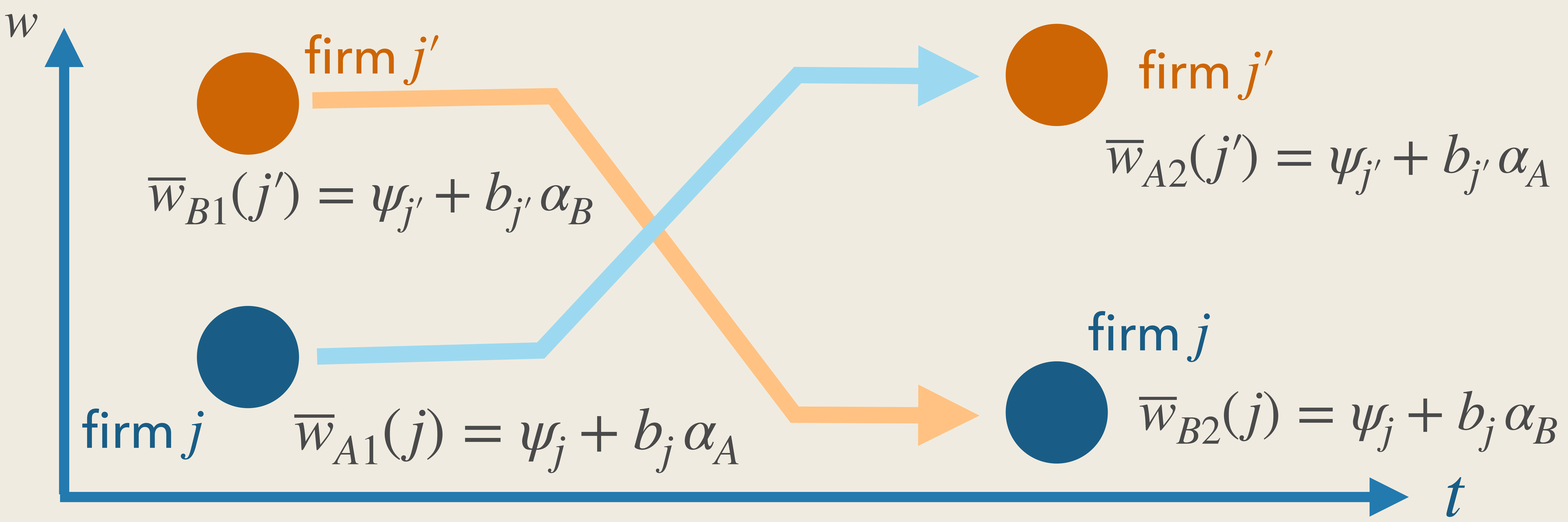
Identification Result



Identification Result



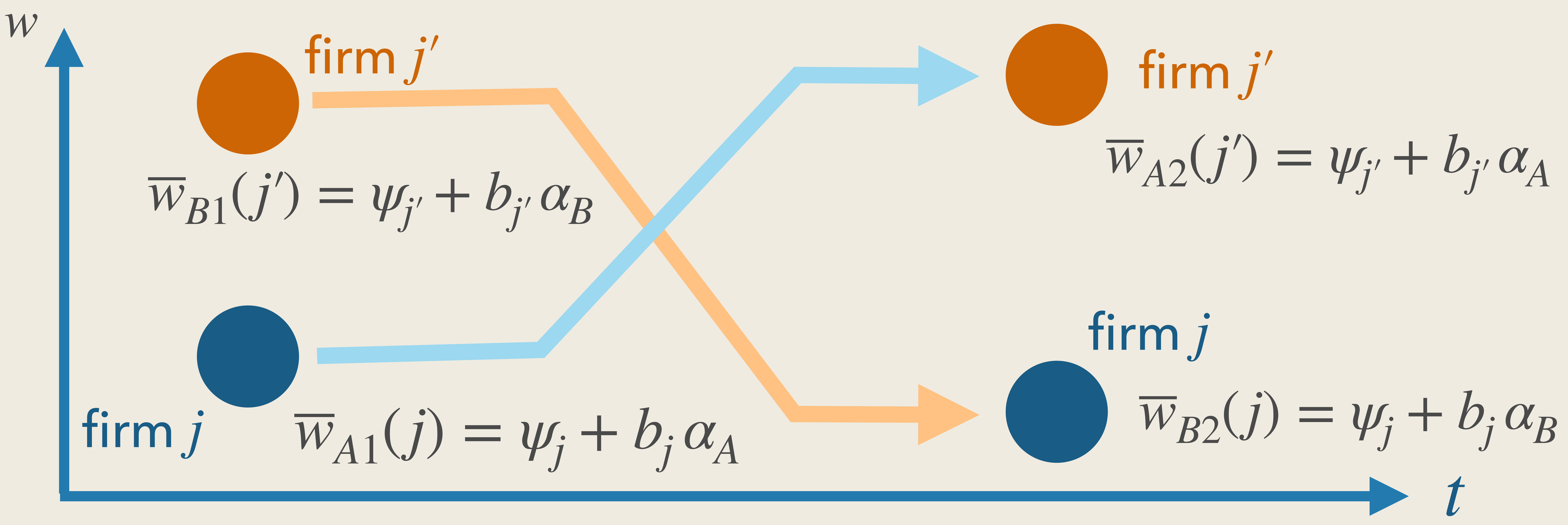
Identification Result



- **Result:** $\{b_j\}$ is identified (up to scale)

$$\overline{w}_{A2}(j') - \overline{w}_{B1}(j') = b_{j'}(\alpha_A - \alpha_B),$$

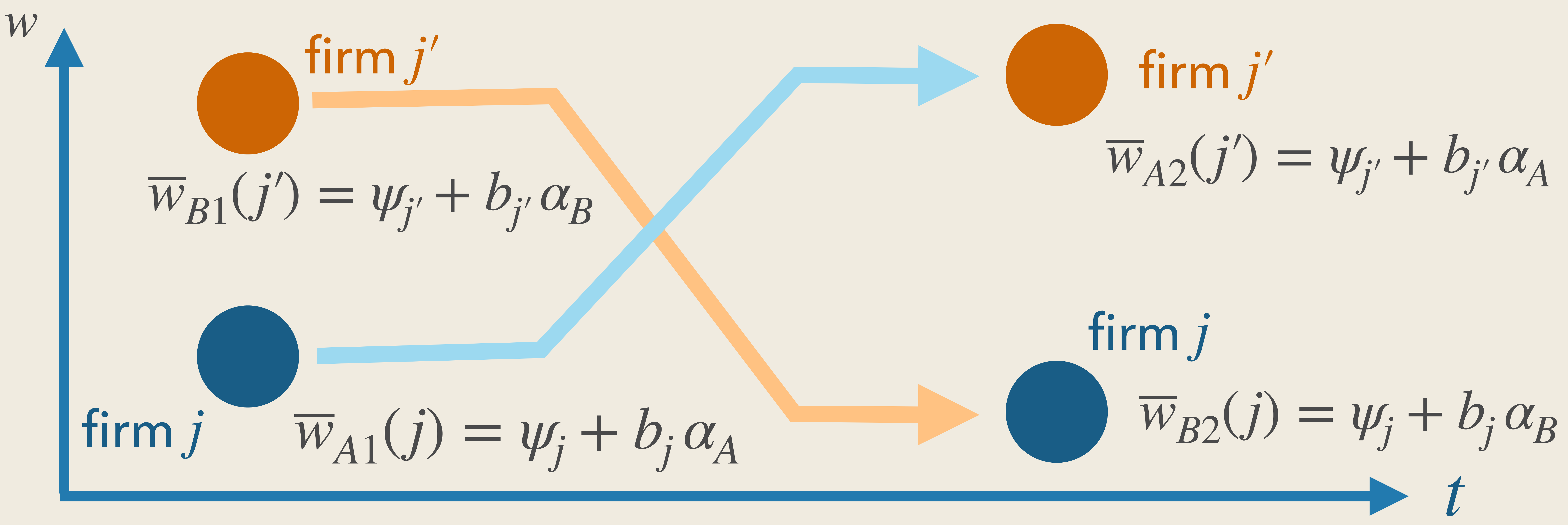
Identification Result



- **Result:** $\{b_j\}$ is identified (up to scale)

$$\overline{w}_{A2}(j') - \overline{w}_{B1}(j') = b_{j'}(\alpha_A - \alpha_B), \qquad \overline{w}_{A1}(j') - \overline{w}_{B2}(j') = b_j(\alpha_A - \alpha_B)$$

Identification Result



- **Result:** $\{b_j\}$ is identified (up to scale)

$$\overline{w}_{A2}(j') - \overline{w}_{B1}(j') = b_{j'}(\alpha_A - \alpha_B), \qquad \overline{w}_{A1}(j') - \overline{w}_{B2}(j') = b_j(\alpha_A - \alpha_B)$$

$$\Rightarrow \quad \frac{b_{j'}}{b_j} = \frac{\bar{w}_{A2}(j') - \bar{w}_{B1}(j')}{\bar{w}_{A1}(j) - \bar{w}_{B2}(j)}$$

Estimation in Practice

- In practice, estimation poses a challenge
- $b_{j'}/b_j$ is identified by comparing workers moving from j to j' and j' to j
- For any firm pairs, there are only a handful of such job-movers
  - Regardless of the sample size of workers/firms
- This causes incidental parameter biases (a.k.a. "limited mobility bias"):
  - Estimates of non-linear models are biased with too little variation
- BLM (2019) solution:
  - Firms are clustered in K discrete firm types ($K = 10$ in the application)
  - Assume $b_j = \tilde{b}_{k(j)}$ and $\psi_j = \tilde{\psi}_{k(j)}$

k-means Clustering

- How do we know which cluster a firm belongs to?
- Use k-means clustering to partition firms:

$$\min_{k(1),\ldots,k(J),H_1,\ldots,H_K} \sum_{j=1}^{J} n_j \sum_{d=1}^{D} \left(\hat{F}_j(w_d) - H_{k(j)}(w_d) \right)^2$$

  - n_j: number of workers in firm j
  - w_d: grid points on wage distribution
  - $\hat{F}_j(w)$: empirical cdf of wage distribution in firm j
- BLM (2022): Even with continuous heterogeneity, k-means provide approximation

Two-Step Estimation

1. Cluster firms into K groups based on wage distributions using k-means
    - any language has a package with efficient algorithm
2. Estimate parameters (in the previous example $\{\psi_k, b_k, \alpha_i\}$)
    - In theory, non-parametric identification is possible (see the paper)
    - In application, assume
        - workers consist of L types indexed by α
        - parameterize the wage distribution in terms of (α, k)
        - estimate using maximum likelihood

In the paper:
"dynamic model" that incorporates hitory dependence in earnings and mobility

Beyond Additive Separability: Application

– Bonhomme, Lamadon, and Manresa (2019)

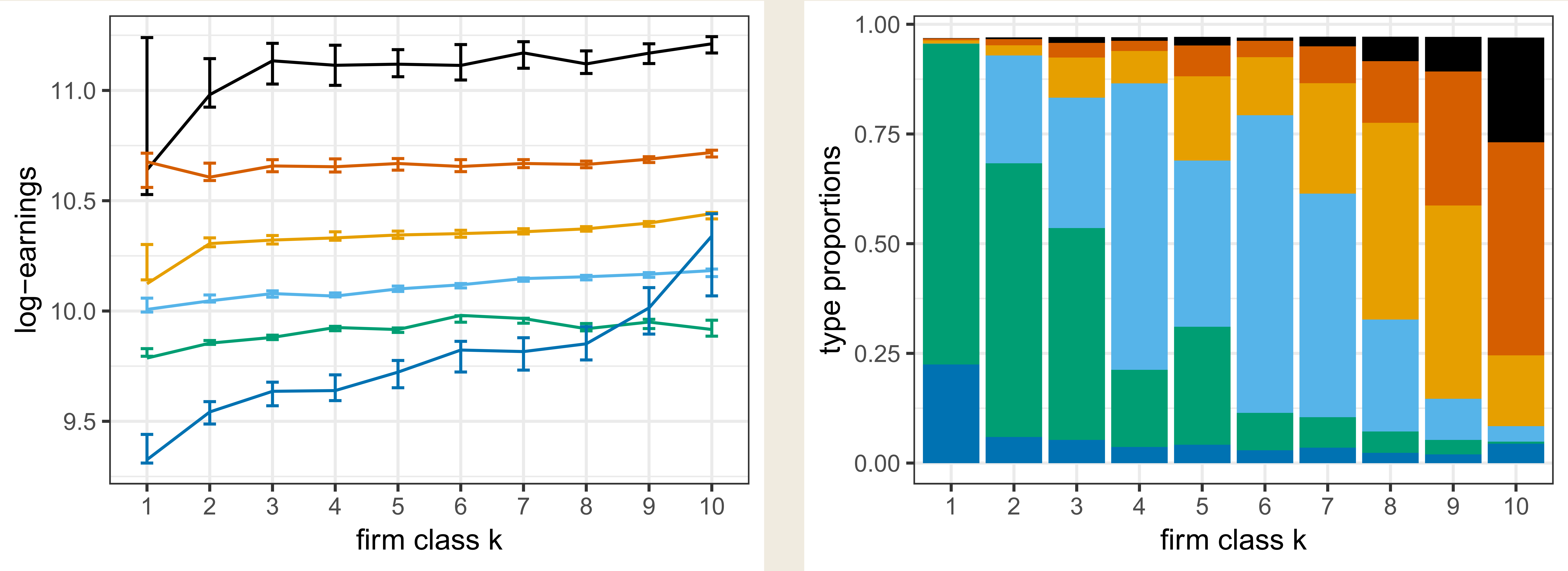
Application to Swedish Labor Market

- Employer-employee data from Sweden, 2002 & 2004
- Focus on full-year employed male: 1,000,000 workers and 60,000 firms
- Parameteric assumptions:
  - $w_{it} \sim LN(\mu_{k\alpha}, \sigma_{k\alpha})$
  - $\pi_{k\alpha}$: proportion of α workers in firm k (non-parameteric)
  - $p_{kk'}(\alpha)$ moving prob. of α workers from firm k to k' (non-parameteric)
- With $K = 10$ and $L = 6$, we have 900 parameters to estimate

Clustering

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

Main Result



1. No strong evidence of complementarity
   - If anything, low-wage workers gain the most from working at high-wage firms
2. Strong sorting between high-wage firms and high-wage workers

Puzzle to Keep in Mind

"the presence of strong sorting, together with the absence of strong complementarities in wages, is difficult to reconcile with models where sorting is driven by complementarities in production"

— Bonhomme, Lamadon, and Manresa (2019)

Summary

- AKM firm fixed effects:
  - accounts for an important share of cross-sectional wage inequality
  - vary systematically with respect to firm characteristics
- BLM relaxes strong parameteric assumption in AKM…
  …yet it turns out AKM was good enough!
- AKM/BLM firm effects are statistical in nature
- What do these objects mean? – a question we tackle next