Superstar Teams

Lukas Freund

Columbia University

April 17, 2025

NYU/Stern

Motivation: people have specialized skills & work in teams

- **Motivation:** production is complex $\rightarrow$ **specialized skills** & **team production**

[Marschak-Radner, 1972; Becker-Murphy, 1992; Garicano, 2000]

Motivation: people have specialized skills & work in teams

- **Motivation:** production is complex $\rightarrow$ **specialized skills** & **team production**

  [Marschak-Radner, 1972; Becker-Murphy, 1992; Garicano, 2000]

- **Q:** how do these fundamental features of production affect **macro outcomes?**
  - what happens to labor market inequality & agg. productivity as skill specificity $\uparrow$?

Motivation: people have specialized skills & work in teams

- **Motivation:** production is complex $\rightarrow$ **specialized skills** & **team production**

  [Marschak-Radner, 1972; Becker-Murphy, 1992; Garicano, 2000]

- **Q:** how do these fundamental features of production affect **macro outcomes?**
  - what happens to labor market inequality & agg. productivity as skill specificity $\uparrow$?

- **Challenge:** macro models typically abstract from firm organization

  [recent steps: Jarosch-Oberfield-RossiHansberg, 2021; Herkenhoff-Lise-Menzio-Phillips, 2024]

Motivation: people have specialized skills & work in teams

- **Motivation:** production is complex $\rightarrow$ **specialized skills** & **team production**

  [Marschak-Radner, 1972; Becker-Murphy, 1992; Garicano, 2000]

- **Q:** how do these fundamental features of production affect **macro outcomes?**
  - what happens to labor market inequality & agg. productivity as skill specificity $\uparrow$?

- **Challenge:** macro models typically abstract from firm organization

  [recent steps: Jarosch-Oberfield-RossiHansberg, 2021; Herkenhoff-Lise-Menzio-Phillips, 2024]

- **This paper:** aggregative **framework** where skill specificity + teams matter for macro

Core idea: specialization $\rightarrow$ team complementarities $\rightarrow$ macro effects

- **Suppose**:
  - production requires many **tasks**
  - workers have **het. task-specific skills**
    - talent $\sim$ absolute skill
    - skill specificity $\sim$ *dispersion* in ind. task-specific skills
  - firms hire multiple workers (**"team"**)
  - hiring involves random **search**

Core idea: specialization → team complementarities → macro effects

- **Suppose**:
  - production requires many **tasks**
  - workers have **het. task-specific skills**
  - firms hire multiple workers (**“team”**)
  - hiring involves random **search**

→ talent ∼ absolute skill
→ skill specificity ∼ *dispersion* in ind. task-specific skills

- **Firm-level production:** **skill specificity** endogenously generates (1) productivity gains from teamwork, and (2) **coworker complementarities**

Core idea: specialization → team complementarities → macro effects

- **Suppose**:
  - production requires many **tasks**
  - workers have **het. task-specific skills**
    - → talent ∼ absolute skill
    - → skill specificity ∼ *dispersion* in ind. task-specific skills
  - firms hire multiple workers (**"team"**)
  - hiring involves random **search**

- **Firm-level production:** **skill specificity** endogenously generates (1) productivity gains from teamwork, and (2) **coworker complementarities**

- **Macro-level:**
  - talent complementarities → assortative matching → **firm-level inequality**
  - specialization gains vs. frictional coworker mismatch → **agg. productivity**

This paper: framework of the firm as a "team assembly"

Literature

1. **Theory**

2. **Measurement**

3. **Applications**

This paper: framework of the firm as a "team assembly"

▸ Literature

1. **Theory**
   - **microfound task-based production fn.** $\rightarrow$ endogenous coworker complementarities
   - tractable enough to characterize equilibrium team formation with search
2. **Measurement**
3. **Applications**

This paper: framework of the firm as a "team assembly"

▸ Literature

1. **Theory**
   - **microfound task-based production fn.** $\rightarrow$ endogenous coworker complementarities
   - tractable enough to characterize equilibrium team formation with search
2. **Measurement**
   - micro panel data on wages+matches $\rightarrow$ identification, estimation & validation
3. **Applications**

This paper: framework of the firm as a "team assembly"

▸ Literature

1 Theory

- **microfound task-based production fn.** $\rightarrow$ endogenous coworker complementarities
- tractable enough to characterize equilibrium team formation with search

2 Measurement

- micro panel data on wages+matches $\rightarrow$ identification, estimation & validation

3 Applications: implications of $\uparrow$ **skill specificity**

- **structural explanation for "firming up inequality"** *[e.g. Card et al., 2013; Bloom et al., 2019]*

This paper: framework of the firm as a "team assembly"

▸ Literature

1. **Theory**
   - **microfound task-based production fn.** $\rightarrow$ endogenous coworker complementarities
   - tractable enough to characterize equilibrium team formation with search

2. **Measurement**
   - micro panel data on wages+matches $\rightarrow$ identification, estimation & validation

3. **Applications**: implications of $\uparrow$ **skill specificity**
   - **structural explanation for "firming up inequality"** *[e.g. Card et al., 2013; Bloom et al., 2019]*
   - **agg. productivity gains limited by labor market frictions**

Roadmap

Theory

Measurement

Applications

Environment: task-based production & frictional matching into teams

- **Agents:** continuums of workers & firms, infinitely-lived, risk-neutral
  - **firms** are ex-ante identical; $n \in \{0, 1, 2\}$ employees
  - **worker** $i \in [0, 1]$ is endowed with time-invariant, task-specific skills, $\{z_i(\tau)\}_{\tau \in [0,1]}$
- **Production:** requires differentiated tasks *[e.g., Acemoglu-Restrepo, 2018]*
- **Labor market**: workers & multi-worker firms meet through random search

[similar to Herkenhoff-Lise-Menzio-Phillips (2024) but with high-dim. skills]

Environment: task-based production & frictional matching into teams

- **Agents:** continuums of workers & firms, infinitely-lived, risk-neutral
  - **firms** are ex-ante identical; $n \in \{0, 1, 2\}$ employees
  - **worker** $i \in [0, 1]$ is endowed with time-invariant, task-specific skills, $\{z_i(\tau)\}_{\tau \in [0,1]}$
- **Production:** requires differentiated tasks *[e.g., Acemoglu-Restrepo, 2018]*
- **Labor market**: workers & multi-worker firms meet through random search

  [similar to Herkenhoff-Lise-Menzio-Phillips (2024) but with high-dim. skills]

- **Game plan**: parametrize skill dist. and...
  1. microfound tractable *reduced-form* firm-level production fn $f(\cdot)$
  2. given $f(\cdot)$, analyze team formation

Parametrized multi-dim. skills: marginal distribution

Assumption: Fréchet dist.

$$\mathrm{P}\left[z_i(\tau) \leq z\right] = \exp\left(-\left(\frac{z}{\iota x_i}\right)^{-\frac{1}{\chi}}\right)$$

with $x_i \in \mathbb{Z}_{++}$ (**"talent"** $\sim$ scale), $\chi \in [0, \infty)$ (**"skill specificity"** $\sim$ inverse shape), $\iota > 0$ (scaling term)

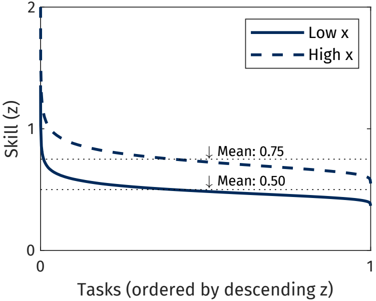
[Eaton & Kortum, 2002]

- In population: x dist. according to some cdf $\tilde{\Psi}_x$

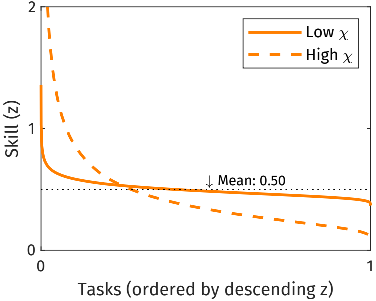
Parametrized multi-dim. skills: marginal distribution

$$P\left[z_i(\tau) \le z\right] = \exp\left(-\left(\frac{z}{\iota x_i}\right)^{-\frac{1}{\chi}} \right)$$

(a) x_i: talent (scale)



(b) χ: specificity (1/shape)



Parametrized multi-dim. skills: multivariate distribution

Brown-Resnick process

- Individuals positioned in a **latent, cylindrical space** → height: talent x → circle: relative specialization
- Conditional on x, individuals uniformly positioned on a **circle** with circumference 2, $\mu_i \in [0, 2)$, with distance metric $d_{il} = d(\mu_i, \mu_l) = \min\{|\mu_i - \mu_l|, 2 - |\mu_i - \mu_l|\}$
- For any pair (i, l), 'correlation' of skills is governed by a Gumbel **copula**

$$C_{il}(u, v) = \exp\left\{-\left[(-\log u)^{1/\xi_{il}} + (\log v)^{1/\xi_{il}}\right]^{\xi_{il}}\right\}$$

with **distance-dependent association** $\xi_{il} = g(d_{il})$ with $g : [0, 1) \to [0, 1)$ increasing

  - $\xi_{il} \to 0$: dependence $\xi_{il} \to 1$: independence

Production with a single team – taking composition as given

- Firm with n workers produces output from **unit continuum of tasks** $\mathcal{T} = [0, 1]$

$$\ln Y = \int_{\mathcal{T}} \ln q(\tau) d\tau \tag{1}$$

- **Task-level aggregation** for task τ:

$$q(\tau) = \sum_{i=1}^{n} y_i(\tau) \tag{2}$$

- **Task production:** i has task-specific skill $z_i(\tau)$, supplies 1 time unit

$$y_i(\tau) = z_i(\tau) l_i(\tau) \tag{3}$$

$$1 = \int_{\mathcal{T}} l_i(\tau)\, d\tau \tag{4}$$

Firm's optimization problem

- **Firm solves mini-planner problem:** $\max_{\mathbf{q},\{\mathbf{y_i}\},\{\mathbf{l_i}\}} Y$ s.t. (1)-(4)

  $\Rightarrow$ derive & characterize *reduced-form* team production function f

$$f(\mathbf{z}_1, \dots, \mathbf{z}_n) = \max Y$$
$$\text{s.t. (1)-(4)}$$

Firm's optimization problem

- **Firm solves mini-planner problem:** max Y s.t. (1)-(4)

$$\mathcal{L}(\cdot) = Y + \lambda \underbrace{\left[\left(\int_{\mathcal{T}} \ln q(\tau) d\tau \right) - \ln Y \right]}_{\text{tasks} \rightarrow \text{output}} + \int_{\mathcal{T}} \lambda(\tau) \underbrace{\left(\sum_{i=1}^{n} y_i(\tau) - q(\tau) \right)}_{\text{task aggregation}} d\tau$$

$$+ \sum_{i=1}^{n} \lambda_i^L \underbrace{\left(\int_{\mathcal{T}} \frac{y_i(\tau)}{z_i(\tau)} d\tau - 1 \right)}_{\text{time constraint + task production}} + \text{non-negativity constraints}$$

- **FOCs** imply

$$\lambda(\tau) = \min_i \left\{ \frac{\lambda_i^L}{z_i(\tau)} \right\}$$

shadow cost of τ

opportunity cost of i's time

i's skill for τ

Firm's optimization problem: canonical task assignment

- **Firm solves mini-planner problem:** max Y s.t. (1)-(4)

$$\mathcal{L}(\cdot) = Y + \lambda \underbrace{\left[\left(\int_{\mathcal{T}} \ln q(\tau) d\tau \right) - \ln Y \right]}_{\text{tasks} \to \text{output}} + \int_{\mathcal{T}} \lambda(\tau) \underbrace{\left(\sum_{i=1}^{n} y_i(\tau) - q(\tau) \right)}_{\text{task aggregation}} d\tau$$

$$+ \sum_{i=1}^{n} \lambda_i^L \underbrace{\left(\int_{\mathcal{T}} \frac{y_i(\tau)}{z_i(\tau)} d\tau - 1 \right)}_{\text{time constraint + task production}} + \text{non-negativity constraints}$$

- **FOCs** imply **task assignment by comparative advantage**

$$\lambda(\tau) = \min_i \left\{ \frac{\lambda_i^L}{z_i(\tau)} \right\} \quad \Rightarrow \quad \mathcal{T}_i = \left\{ \tau \in \mathcal{T} : \frac{z_i(\tau)}{\lambda_i^L} \geq \max_{k \neq i} \frac{z_k(\tau)}{\lambda_k^L} \right\}$$

Micro-founded production function: aggregation result

▸ Lemma ▸ General n

Proposition: Aggregation result

If skills are distributed multivariate Fréchet, then talents $\{x_i\}$ and horizontal distances $\{\xi_{il}\}$ are sufficient statistics for team output Y given parameter χ:

$$Y = f(\{x_i\}, \{\xi_{il}\}; \chi)$$

- **Proof sketch:** Fréchet max-stability property yields closed-form characterization of dist. of $\{\lambda(\tau)\}$, task shares, cost index λ, $\{\lambda_i^L\}_i \rightarrow$ integrate over task continuum & workers, find f after normalizing $\lambda = 1$

- Next consider closed-form expression for $n = 2$ case

Micro-founded production function: properties

General n

$$Y = f(x_i, x_l, \xi_{il}; \chi) = \left((x_i)^{\frac{1}{1+\chi\xi_{il}}} + (x_l)^{\frac{1}{1+\chi\xi_{il}}} \right)^{1+\chi\xi_{il}}$$

- **Benchmark** without division of labor: $Y = \sum_{i=1}^{n} x_i$

Micro-founded production function: properties

▸ General n

$$Y = f(x_i, x_l, \xi_{il}; \chi) = \left((x_i)^{\frac{1}{1+\chi\xi_{il}}} + (x_l)^{\frac{1}{1+\chi\xi_{il}}} \right)^{1+\chi\xi_{il}}$$

1. **Super-additive**: $f(\cdot) > \sum_{i=1}^{n} x_i$
2. **Super-modular**: $\frac{\partial^2 f}{\partial x_i \partial x_l} > 0$

Gains from team production are increasing in skill specificity

$$f(x_i, x_l, \xi_{il}; \chi) = \underbrace{2^{1+\chi\xi_{il}}}_{\text{efficiency gains}} \times \left(\frac{1}{2}(x_i)^{\frac{1}{1+\chi\xi_{il}}} + \frac{1}{2}(x_l)^{\frac{1}{1+\chi\xi_{il}}} \right)^{1+\chi\xi_{il}}$$

1. **Gains from team production** increasing in skill specificity (χ)
   - realizing gains requires coworkers being *horizontally distant* (ξ_{il})

▸ Intuition

Skill specificity implies that productivity is lowered by talent dispersion

$$f(\mathbf{x}, \xi; \chi) = \underbrace{2^{1+\chi\xi_{il}}}_{\text{efficiency gains}} \times \underbrace{\left(\frac{1}{2}(x_i)^{\frac{1}{1+\chi\xi_{il}}} + \frac{1}{2}(x_l)^{\frac{1}{1+\chi\xi_{il}}} \right)^{1+\chi\xi_{il}}}_{\textbf{talent complementarity}},$$

1. Gains from team production increasing in skill specificity (χ) ▸ Intuition
2. **Coworker talent complementarities** increasing in skill specificity (χ) ▸ Intuition
   - $\frac{\partial\left(\partial^2 f(\cdot)/\partial x_i \partial x_l\right)}{\partial \chi} > 0$

Roadmap & key takeaways

Theory

1. Economics: **skill specificity** *endogenously* generates (1) **gains** from team production & (2) **coworker talent complementarities**
   - methodology: Fréchet + optimal assignment $\rightarrow$ low-dim. $f(\cdot)$ despite high-dim. skills
2. Next: given $f(\cdot)$, what is the endogenous composition of different teams?

Endogenous team composition: frictional matching

- **More details on environment:**
  - random search with firm size $n \in \{0, 1, 2\}$ *[cf. Herkenhoff-Lise-Menzio-Phillips, 2024]*
  - exogenous separations, matching decision endogenous
  - employment states: unemp., employed alone, employed with one coworker
  - Nash wage bargaining with continuous renegotiation ▸ Details
  - only unemployed search
- ξ_{il} interpretable as **unobserved match-quality component**
  - tractable b/c by Prop. 1, $(\mathbf{x}, \xi)$ is sufficient statistic
  - distribution of ξ_{il}, H, satisfies $H = g^{-1}$ as $\xi_{il} = g(d_{il})$
- **Stationary equilibrium** ▸ HJBs ▸ KFEs ▸ Definition

Surplus max. determines which teams are formed

- Joint value of firm with 1 worker of talent x satisfies:

$$\rho\Omega_1(x) = f(x) + \delta(x)\Big[-\Omega_1(x) + V_u(x) + V_{f.0}\Big]$$
$$+ \lambda_{v.u} \int\int \frac{d_u(x')}{u} \max\Big\{\underbrace{-\Omega_1(x) + V_{e.2}(x|x',\tilde{\xi}) + V_{f.2}(x,x',\xi)}_{(1-\omega)S(x'|x,\xi)}, 0\Big\} dH(\tilde{\xi})dx'$$

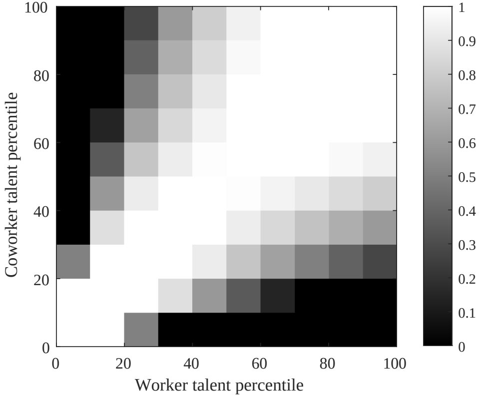
$V_u(x)$: value for unemp. worker; $V_{f.0}$: value for vacant firm; $d_u(x)$: density of unemployed workers; $u = \int d_u(x)dx$; ω: worker bargaining wgt; $\delta(x)$: sep. hazard; $\lambda_{v.u}$: hazard rate of vacancy meeting unmatched worker; H: cdf of ξ

- Surplus $S(x|x',\xi)$ reflects production complementarities

$$S(x|x',\xi)(\rho + \delta(x) + \delta(x')) = f(x,x',\xi) - \textit{outside options}$$

Equilibrium properties: conditional matching probabilities for given χ

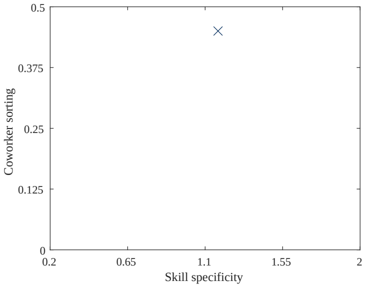
- Team composition determined by tradeoff between **match quality vs. search costs**
  $\Rightarrow$ cond. match probabilities $\mathrm{P}\left\{S(x'|x,\xi) > 0\right\}$



Skill specificity ⇒ coworker talent sorting

- **Mechanism:** skill specificity ⇒ **complementarities ⇒ coworker talent sorting**

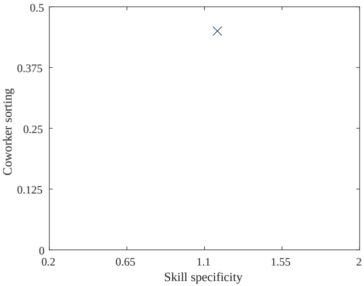
(a) Eqm. sorting



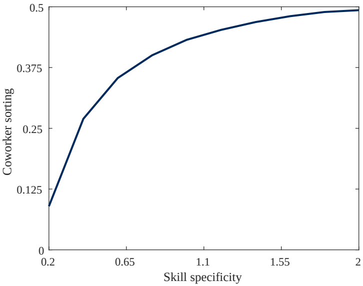
Skill specificity $\uparrow \Rightarrow$ coworker talent sorting $\uparrow$

- **Mechanism:** skill specificity $\uparrow \Rightarrow$ **complementarities $\uparrow \Rightarrow$ coworker talent sorting $\uparrow$**

(a) Eqm. sorting



(b) Comparative statics



Roadmap & key takeaways

Theory

1. **Skill specificity** *endogenously* generates **coworker complementarities**
2. **Talent complementarities** lead to **positive assortative matching**

Next: confront theory with data

Taking the model to the data: overview

- **Data:** SIEED matched employer-employee panel for West Germany
  - for now: 2010-2017; later: 1985-2017
- **Mapping & estimation**
  - worker i' s talent type $\hat{x}_i \approx$ decile in lifetime wage dist. ▸ Details
    – solve model numerically with talent types $\hat{x}_i \in \{1, \dots, 10\}$
  - "representative coworker type" $\hat{x}_{-it}$: avg. $\hat{x}$ of workers in same estab.-yr. ▸ Details
  - external: discount rate ρ, bargaining weight ω
  - estimated offline: job separation hazards $\delta(x)$
  - indirect inference: meeting rate, unemp. flow benefit, χ, mapping $\hat{x} \to x$
- **Main challenge:** skill specificity χ not directly observable
  - evidence for task-specific skills *[cf. Deming, 2023]* but no cardinal measure of specificity

Indirect inference

Reg. table

- Indirect inference approach to discipline χ

Sketch | Equation for $\bar{w}(x|x')$

$$\boxed{\chi \longrightarrow \frac{\partial^2 f(x,x',\xi)}{\partial x \partial x'} \propto \frac{\partial^2 w(x|x',\xi)}{\partial x \partial x'}.}$$

- Approximate $\frac{\partial^2 \bar{w}(x|x')}{\partial x \partial x'}$ using **regression with interaction term**

Long robustness list (it's a JMP...)

$$\frac{W_{it}}{\overline{W}_t} = \beta_0 + \sum_{d=2}^{10} \beta_{1d} \mathbf{1}\{\hat{x}_i = d\} + \sum_{d'=2}^{10} \beta_{2d'} \mathbf{1}\{\hat{x}_{-it} = d'\} + \beta_c(\hat{x}_i \times \hat{x}_{-it}) + \psi_{j(i,t)} + \nu_{o(i,t)t} + \xi_{s(i,t)t} + \epsilon$$

- **Estimation of structural model:** replicate semi-structural regression with model-generated data, infer χ by matching empirical $\hat{\beta}_c$

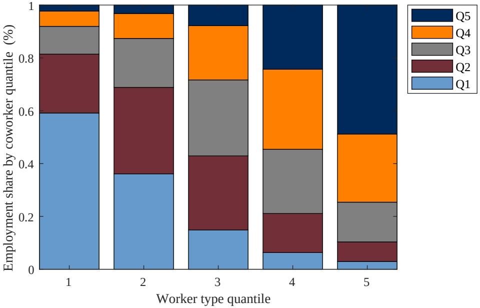
Parameter values

Model matches (untargeted) coworker talent sorting patterns in the data

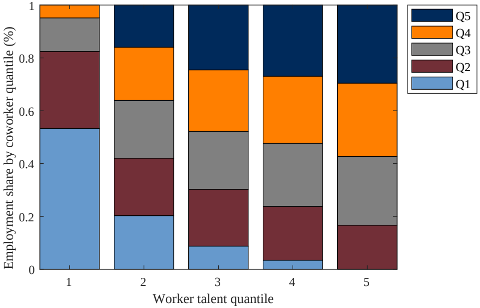
1. ✓More talented workers experience lower unemp. rates due to lower separation rates but job finding rates don't increase much with talent *[e.g. Cairo & Cajner, 2018]*

2. ✓**Coworker talent sorting patterns**

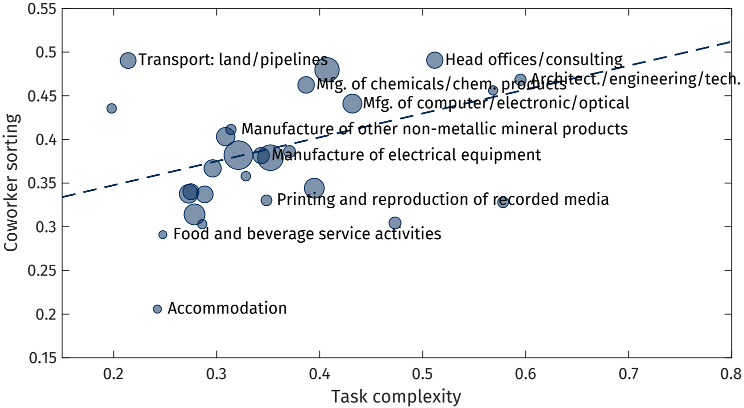
(a) Data



(b) Theory



Validation: industries with ↑ task complexity feature more sorting

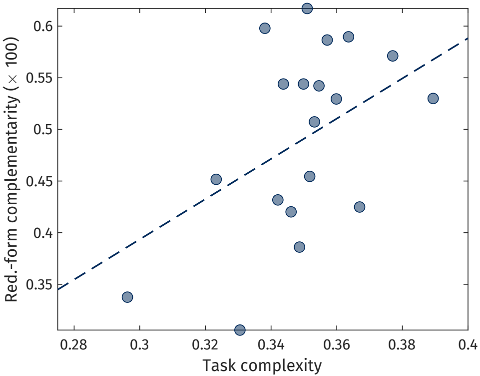


Notes. Task complexity: occupation-specific measure of the share of cognitive non-routine tasks, weighted by industry-specific occ. employment weights. Weighted linear best fit. Data: SIEED + BIBB.

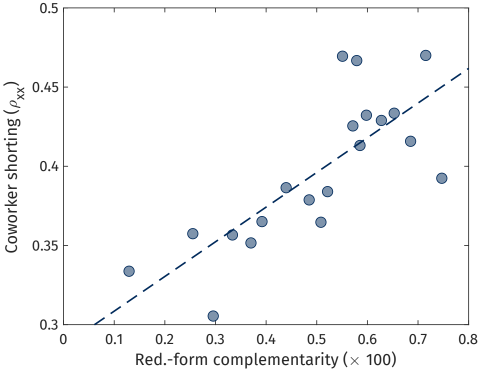
Validation: this pattern is consistent with the model

▸ W/o industry FEs ▸ Model vs. Data

(a) Skill specificity proxy → Complementarities



(b) Complementarities → sorting



Notes. Binned scatterplots, with industry FEs, so variation is within-industry over time. Moments estimated separately for 2-digit industries over 5 sample periods. Data: SIEED + BIBB.

Roadmap & key takeaways

Theory

1. Skill specificity *endogenously* generates coworker complementarities
2. Talent complementarities lead to positive assortative matching

Model Meets Data

3. Estimated model implies large ex-post differences across ex-ante identical firms

Next: applications

4. Structural explanation for the “firming up of inequality”
5. Implications for aggregate productivity

Hypothesis: growing skill specificity ($\chi \uparrow$)

Task movements

1. Δ **Task composition:** fewer routine (low-χ), more complex (high-χ) tasks *[Deming, 2017]*

DE evidence

2. **Burden of knowledge**: increasing cost of reaching the frontier – necessitates increasingly narrow individual expertise *[Jones, 2009]*

Medical specialization

3. **Education:** if education augments task-specific skills randomly, then the trend toward more (secondary & tertiary) education fosters $\uparrow$ dispersed task-specific skills

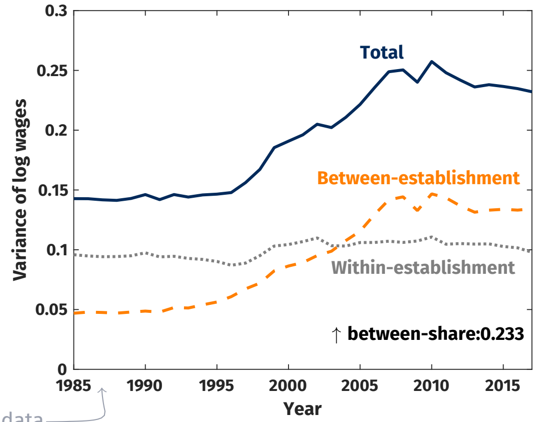
Formalization & edu data

Application 1: structural explanation for the “firming up” of inequality

▸ Details

“the variance of firm [wages] explains an increasing share of total inequality in a range of countries”

[Song-Price-Guvenen-Bloom-von Wachter, 2019]

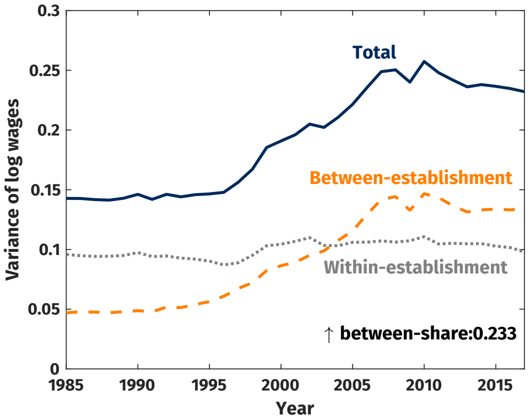


German matched employer-employee data

Application 1: structural explanation for the “firming up” of inequality

▸ Details

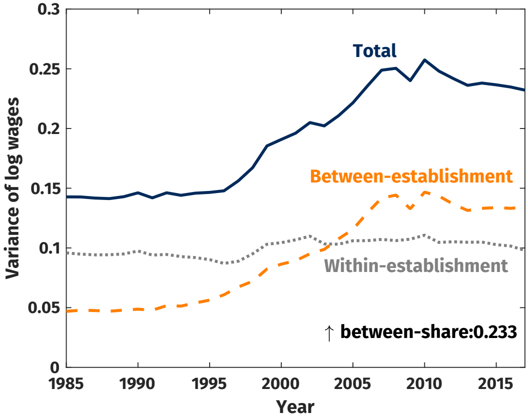
What are the causal driver(s)?



Application 1: structural explanation for the "firming up" of inequality

▸ Details

1. The set of tasks any one worker can perform well has narrowed: **skill specificity ↑**
2. **Coworker talent complementarities ↑**
3. **Coworker talent sorting ↑**
4. Greater ***firm-level*** **productivity & wage dispersion**



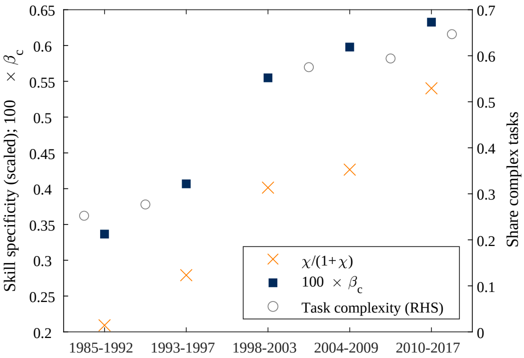
Estimate model for several periods: skill specificity ↑

▸ Schooling ▸ Peer effect trends

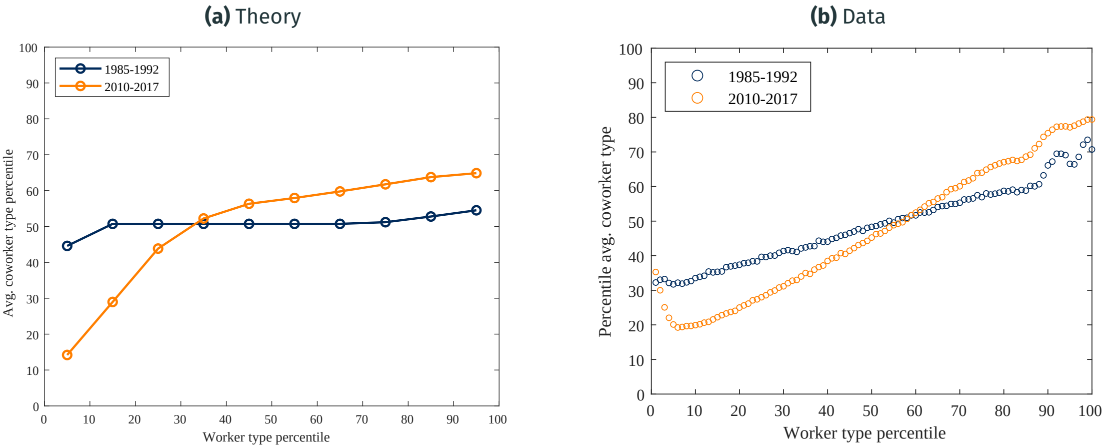
- **Method:** estimate reduced-form coefficient β_c for 5 sample periods
  $\Rightarrow$ re-estimate structural model

- **Skill specificity has intensified** ($\chi \uparrow$)

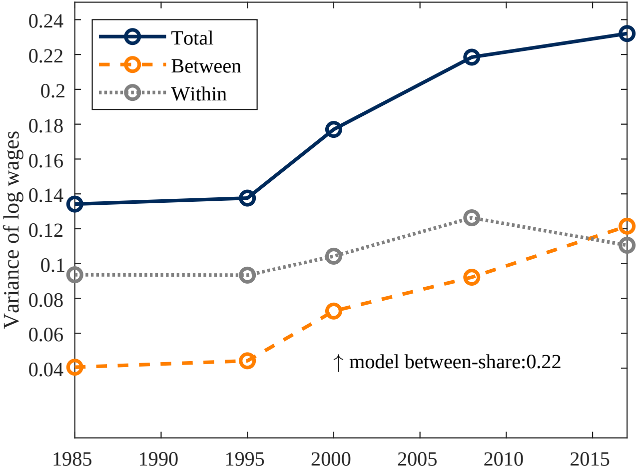
- Implied complementarities ↑
  - $\frac{f(x^{p80},x^{p80},1)+f(x^{p20},x^{p20},1)}{f(x^{p80},x^{p20},1)+f(x^{p80},x^{p20},1)}$: 1.05 $\nearrow$ 1.16



Talent sorting has intensified

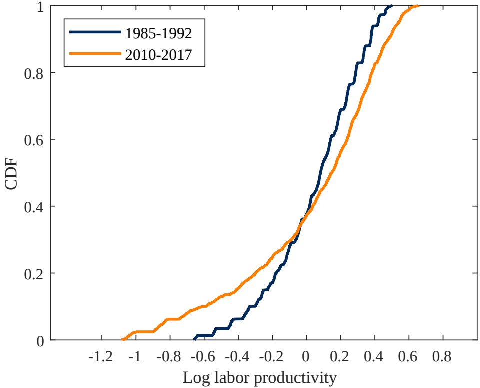


Model replicates observed ↑ firm-level wage inequality



Model endogenously generates ↑ firm-level productivity dispersion

- Firm dynamics literature: increased productivity dispersion *[Autor et al., 2020; de Ridder, 2024]*, correlated with wage & talent dispersion *[Berlingieri et al., 2017; Sorkin-Wallskog, 2020]*



Skill specificity $\chi \uparrow$ explains large share of “firming up” $\uparrow$

- **Q:** How much of $\uparrow$ between-firm share of wage var. is due to $\chi \uparrow$?
- **Counterfactual:** between-firm share in 2010s absent $\chi \uparrow$ since ’85-’92
- **A:** $\chi \uparrow$ **accounts for 65%** of model-predicted $\Delta \leftrightarrow \approx$ **59% of empirical Δ**
- **Robustness** exercises

▸ Within-industry ▸ Outsourcing

- Effect of $\downarrow$ search frictions *[e.g., Martellini-Menzio, 2021]* $\sim$ 11% of model-predicted Δ
  - search effort plausibly endogenous to χ

Roadmap & key takeaways

Jump to conclusion

Theory

1. Skill specificity *endogenously* generates coworker complementarities
2. Talent complementarities lead to positive assortative matching

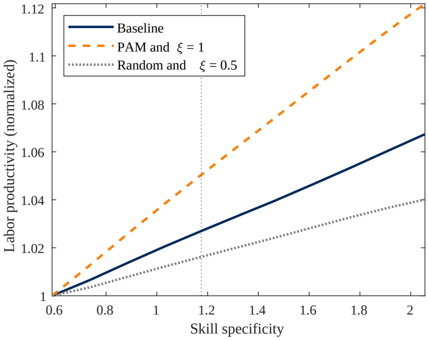
Model Meets Data

3. Estimated model endogenously generates realistic ex-post firm heterogeneity

Applications

4. **Increased skill specificity – leading to stronger complementarities and, hence, sorting – helps explain the “firming up” of inequality**
5. **Next:** productivity application

Labor market frictions impede productivity gains from specialization



- Gains from the division of labor are limited by the functioning of the labor market
  - microfoundation for recent econ-dev findings *[Bandiera-Kotia-Lindenlaub-Moser-Prat, 2024]*
  - labor market frictions may inhibit specialization *[cf. Atencio et al., 2024; Bassi et al., 2024]*

Conclusion: how firms organize teams matters for macro

▸ Literature

- **Core contribution**: tractable framework of **the firm as a "team assembly"**
- **3 takeaways:**
  1. specialization + teams foster firm-level heterogeneity due to **complementarities**
  2. $\uparrow$ skill specialization helps explain **"firming up" of inequality**
  3. **productivity gains** from specialization hinge on teams being well-matched
- **Follow-up work in progress:**
  - skill specialization shapes the earnings effects of AI *[with L. Mann]*
  - firms as foragers: exploration vs. exploitation *[with V. Carvalho]*
  - the talent origins of top firms *[with T. Ifergane]*

Thank You!

Extra Slides

What's new?

Intro Conclusion

- **Firms: task-based microfoundation for complementarities**

  Firms & teams: *Lucas, 1978; Becker & Murphy, 1992; **Kremer, 1993**; Kremer & Maskin, 1996; **Garicano, 2000**; **Garicano & Rossi-Hansberg, 2006**; Porzio, 2017; **Jarosch et al., 2021**; Kuhn et al., 2023*

  Task assignment: *Costinot & Vogel, 2010; **Acemoglu & Restrepo, 2018**; Ocampo, 2021*

- **Sorting: parsimonious model of matching into teams with multi-dim. skill het.**

  Multi-dim. skill heterogeneity: *Kambourov-Manovskii, 2008; Gathman-Schoenberg, 2010; Lindenlaub, 2017; Guvenen et al., 2020; Lise & Postel-Vinay, 2020; Baley et al., 2022; Grigsby, 2024; Rubbo, 2024*

  Frictional matching: *Shimer & Smith, 2000; Cahuc et al., 2006; Eeckhout & Kircher, 2011/2018; Hagedorn et al., 2017; de Melo, 2018; Lindenlaub & Postel-Vinay, 2023; **Herkenhoff et al., 2024**; Bandiera et al., 2024*

- **Wage inequality: technological explanation for ↑ firm-level inequality**

  Technology: *Katz & Murphy, 1992; Krusell et al., 2000; Autor et al., 2003; Acemoglu & Restrepo, 2018*

  Firms: *__Card et al., 2013__; Barth et al., 2016; Alvarez et al., 2018; __Bloom et al., 2019__; Sorkin & Wallskog, 2023*

What's the value-added of the micro-founded production function?

- **Concern:** the microfoundation isn't used for measurement — i.e. measure $z_i(\tau)$'s directly and then 'aggregate up' to recover complementarities – so what's the point?

- **Value-added #1**: tractable model of team production with multi-dimensional skills
  - reduces dimensionality of matching into team with multi-d. skills

- **Value-added #2**: relative to a r-f CES fn. with 1-dim. skill *[e.g. Herkenhoff et al., 2024]*
  1. explanation for why talent complementarities exist & may change over time
  2. the two models are not observationally equivalent
     - benefit from team production is also increasing with χ, hence this term co-moves with talent complementarities (and it affects sorting differently)
     - selection effects due to ξ: when we observe low and high x workers together, they are likely to be a good match in terms of their task-specific skills *[cf. Borovickova-Shimer, 2024]*

‘

Spatial model (1)

Main

- Consider a non-negative random field $\{Z(s) : s \in \mathcal{S}\}$, where $\mathcal{S}$ is a *cylinder:*

$$\mathcal{S} = \mathcal{R} \times \mathbb{R}^+, \quad \text{where } \mathcal{R} = [0, 1)$$

  and $r \in \mathcal{R}$ represents the circular coordinate, and $h \in \mathbb{R}^+$ represents the height

- $Z(s)$ is constructed as the product of a height-dependent scale parameter $x(h) > 0$

$$Z(s) = x(h)\tilde{Z}(r), \quad s = (r, h) \in \mathcal{S}.$$

  where $\tilde{Z}(r)$ is a **Brown-Resnick max-stable process**
  - w.l.o.g., work with unit shape parameter *[Resnick, 1987]*

Spatial model (2): Brown-Resnick process

- de Haan spectral representation: the Brown-Resnick process can be written as

$$\tilde{Z}(r) = \max_{i \geq 1} \zeta_i W_i(r), \quad r \in \mathcal{R}.$$

  - $\{\zeta_i\}$ are points of a **Poisson process** on $(0, \infty)$ with intensity $\zeta^{-2} d\zeta$.
  - $W_i(r)$ are **independent copies of the spectral function**, defined as:

$$W(r) = \exp\{\varepsilon(r) - \gamma(r)\}.$$

  - $\varepsilon(r)$ is a **stationary Gaussian process with mean zero and stationary increments**
  - semivariogram $\gamma(r_1, r_2)$ determines dependence and is isotropic:

$$\gamma(r_1, r_2) = \gamma(d(r_1, r_2))$$

  e.g.

$$\gamma(d) = \lambda d^{\kappa}, \quad \lambda > 0, \quad 0 < \kappa \leq 2$$

- Equip circle with a distance function $d : \mathcal{R} \times \mathcal{R} \to \mathbb{R}^+$

$$d(r_1, r_2) = \min(|r_1 - r_2|, 1 - |r_1 - r_2|).$$

Spatial model (3): finite-dimensional distributions

- By construction, the process $\tilde{Z}(r)$ has **unit Fréchet** marginals:

$$P(\tilde{Z}(r) \leq z) = \exp(-z^{-1}), \quad z > 0.$$

- For any pair locations $\{s_1, s_2\} \subset \mathcal{S}$ the bivariate exponent function is

$$V(z_1, z_2) = \frac{1}{z_1}\Phi\left(\frac{1}{a}\log\frac{z_2}{z_1} + \frac{a}{2}\right) + \frac{1}{z_2}\Phi\left(\frac{1}{a}\log\frac{z_1}{z_2} + \frac{a}{2}\right),$$

where $a^2 = \gamma(d(r_1, r_2))$.

- NB: for a max-stable random vector with unit Fréchet margins, the joint dist is

$$\Pr(Z_1 \leq z_1, \ldots, Z_d \leq z_d) = \exp\{-V(z_1, \ldots, z_d)\}, \quad z_i > 0,$$

where $V(z_1, \ldots, z_d)$ is called the exponent function and is homogeneous of order -1.
- Higher-dimensional V *[Huser and Davison, 2013]*

Spatial model (4): approximation argument

- The bivariate margins of the Brown-Resnick process are of the Hüsler–Reiss type, rather than Gumbel-Hougaard
- *However,* in finite samples, the two copulas are statistically indistinguishable *[Genest et al., 2011]*
- **Therefore:**
  1. we can model the joint skill distribution as a Brown-Resnick process
  2. we can approximate the bivariate copula using an analytically convenient Gumbel-Hougaard copula
- Note also: the optimal task assignment can be solved for *any* max-stable process/e-v copula/finite-dim. multivariate distribution. But the power-law structure of the G-H is critical to obtain *explicit formula* for $f(\cdot)$

Spatial model (4b): approximation argument

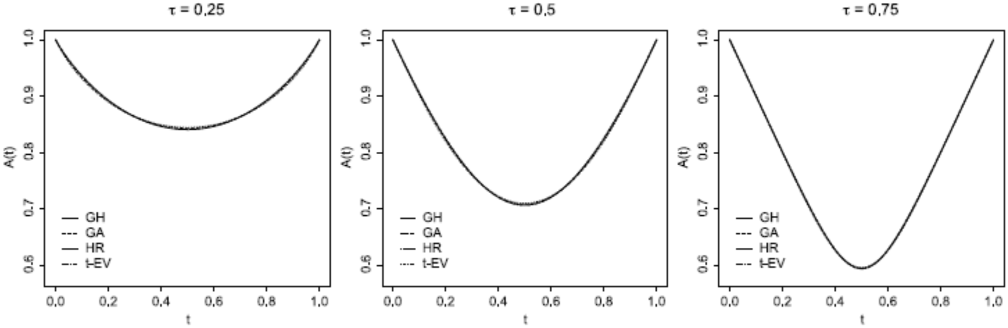


Figure 2. Pickands dependence functions of the Gumbel–Hougaard, Galambos, Hüsler–Reiss and t-EV copulas when $\tau = 0.25$, $\tau = 0.50$ and $\tau = 0.75$.

Production setting: $n \in \mathcal{Z}$

▸ Main

- Main setting: closed-form expression for $f(\cdot)$ assuming $n = 2$
- Can construct aggregation result for any $n \in \mathbb{Z}_{++}$ as long as skills at the population level are the realizations of a max-stable processes
  - Brown–Resnick process defined on a cylinder
- *Closed-form* expression for $f(\cdot)$ requires that the finite-dimensional distributions are multivariate GEV with Fréchet marginals and Gumbel-Hougaard copula
- Beyond the $n = 2$ case this works for $n > 2$ if $\xi_{il} = \xi$ for any $i \neq l$

$$Y = f(\mathbf{x}, \xi; \chi) = \left(\sum_{i=1}^{n} x_i^{\frac{1}{\chi\xi+1}} \right)^{\chi\xi+1}$$

Lemma

Main

Lemma: Lemma

Implied task share and shadow-cost index equal

$$\pi_i = \frac{\left(x_i/\lambda_i^L\right)^{\frac{1}{\chi\xi}}}{\sum_{k=1}^{n}\left(x_i/\lambda_i^L\right)^{\frac{1}{\chi\xi}}} \quad x; \lambda = \left(\sum_{i=1}^{n}\left(\frac{x_i}{\lambda_i^L}\right)^{\frac{1}{\chi\xi}}\right)^{-\chi\xi}$$

Intuition: features of optimal organization

- **What is the intuition for these properties?**
- Solution of firm's mini-planner problem implies:

1. Complete division of labor, with tasks assigned by comparative advantage
   - i's task set $\mathcal{T}_i = \left\{ \tau \in \mathcal{T} : \frac{z_i(\tau)}{\lambda_i^L} \geq \max_{k \neq i} \frac{z_k(\tau)}{\lambda_k^L} \right\}$
   - classic source of efficiency gains
2. **i's share of tasks $\uparrow$ in i's talent, $\downarrow$ in coworkers' talent**
   - i's task share $\pi_i = \left(x_i^{\frac{1}{1+\chi\xi}}\right)\left(\sum_{k=1}^{n}(x_k)^{\frac{1}{1+\chi\xi}}\right)^{-1}$

Intuition: comparative statics for task shares

Main

- Suppose that $x_i > x_j$. Then

  1. i performs a strictly larger share of tasks than j for $\chi < \infty$

Intuition: comparative statics for task shares

- Suppose that $x_i > x_j$. Then
  1. i performs a strictly larger share of tasks than j for $\chi < \infty$
  2. **the difference in task shares is decreasing in χ**

$\Rightarrow$ **Greater skill specialization implies a larger share of tasks is performed by relatively less talented team members** – more talented coworkers can't easily compensate

Surplus sharing protocol

Main

- The wage of a worker of type x employed alone satisfies

$$(1-\omega)\big(V_{e.1}(x) - V_u(x)\big) = \omega\big(V_{f.1}(x) - V_{f.0}\big), \tag{5}$$

- The wage $w(x|x', \xi)$ of a type-x worker with a coworker of type x' given shock ξ satisfies

$$(1-\omega)\big(V_{e.2}(x|x', \xi) - V_u(x)\big) = \omega\big(V_{e.2}(x'|x, \xi) + V_{f.2}(x, x', \xi) - V_{e.1}(x') - V_{f.1}(x')\big). \tag{6}$$

HJB: unmatched

Main

- Unmatched firm:

$$\rho V_{f.0} = (1-\omega)\lambda_{v.u} \int \frac{d_u(x)}{u} S(x)^+ dx, \tag{7}$$

- Unmatched worker:

$$\rho V_u(x) = b(x) + \lambda_u \omega \left[\frac{d_{f.0}}{v} S(x)^+ + \int \int \frac{d_{m.1}(\tilde{x}')}{v} S(x|\tilde{x}', \tilde{\xi})^+ dH(\tilde{\xi}) d\tilde{x}' \right] \tag{8}$$

Joint values

- Joint value of firm with x and x', ξ

$$\rho\Omega_2(x, x', \xi) = f_2(x, x', \xi) - \delta S(x|x', \xi) - \delta S(x'|x, \xi) \tag{9}$$

- Joint value of firm with x

$$\begin{aligned}\rho\Omega_1(x) &= f_1(x) + \delta\big[-\Omega_1(x) + V_u(x) + V_{f.0}\big] \\ &\quad + \lambda_{v.u} \int\int \frac{d_u(\tilde{x}')}{u} \Big(\underbrace{-\Omega_1(x) + V_{e.2}(x|\tilde{x}', \tilde{\xi}) + V_{f.2}(x, \tilde{x}', \tilde{\xi})}_{(1-\omega)S(\tilde{x}'|x,\tilde{\xi})}\Big)^{+} dH(\tilde{\xi}) d\tilde{x}'.\end{aligned} \tag{10}$$

HJB: surpluses

- Surplus of coalition of firm with worker x

$$(\rho+\delta)S(x) = f_1(x) - \rho\big(V_u(x) + V_{f.o}\big) + \lambda_{v.u}(1-\omega)\int \frac{d_u(\tilde{x}')}{u} S(\tilde{x}'|x,\tilde{\xi})^+ dH(\tilde{\xi})\tilde{x}'. \quad (11)$$

- Surplus from adding x to x' with xi

$$S(x|x',\xi)\big(\rho+2\delta\big) = f_2(x,x',\xi) - \rho\big(V_u(x) + V_u(x') + V_{f.o}\big) + \delta S(x) - (\rho+\delta)S(x'). \quad (12)$$

KFE: unemployed

$$\delta\left(d_{m.1}(x) + \int d_{m.2}(x,\tilde{x}')d\tilde{x}'\right) = d_u(x)\lambda_u\left(\int \frac{d_{f.0}}{v}h(x,\tilde{y}) + \int \frac{d_{m.2}(\tilde{x}')}{v}h(x|\tilde{x}')d\tilde{x}'\right). \tag{13}$$

KFE: one-worker matches

$$d_{m.1}(x)\left(\delta + \lambda_{v.u} \int \frac{d_u(\tilde{x}')}{u} h(\tilde{x}'|x) d\tilde{x}'\right) = d_u(x)\lambda_u \frac{d_{f.0}}{v} h(x) + \delta \int d_{m.2}(x, \tilde{x}') d\tilde{x}'. \qquad (14)$$

KFE: two-worker matches

$$2\delta d_{m.2}(x,x') = d_u(x)\lambda_u \frac{d_{m.1}(x')}{v} h(x|x') + d_u(x')\lambda_u \frac{d_{m.1}(x)}{v} h(x'|x). \tag{15}$$

Matching – stationary equilibrium

Main

- HJ-Bellman equations $\rightarrow$ **values & matching policies** HJBs
- Flows between/**distribution** over types $\times$ employment states KFEs

Definition: Stationary equilibrium

A stationary eqm. consists of a production function, value functions & a distribution of agents, s.t.

1. the production function is consistent with the optimal assignment of tasks;
2. the value functions satisfy the HJB equations given the distribution;
3. the distribution is stationary given the policy fn's implied by the value fn's.

A useful identification result

Monte Carlo: χ | Identification validation | Main

- **Challenge:** skill specificity χ *not* directly observable
  - evidence for task-specific skills *[cf. Deming, 2023]* but no cardinal measure of specificity
  - inferring χ from observed sorting patterns could load too much onto χ
- **Structural identification:** χ identifiable from $w(x|x')$ given x and x'

Sketch | Equation for $\bar{w}(x|x')$

$$\boxed{\chi \longrightarrow \frac{\partial^2 f(x, x', \xi)}{\partial x \partial x'} \propto \frac{\partial^2 w(x|x', \xi)}{\partial x \partial x'}.}$$

- Motivates measuring $\dfrac{\partial^2 \bar{w}(x|x')}{\partial x \partial x'}$

Reduced-form regression to identify χ (2010-2017)

- Approximate $\frac{\partial^2 \bar{w}(x|x')}{\partial x \partial x'}$ using **regression with interaction term**

$$\frac{w_{it}}{\overline{w}_t} = \beta_0 + \sum_{d=2}^{10} \beta_{1d} \mathbf{1}\{\hat{x}_i = d\} + \sum_{d'=2}^{10} \beta_{2d'} \mathbf{1}\{\hat{x}_{-it} = d'\} + \beta_c (\hat{x}_i \times \hat{x}_{-it}) + \psi_{j(i,t)} + \nu_{o(i,t)t} + \xi_{s(i,t)t} + \epsilon$$

- **Reduced-form estimate:** $\hat{\beta}_c$ = 0.0063
  - robust: schooling as non-wage measure, small teams, lagged types, excl managers, ...

Reg. table

Long robustness list (it's a JMP...)

Reduced-form regression to identify χ (2010-2017)

- Approximate $\frac{\partial^2 \bar{w}(x|x')}{\partial x \partial x'}$ using **regression with interaction term**

$$\frac{w_{it}}{\overline{w}_t} = \beta_0 + \sum_{d=2}^{10} \beta_{1d} \mathbf{1}\{\hat{x}_i = d\} + \sum_{d'=2}^{10} \beta_{2d'} \mathbf{1}\{\hat{x}_{-it} = d'\} + \beta_c(\hat{x}_i \times \hat{x}_{-it}) + \psi_{j(i,t)} + \nu_{o(i,t)t} + \xi_{s(i,t)t} + \epsilon$$

- **Reduced-form estimate:** $\hat{\beta}_c$ = 0.0063

  Reg. table

  - robust: schooling as non-wage measure, small teams, lagged types, excl managers, ...

  Long robustness list (it's a JMP...)

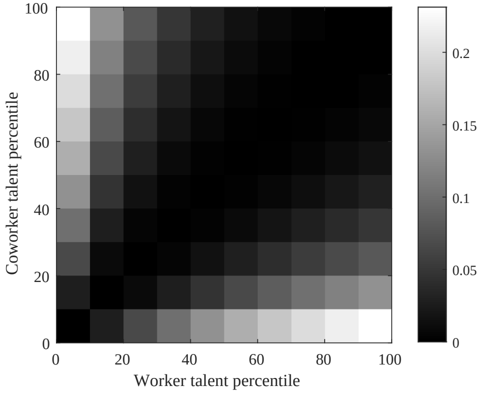
- **Estimation of structural model:** replicate semi-structural regression with model-generated data, infer χ from matching empirical $\hat{\beta}_c$

  Parameter values

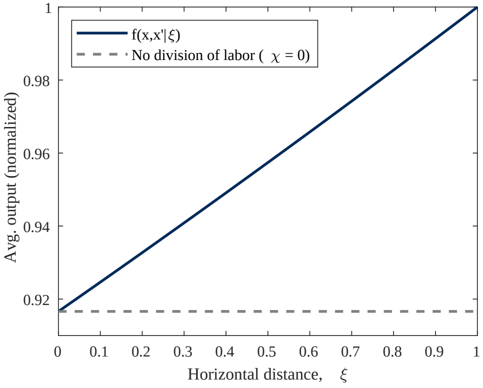
Properties of the estimated r.-f. production function

JH validation

(a) Talent complementarity



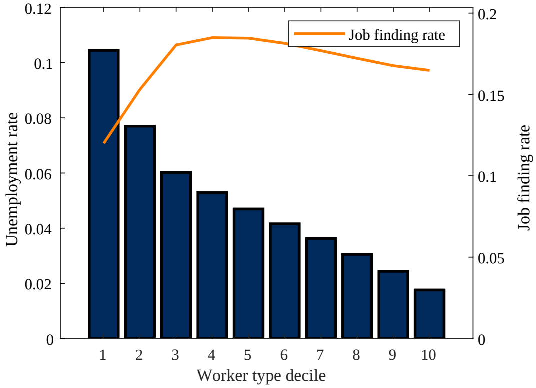
(b) Horizontal distance



Macro properties (untargeted): job finding/separation rates

▸ Parameter values

1. ✓Higher-x workers experience lower unemp. rates due to lower separation rates but job finding rates don't increase much with talent *[e.g., Cairo & Cajner, 2018]*



Macro properties (untargeted): firm-level wage dispersion

1. ✓Higher-x roworkers experience lower unemp. rates due to lower separation rates but job finding rates don't increase much with talent *[e.g., Cairo & Cajner, 2018]*
2. ✓**Match coworker sorting patterns**
   - $\rho_{xx} = 0.45$ (vs. 0.64 in data)

▸ Avg. coworker figure

3. ✓**Between-firm wage inequality**
   - between-share 0.55 (vs. 0.57 in data)
   - mirrors endogenous firm-level productivity dispersion

▸ Figure

$\Rightarrow$ **Model endogenously generates ex-post heterogeneity among ex-ante identical firms**

A useful identification result

Monte Carlo: χ | Identification validation | Selection

- **Challenge:** skill specificity χ *not* directly observable
  - evidence for task-specific skills *[cf. Deming, 2023]* but no cardinal measure of specificity
  - inferring χ from observed sorting patterns could load too much onto χ
- **Structural identification:** Proposition 1 monotonically relates χ to $\frac{\partial^2 f(\cdot)}{\partial x \partial x'}$, which we c from $w(x|x')$ given x and x'
  - intuition: outside options influence *level* of w *[Eeckhout-Kircher, 2011]* but enter separably

Sketch | Equation for $\bar{w}(x|x')$

$$\frac{\partial^2 f(x, x', \xi)}{\partial x \partial x'} \propto \frac{\partial^2 w(x|x', \xi)}{\partial x \partial x'}.$$

- Motivates measuring $\frac{\partial^2 \bar{w}(x|x')}{\partial x \partial x'}$

Reduced-form regression to identify χ (2010-2017)

- Approximate $\frac{\partial^2 \bar{w}(x|x')}{\partial x \partial x'}$ using **regression with interaction term**

$$\frac{w_{it}}{\overline{w}_t} = \beta_0 + \sum_{d=2}^{10} \beta_{1d} \mathbf{1}\{\hat{x}_i = d\} + \sum_{d'=2}^{10} \beta_{2d'} \mathbf{1}\{\hat{x}_{-it} = d'\} + \beta_c (\hat{x}_i \times \hat{x}_{-it}) + \psi_{j(i,t)} + \nu_{o(i,t)t} + \xi_{s(i,t)t} + \epsilon$$

- **Reduced-form estimate:** $\hat{\beta}_c$ = 0.0063

Reg. table

  - robust: schooling as non-wage measure, small teams, lagged types, excl managers, ...

Long robustness list (it's a JMP...)

Reduced-form regression to identify χ (2010-2017)

- Approximate $\frac{\partial^2 \bar{w}(x|x')}{\partial x \partial x'}$ using **regression with interaction term**

$$\frac{w_{it}}{\overline{w}_t} = \beta_0 + \sum_{d=2}^{10} \beta_{1d} \mathbf{1}\{\hat{x}_i = d\} + \sum_{d'=2}^{10} \beta_{2d'} \mathbf{1}\{\hat{x}_{-it} = d'\} + \beta_c (\hat{x}_i \times \hat{x}_{-it}) + \psi_{j(i,t)} + \nu_{o(i,t)t} + \xi_{s(i,t)t} + \epsilon$$

- **Reduced-form estimate:** $\hat{\beta}_c$ = 0.0063

Reg. table

  - robust: schooling as non-wage measure, small teams, lagged types, excl managers, ...

Long robustness list (it's a JMP...)

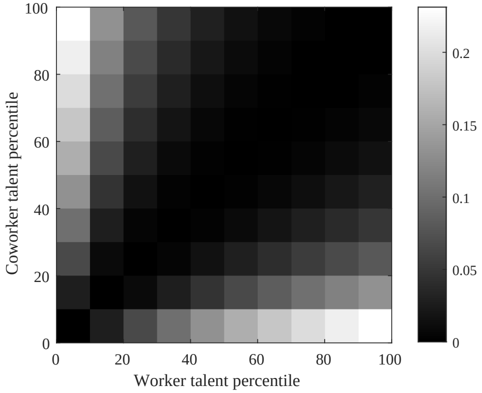
- **Estimation of structural model:** replicate semi-structural regression with model-generated data, infer χ from matching empirical $\hat{\beta}_c$

Parameter values

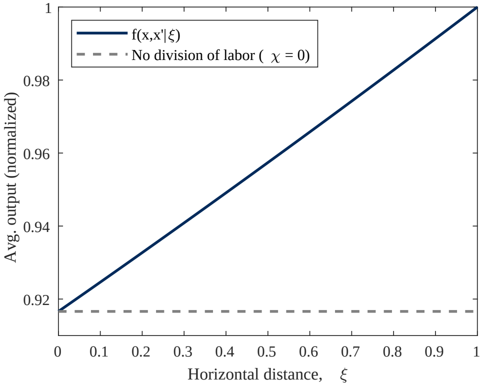
Properties of the estimated r.-f. production function

JH validation

(a) Talent complementarity



(b) Horizontal distance



Mapping theory to data: worker & coworker types

▸ Main

- **Theory:** wage monotonically $\uparrow$ in x, so higher types have higher expected/lifetime earnings
- **Implementation:** standard methods
  - worker fixed effect (FE) in Mincerian wage regression
    – baseline: AKM [Abowd et al., 1999] with pre-est. k-means clustering to address limited mobility bias [Bonhomme et al., 2019]

$\Rightarrow$ **Worker i's talent type $\hat{x}_i$: decile rank of i's FE**

  - baseline: economy-wide rank; robustness: within 2d-occupation

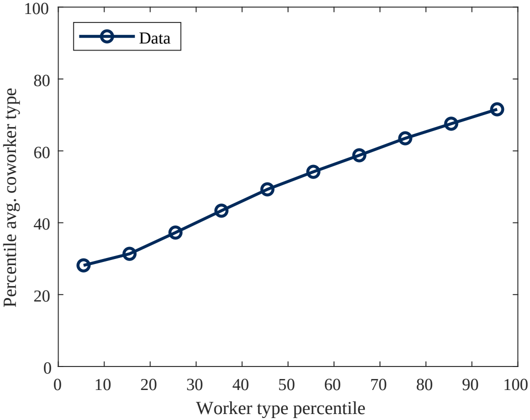
- **"Representative coworker type"** $\hat{x}_{-it}$: avg. $\hat{x}$ of workers in same estab.-yr.

Mapping model to data: coworker types

- Defining $S_{-it} = \{k : j(kt) = j(it), k \neq i\}$ as the set of i's coworkers in year t, compute the average type of i's coworkers in year t as $\hat{x}_{-it} = \frac{1}{|S_{-it}|} \sum_{k \in S_{-it}} \hat{x}_k$.
- **Coworker group:**
  - alternative: same establishment-occupation-year cell
  - but CC arise precisely when workers have *differentiated* task-specific skills
- **Averaging step:**
  - equally-weighted averaging ignores non-linearity in coworker aggregation
  - paper: show using non-linear averaging method that baseline results in bias, but it's minor in magnitude
- **Firm size variation:** averaging ensures that a single move will induce a smaller change in the *average* coworker quality in a large team than in a small one

Mapping theory to data: talent sorting in the data

- Measures of $\hat{x}_i$ and $\hat{x}_{-it}$ sufficient to measure empirical talent sorting



Measurement: a useful identification result

Main

- **Q:** How to quantify $\frac{\partial^2 f(x,x')}{\partial x \partial x'}$?
- **Proposition:** production complementarities are proportional to wage compl.
- **Proof sketch:** wage level for worker x with coworker x'

$$w(x|x',\xi) = \omega\big(f(x,x',\xi) - f(x')\big) + (1-\omega)\rho V_u(x) - \omega(1-\omega)\lambda_{v.u} \int\int \frac{d_u(\tilde{x}'')}{u} S(\tilde{x}''|x',\tilde{\xi})^+ dH(\tilde{\xi})$$

$$= \omega f(x,x',\xi) + g(x) - h(x')$$

where $g : [0,1] \to \mathbb{R}$ and $h : [0,1] \to \mathbb{R}$ are strictly increasing

$\Rightarrow$ *outside options* are *separable*: affect *level* of wage but not the *cross-partial*

- Integrating over ξ using optimal decision rules $h(\cdot) \Rightarrow$ average *realized* wage

Expected (log) wage level

- Expected wage, given threshold $\bar{\xi}$ and cond. exp. value $\xi^*(k) = \frac{\int_k^1 \xi dH(\xi)}{1-H(k)}$

$$\overline{w}(x|x') = \mathbb{E}_\xi\left[w(x|x',\xi)\right] \underbrace{\frac{d_u(x)\lambda_u \frac{d_{m.1}(x')}{v} h(x|x')}{d_u(x)\lambda_u \frac{d_{m.1}(x')}{v} h(x|x') + d_u(x')\lambda_u \frac{d_{m.1}(x)}{v} h(x'|x)}}_{p(x|x')} \times w\left(x|x', \xi^*(\bar{\xi}(x|x'))\right)$$

$$+ \frac{d_u(x')\lambda_u \frac{d_{m.1}(x)}{v} h(x'|x)}{d_u(x)\lambda_u \frac{d_{m.1}(x')}{v} h(x|x') + d_u(x')\lambda_u \frac{d_{m.1}(x)}{v} h(x'|x)} \times w\left(x|x', \xi^*(\bar{\xi}(x'|x))\right).$$

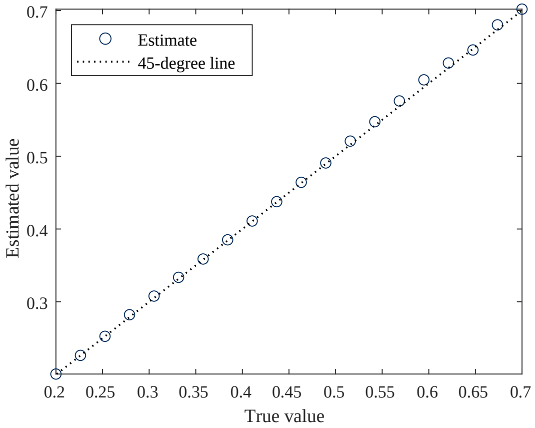
- Expected log wage, with $B^\xi(x|x') = \{\xi : S(x|x',\xi) > 0\}$

$$\mathbb{E}_\xi\left[\ln w(x|x',\xi)\right] = \overline{\ln w}(x|x') = p(x|x') \times \left(\frac{1}{1-h(x|x')} \times \int_{\xi \in B^\xi(x|x')} \ln w(x|x',\xi) dH(\xi)\right)$$

$$+ p(x'|x) \times \left(\frac{1}{1-h(x'|x)} \times \int_{\xi \in B^\xi(x'|x)} \ln w(x|x',\xi) dH(\xi)\right),$$

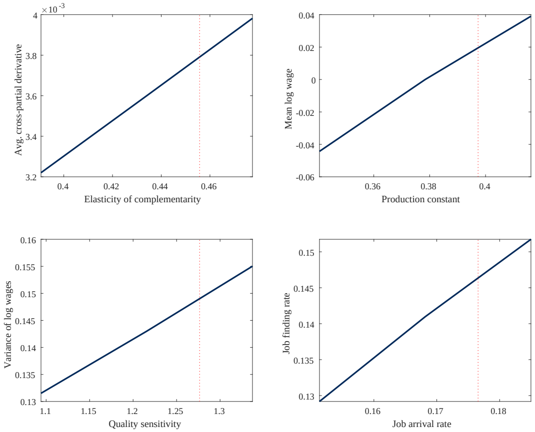
Monte Carlo study: identifying χ

Main



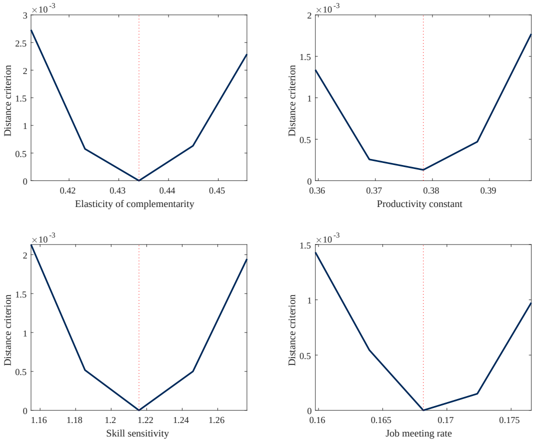
Identification validation exercise 1

▸ Main



Notes. This figure plots the targeted moment against the relevant parameter, holding constant all other parameters.

Identification validation exercise 2



Notes. This figure plots the distance function $\mathcal{G}(\psi_i, \psi^*_{-i})$ when varying a given parameter ψ_i around the estimated value ψ^*_i. The remaining parameters are allowed to adjust to minimize $\mathcal{G}$.

Regression estimates

Main

	(1)	(2)	(3)	(4)	(5)
Interaction coefficient ($\hat{\beta}_c$)	0.0067***	0.0067***	0.0063***	0.0063***	0.0059***
	(0.0005)	(0.0004)	(0.0005)	(0.0005)	(0.0008)
Employer FEs	No	No	Yes	Yes	Yes
Industry-year FEs	No	Yes	No	Yes	Yes
Occupation-year FEs	No	No	Yes	Yes	Yes
Type ranking	Economy	Economy	Economy	Economy	Occupation
Obs. (1000s)	3,606	3,606	3,606	3,606	3,606
Adj. R^2	0.788	0.800	0.801	0.813	0.769

Notes. Employer-clustered standard errors are given in parentheses. Observations are weighted by the inverse employment share of the respective type and (rounded) coworker type cell. Observation count rounded to 1000s. * p<0.1; ** p<0.05; *** p<0.01.

Robustness: reduced-form coworker complementarity

Main

- Types from non-parametric ranking algorithm instead of AKM-based Jump
- Schooling as a non-wage measure of types Jump
- Lagged types Jump
- Small teams Jump
- Movers Jump
- Non-parametric, finite-differences approximation Jump
- Excluding managers Jump
- Log specification Jump

Coworker complementarity: lagged types

Robustness overview

- Concern with both regression approach and non-parametric FD approach: mechanical relationship between wages ("LHS") and (within-period time-invariant) worker types, which are estimated from wages themselves ("RHS")
- Robustness check #1: years of schooling as type measure Jump
- Robustness check #2: assign to each individual i in periods $p \in \{2, 3, 4, 5\}$ the FE estimated for i in period $p - 1$; re-compute worker deciles and average coworker types, $\hat{x}_i^{p-1}$ and $\hat{x}_{-it}^{p-1} = (|S_{-it}|)^{-1} \sum_{k \in \mathcal{S}} \hat{x}_k^{p-1}$; re-estimate wage regression
- Results (see paper): magnitude of estimated $\hat{\beta}_c$ around 50% smaller when using lagged types, but evolution over time similar to baseline

Complementarity estimates using years of schooling

Robustness overview

	'85-'92	'93-'97	'98-'03	'04-'09	'10-'17
Interaction	0.0063***	0.0060***	0.0099***	0.0112***	0.0129***
	(0.0008)	(0.0007)	(0.0008)	(0.0007)	(0.0009)
Obs. (1000s)	3,613	2,508	2,694	3,836	4,376
R^2	0.5033	0.5451	0.5746	0.6330	0.6425

Notes. Dependent variable is the wage level over the year-specific average wage. Independent variables are a constant, years of schooling, coworker years of schooling, and the interaction between those two terms. All regressions include industry-year, occupation-year and employer fixed effects. Employer-clustered standard errors in parentheses. Observations are unweighted. The sample is unchanged from the main text, except that 96,517 observations with missing years of schooling are dropped. Observation count rounded to 1000ss.

Within-industry empirical analysis

Overview: robustness
Within-industry calibration

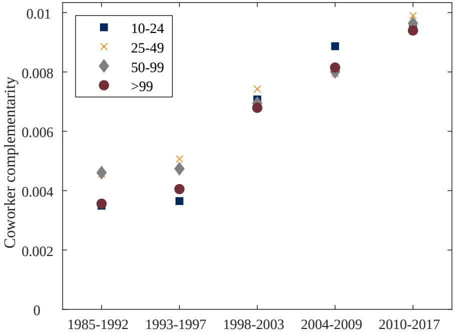
	Baseline				Within-industry avg.			
Sample Period	σ_w^2	$\sigma_{\bar{w}}^2/\sigma_w^2$	ρ_{xx}	$\hat{\beta}_c$	σ_w^2	$\sigma_{\bar{w}}^2/\sigma_w^2$	ρ_{xx}	$\hat{\beta}_c$
1	0.143	0.337	0.427	0.0036	0.125	0.249	0.333	0.00283
2	0.148	0.391	0.458	0.0042	0.125	0.288	0.351	0.00342
3	0.191	0.456	0.495	0.0070	0.150	0.324	0.369	0.00585
4	0.234	0.547	0.547	0.0082	0.168	0.388	0.405	0.00738
5	0.241	0.568	0.617	0.0091	0.171	0.412	0.464	0.00823

Notes. Within-industry avg. is person-year weighted average across OECD STAN-A38 (2-digit) industries.

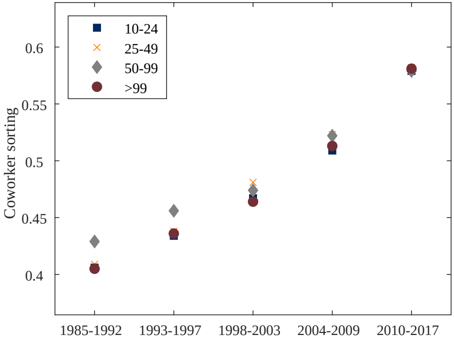
Coworker complementarity & sorting by team size

Robustness

(a) Complementarity



(b) Sorting



Sorting & complementarity based on non-parametric ranking algorithm

▸ Main

- Instead of ranking workers based on AKM worker FEs, use non-param. ranking algo *[Hagedorn et al., 2017]*

	Sorting		Complementarities	
Period	Spec. 1	Spec. 2	Spec. 1	Spec. 2
1985-1992	0.47	0.38	0.001	0.000
1993-1997	0.56	0.46	0.002	0.001
1998-2003	0.60	0.48	0.004	0.002
2004-2009	0.65	0.50	0.005	0.002
2010-2017	0.68	0.51	0.005	0.004

Notes. This table indicates, under the column "Sorting" the correlation between a worker's estimated type and that of their average coworker, separately for five sample periods. The column "Complementarities" indicates the point estimate of the regression coefficient β_C. Under "Specification 1" workers are ranked economy wide, while under "Specification 2" they are ranked within two-digit occupations. Worker rankings are based on the non-parametric method.

Coworker complementarity: excluding managers

▸ Robustness overview

- **Concern** regarding complementarity estimates: driven by managers?
  - only managers benefit from team quality, e.g. via larger span of control
  - the only coworkers that matter are managers

Period	Baseline	Exclude as recipients	Exclude entirely
1985-1992	0.0036***	0.0036***	0.0038***
1993-1997	0.0042***	0.0041***	0.0043***
1998-2003	0.0070***	0.0074***	0.0076***
2004-2009	0.0082***	0.0084***	0.0092***
2010-2017	0.0091***	0.0097***	0.0093***

Notes. Managed are defined based on KldB-1988-3d, as in Jarosch et al. (2023).

Coworker complementarity: movers

Robustness overview

- Consider sub-samples of job movers, job movers with contiguous employment spells ($t \to t+1$), and job movers with non-contiguous E spells ($t \to t+s$, $s > 1$)
- Caveat: annual panel given data size, no direct observation of U/N spells in SIEED

Period	Baseline	All movers	Contig. E spells	Non-contig. E spells
1985-1992	0.0043***	0.0043***	0.0045***	0.0039***
1993-1997	0.0049***	0.0052***	0.0052***	0.0051***
1998-2003	0.0078***	0.0085***	0.0083***	0.0082***
2004-2009	0.0090***	0.0107***	0.0104***	0.0102***
2010-2017	0.0088***	0.0103***	0.0101***	0.0090***
Obs. in '10-'17 (1000s)	4,410	538	355	375

Notes. Unweighted observations. Regressions include FEs for employer; occupation-year; industry-year. Employer-clustered standard errors in parentheses.

Parametrization (2010-2017)

Main

Parameter	Description	Value	Source	m	$\hat{m}$
ρ	Discount rate	0.008	External		
ω	Worker barg. weight	0.50	External		
δ_0	Sep. rate, constant	0.0147	Offline est.		
δ_1	Sep. rate, scale	-0.84	Offline est.		
$\bar{n}$	Team size	14	Offline est.		
χ	Skill specificity	1.17	Internal: β_c	0.0063	0.0063
a_0	Production, constant	0.26	Internal: normalized wage	1	1
a_1	Production, scale	1.49	Internal: Var. log wages	0.23	0.23
$\bar{b}$	Unemp. flow utility, scale	0.64	Internal: replacement rate	0.63	0.63
λ_u	Meeting rate	0.23	Internal: job finding rate	0.16	0.16

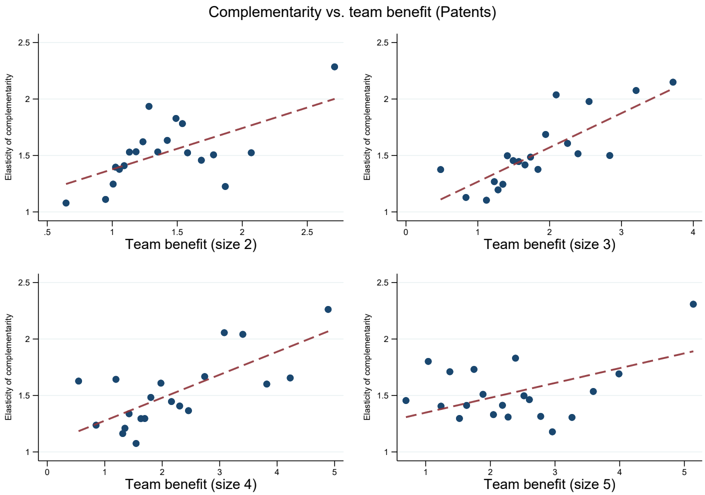
Model Meets Data: types and production function

$$f(x, x', \xi) = 2 \times \left(\frac{\bar{n}}{\bar{n}-1}\right)^{\chi\xi} \left(\frac{1}{2}(x)^{\frac{1}{\chi+1}} + \frac{1}{2}(x')^{\frac{1}{\chi+1}}\right)^{\chi+1}$$

1. Estimated 'talent types' are in *ordinal* space, $\tilde{x} \in [0,1]$. Mapping $x_i = a_0 + a_1\tilde{x}_i$
   - next iteration: allow for higher-order terms
   - (a_0, a_1) captures (i) "talent-biased technological change," and (ii) Δ talent distribution
     – nb: Hakanson et al (2021) find no evidence of $\uparrow$ dispersion in test scores
2. What the model treats as the second hire shows up, in the production function, as the $\bar{n}$-th hire
3. Baseline: ξ as $\sim$ match-specific shock that doesn't affect talent complementarities

Validation: Production functions estimated by Ahmadpoor-Jones (2019)

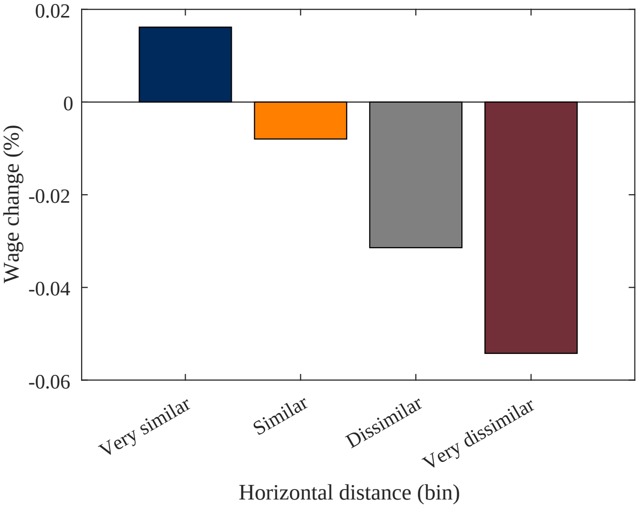
Main



Notes. Source data from Ahmadpoor and Jones (2019, PNAS). Own calculations. Binscatter plot for subsample with complementarity <= 5.

Validation: Structural interpretation of Jaeger-Heining (2022)

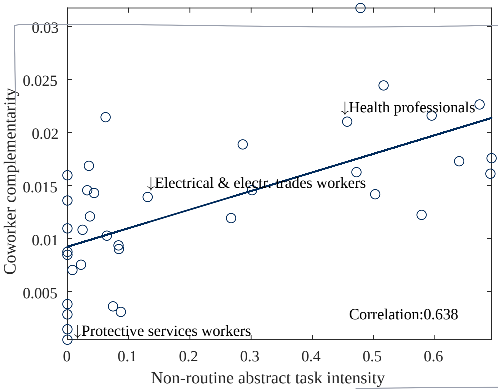
Main



X-sectional validation (occ's): tasks $\Rightarrow$ complementarity

Main

- $\uparrow$ **Non-routine abstract task intensity $\Rightarrow \uparrow$ coworker talent complementarity**



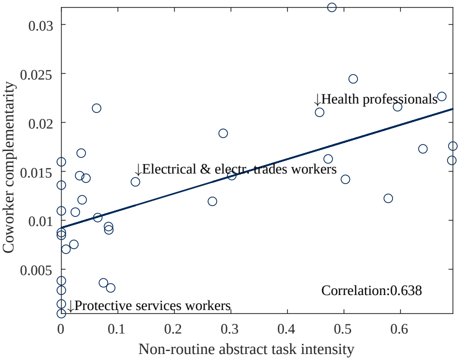
$\hat{\beta}_c$ estimated separately for each occupation

Non-routine abstract task intensity: proxy for χ

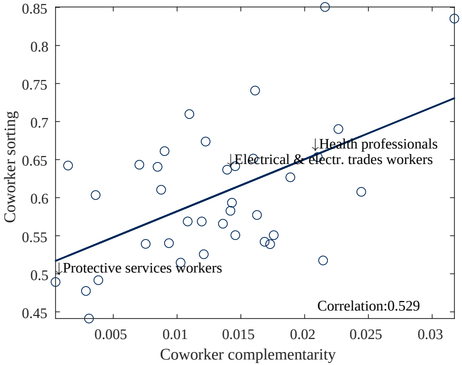
Notes. Quadros de Pessoal microdata. Analysis at ISCO-08-2d level.

X-sectional validation (occ's): tasks ⇒ complementarity ⇒ sorting

- ↑ Non-routine abstract task intensity ⇒ ↑ coworker talent complementarity



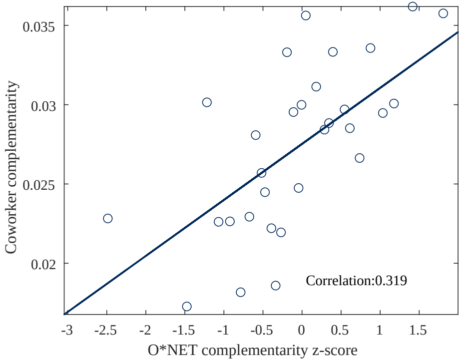
- **↑ Coworker talent complementarity ⇒ ↑ coworker sorting**



Industries: coworker importance ⇒ complementarity ⇒ sorting

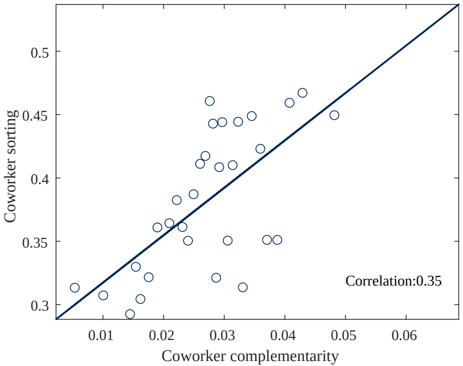
▸ Main

- ↑ **Teamwork** *[Bombardini et al., 2012]*
  ⇒ ↑ **coworker wage complementarity**



Notes. Horizontal axis measures the industry-level weighted mean score of an occupation-level index constructed from O*NET measuring the importance of: teamwork, impact on coworker output, communication, and contact.

- ↑ **Coworker wage complementarity**
  ⇒ ↑ **coworker sorting**



Notes. NACE-4-digit industries.

EE transitions in theory and data

▸ Validation overview

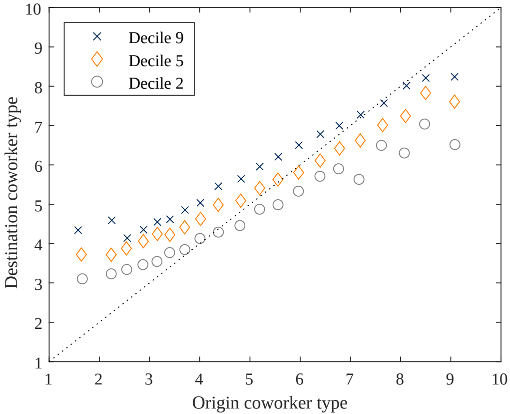
- **Theoretical prediction:** EE transitions move workers in surplus-maximizing direction $\Rightarrow \Delta\hat{x}_{-it} = \hat{x}_{-i,t} - \hat{x}_{-i,t-1}$ should be *positively* correlated with $\hat{x}_i$
  - $h_{2.1}(x, x''|x') = 1$ – worker x in a two-worker firm with coworker x'' would move to an employer that currently has one employee of type x' – if $S(x|x') - S(x|x'') > 0$
- **Empirical analysis**: use SIEED *spell* data to create worker-originMonth-destinationMonth-originJob-destinationJob panel, with information on characteristics of origin and destination job
  - subsample period 2008-2013 (huge panel at monthly frequency)
  - count as "EE" if employer change between two adjacent months
- **Regression analysis:** regress $\Delta\hat{x}_{-it}$, scaled by std. σ_Δ of coworker quality changes, on *own* type and *origin* coworker type

$$\frac{\Delta\hat{x}_{-it}}{\sigma_\Delta} = \beta_0 + \beta_1\hat{x}_i + \beta_2\hat{x}_{-i,t-1} + \epsilon_{it}$$

Empirical coworker sorting changes due to EE moves

Validation overview

- **EE transitions push toward greater coworker sorting:** for given origin, higher *x*-workers move to places with better coworkers than lower-*x* workers do
- Limitation: empirically, EE transitions "move up" low types more than theory predicts
- "**Coworker job ladder**" with both absolute and type-specific dimension?
- **Next:** change in the job ladder *[e.g., Haltiwanger-Spetzler, 2021]*



Evidence that EE *increasingly* reallocate toward PAM: in data & model

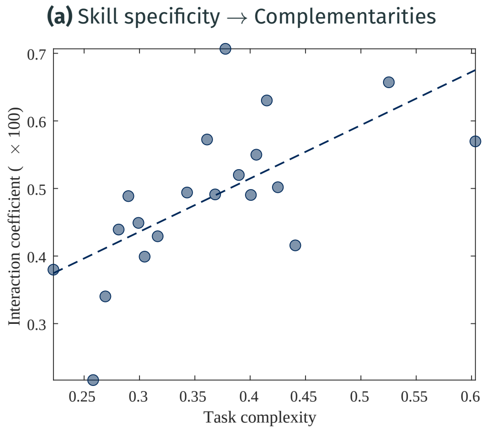
	Data		Model	
Change in coworker type	'85-'92	'10-'17	Period-1	Period-2
Own type	**0.0883*****	**0.118*****	**0.214**	**0.270**
	(0.000799)	(0.000918)		
Controls	Year FEs, Origin	Year FEs, Origin	Origin	Origin
N	196,098	282,718	∞	∞
adj. R^2	0.284	0.204		

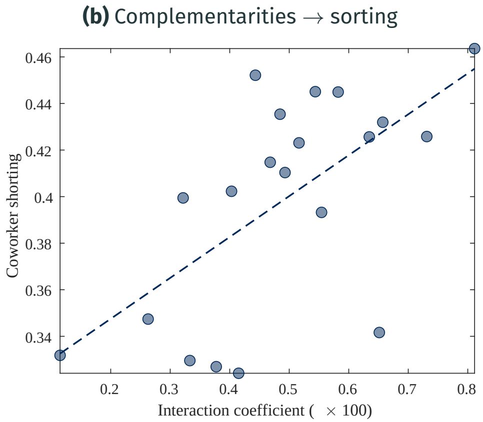
Table 1: Change in coworker type due to EE moves positively related to own type – increasingly so

Notes. For the data columns, individual-level clustered standard errors are given in parentheses. Model counterparts are computed simulation-free in population. Dependent variable is scaled throughout by the standard deviation of the change in coworker type.

Industry-level analysis: mechanisms, w/o industry FEs

Main

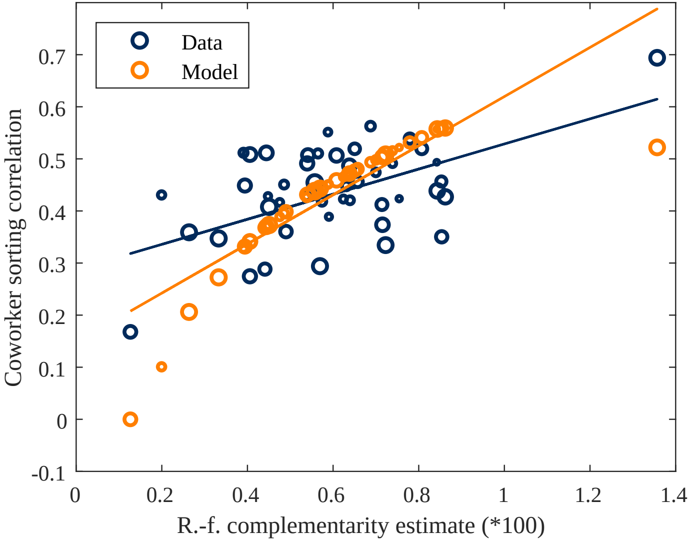




Notes. Binned scatterplots. Moments estimated separately for 2-digit industries over 5 sample periods. Includes period FEs. Data: SIEED + BIBB (task proxies).

Industry-level analysis: model vs. data

Main



Hypothesis: growing skill specificity ($\chi \uparrow$)

Main

1. Δ **Task composition:** fewer routine (low-χ), more complex (high-χ) tasks *[Deming, 2017]*

DE evidence

2. **Burden of knowledge**: increasing cost of reaching the frontier – necessitates increasingly narrow individual expertise *[Jones, 2009]*

Medical specialization

3. **Education:** if education augments task-specific skills randomly, then the trend toward more (secondary & tertiary) education fosters $\uparrow$ dispersed task-specific skills

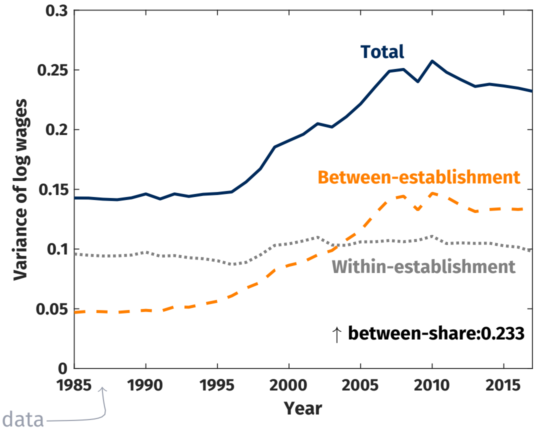
Formalization & edu data

Wage inequality has risen – and firms appear to play a key role

▸ Details

“the variance of firm [wages] explains an increasing share of total inequality in a range of countries”

[Song-Price-Guvenen-Bloom-von Wachter, 2019]

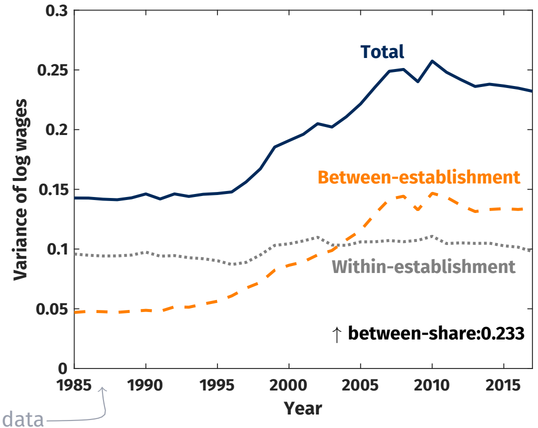


German matched employer-employee data

Applied question

Details Main

What are the causal driver(s)?



German matched employer-employee data

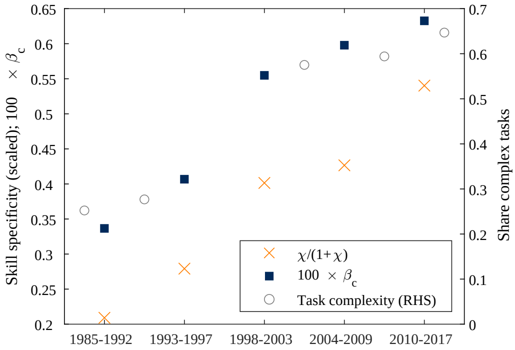
Overview of argument

1. The set of tasks any one worker can perform well has narrowed: **skill specificity** ↑
2. **Coworker talent complementarities** ↑
3. Workers of similar talent increasingly work together (**coworker sorting** ↑)
4. Greater ***firm-level*** **productivity & wage dispersion**

Estimate model for several periods: skill specificity ↑

▸ Schooling ▸ Peer effect trends

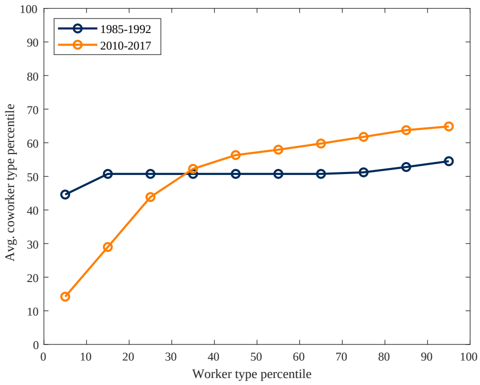
- **Method:** estimate reduced-form coefficient β_c for 5 sample periods
  $\Rightarrow$ re-estimate structural model
- **Skill specificity has intensified** ($\chi \uparrow$)
  [consistent with Grigsby's (2024) US estimates]
- Implied complementarities ↑
  - $\frac{f(x^{p80},x^{p80},1)+f(x^{p20},x^{p20},1)}{f(x^{p80},x^{p20},1)+f(x^{p80},x^{p20},1)}$: 1.05 ↗ 1.16



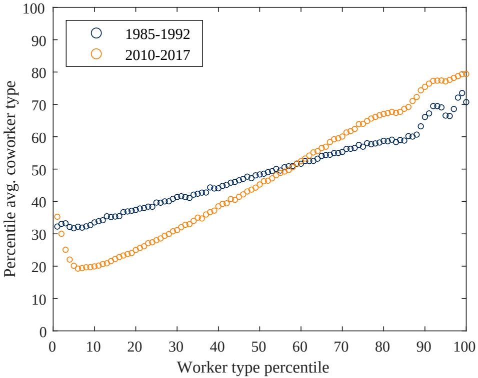
Talent sorting has intensified: theory & data

▸ Details ▸ Model Meets Data

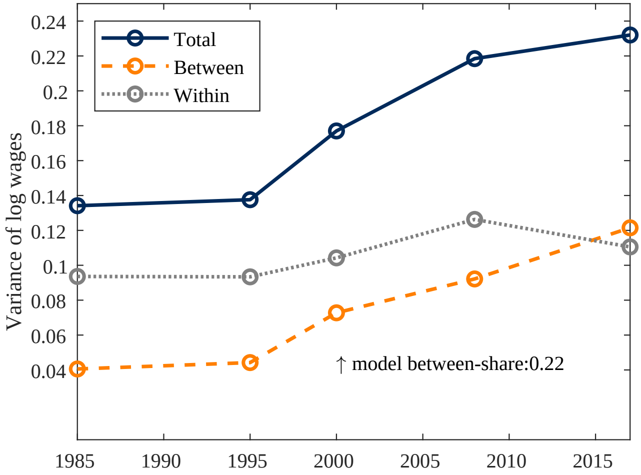
(a) Theory



(b) Data



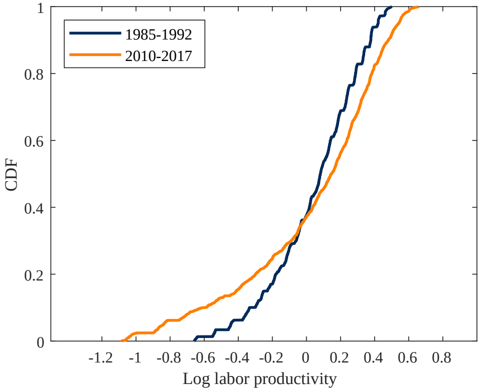
Model replicates observed ↑ firm-level wage inequality



Productivity dispersion

Main

- Firm dynamics literature: increased productivity dispersion *[Autor et al., 2020]*, correlated with wage & talent dispersion *[Berlingieri et al., 2017; Sorkin-Wallskog, 2020]*



Skill specificity $\chi \uparrow$ explains large share of between-share $\uparrow$

Main

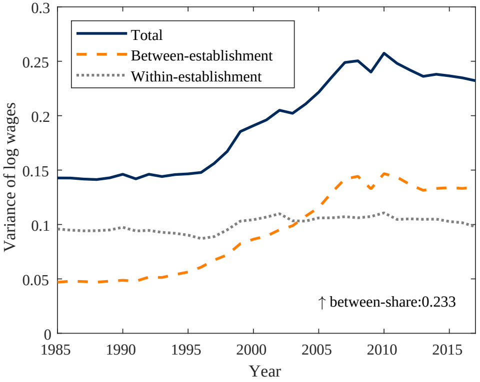
- **Q:** How much of $\uparrow$ between-firm share of wage var. is due to $\chi \uparrow$?
- **Counterfactual:** between-firm share in 2010s absent $\chi \uparrow$ since ’85-’92
- **A:** $\chi \uparrow$ **accounts for 65%** of model-predicted $\Delta \leftrightarrow \approx$ **59% of empirical** Δ
- **Robustness** exercises — Within-industry | Outsourcing
- Effect of $\downarrow$ search frictions *[e.g., Martellini-Menzio, 2021]* $\sim$ 11% of model-predicted Δ
  - search effort plausibly endogenous to χ

Fact #1: ↑ between-firm share of wage inequality

▸ Application ▸ Intro

- Large empirical literature: "firming up inequality" *[e.g., Card et al., 2013; Song et al., 2019]*
    - "superstar firms" *[e.g., Autor et al., 2020]*
- **Fact 1: ↑ wage inequality primarily due to between-component**
- Robust pattern

▸ Cross-country ▸ Panel est. ▸ Wage resid. alternatives
▸ Within-occ ▸ Within-ind



Notes. Model-free statistical decomposition, where the "between" component corresponds to the person-weighted variance of est.-level avg. log wage.

Why might χ have increased over time: schooling argument

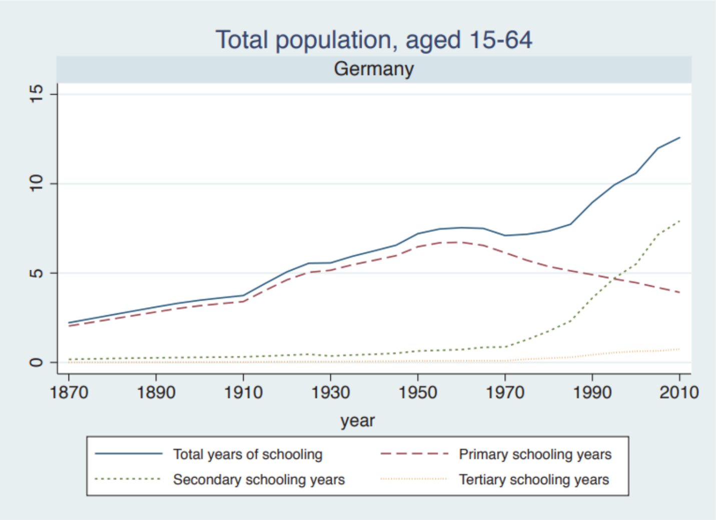
Main

- **Data:** trend toward more (secondary) education
- **Intuition:** if education augments task-specific skills randomly, then longer education leads to more dispersion in task-specific skills

Remark: Fréchet skill dispersion

Let Z be a Fréchet random variable (r.v.) with shape parameter $\theta > 0$ and scale parameter $x > 0$, and let $\{B_n\}_{n\geq 1}$ be a sequence of independent r.v.'s defined recursively as $B_n = \exp(-b_n/(\alpha\theta_{n-1}))$ where $\alpha \in (0,1)$, $\theta_0 = \theta$, $\theta_n = \theta_{n-1}\alpha = \theta\alpha^n$ for $n \geq 1$, $\{b_n\}_{n\geq 1}$ are independent r.v.'s such that $\exp(b_n/\alpha)$ are i.i.d. positive α-stable r.v.'s. Assume Z and $\{B_n\}$ are independent. Define the r.v.'s $\{Z^{(n)}\}_{n\geq 1}$ recursively as $Z^{(0)} = Z, \quad Z^{(n)} = Z^{(n-1)} \times B_n, \quad n \geq 1$. Then for each $n \geq 1$, $Z^{(n)}$ is a Fréchet r.v. with scale x and shape $\theta_n = \theta\alpha^n$.

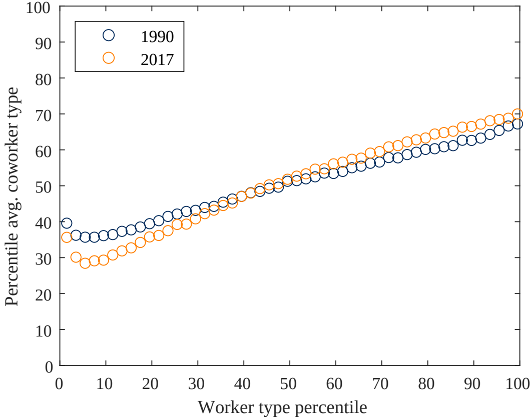
Barro Lee data for Germany



Evolution of coworker sorting: within-occupation ranking

Main

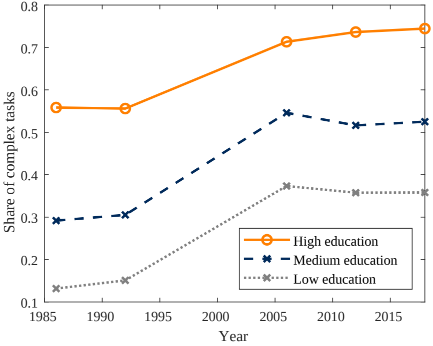
- The most talented within each occupation – the best engineer, PA, economist, manager, ... – tend to work together, and increasingly so



Task composition changes

▸ Main

- **Task complexity ↑**: "extensive margin" of χ
  - DE longitudinal task survey (BIBB)
  - "complex": cognitive non-routine (e.g., organizing, researching)



Workers increasingly tend to perform similar tasks across different jobs

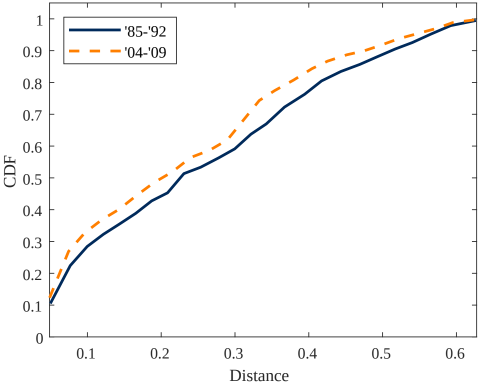
▸ Back

- ✓Workers move to jobs with similar tasks, rather than randomly

▸ Comparison

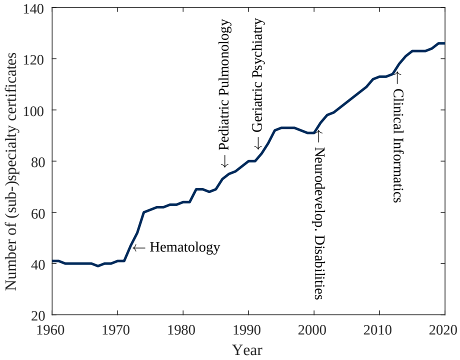
- **Q:** are workers becoming *more* likely to perform similar tasks across jobs?

- **Yes:** distribution of moves in ('04-'09) is stochastically dominated by that in ('85-'92)
  - uncond. average: 0.253 → 0.227: 10% decline
- Robust in regression design
  - quantile regressions



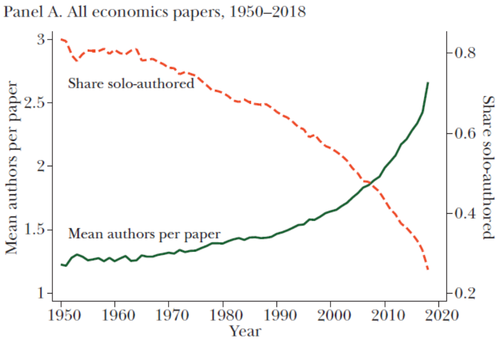
Examples: rising specialization

- **Deepening medical specialization**



Notes. Data from American Board of Medical Specialities. For each year, it shows the number of unique speciality or sub-speciality certificates that have been approved and issued at least once by that year and which are are still being issued.

- **Rise of research teams** *[Jones, 2021]*

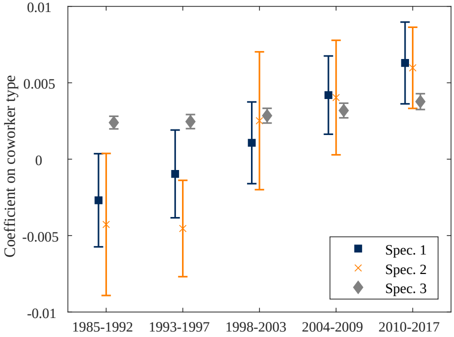


Coworker effects: log wage regression

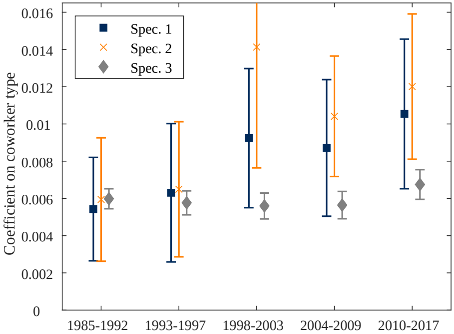
Back: cross-section Back: time series

$$\ln w_{it} = \beta_0 + \beta_1 \hat{x}_i + \beta_2 \hat{x}_{-it} + \psi_{j(it)} + \nu_{o(i)t} + \xi_{s(i)t} + \epsilon_{it}$$

(a) AKM types



(b) NP types

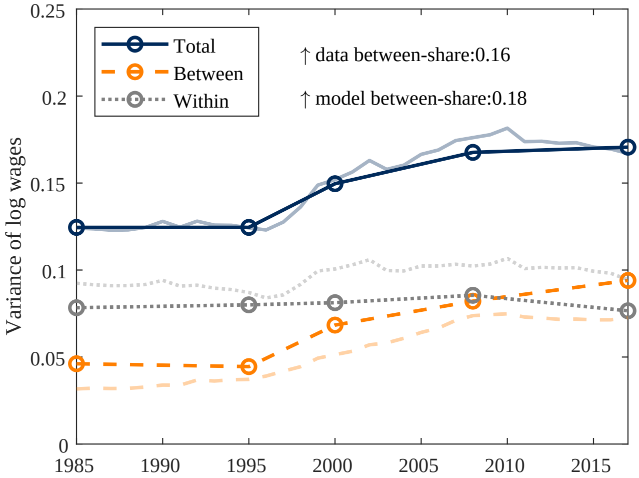


Notes. Specifications vary by ranking method – within-economy (spec. 1) vs. within-occupation (spec. 2/spec.3) and coworker group definition – establishment-year (spec. 1/spec.2) vs. establishment-occupation-year (spec.3).

Within-industry calibration: model fit & counterfactual

Main

- Counterfactual: $\chi \uparrow$ explains 83% of model-implied $\uparrow$ in between-share



Outsourcing & within-occupation ranking analysis

Main

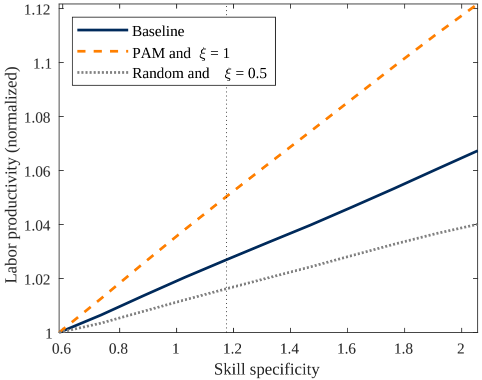
- **Concern:** confounding shifts in labor boundary of firm, e.g. outsourcing
- **Address this concern in multiple steps:**
  1. empirically rank workers *within* occupation ("good engineer vs. mediocre engineer")
  2. empirically re-estimate coworker sorting & complementarity (lower but similar $\uparrow$)
  3. re-estimate model for both periods & re-do counterfactual exercises
- **Result:** qualitatively & quantitatively similar findings

	Δ model	Implied % Δ model due to Δ parameter
Model 2: within-occ. ranking	0.198	-
Cf. a: fix period-1 χ	0.076	61.47

Realizing gains from specialization requires well-matched teams

▸ Eliminating mismatch

▸ Conclusion



- Gains from the division of labor are limited by the functioning of the labor market
  - microfoundation for recent econ-dev findings *[Bandiera-Kotia-Lindenlaub-Moser-Prat, 2024]*

Implications for aggregate productivity

Main

- Production complementarities imply sorting matters for agg productivity – search frictions induce misallocation
- **Quantify** mismatch costs: compare eqm outcome to productivity under pure talent-PAM and different values of ξ – given param's for 2010s

	Labor productivity
Baseline (norm.)	100
PAM + $\xi = 1$	102.6
PAM	101.1
$\xi = 1$	101.4

- Productivity gains from eliminating mismatch are of **limited magnitude**. But...