Superstar Teams

Lukas Freund

Columbia Business School

19 September 2024

Columbia Econ Lunch Seminar

Motivation: most people work in teams – but not in macro

- **Most people work in teams:** production $\neq$ sum of separable worker outputs
  [definition: Alchian-Demsetz, 1972; survey evidence: Lazear and Shaw, 2007; Weidmann & Deming, 2021]

Motivation: most people work in teams – but not in macro

- **Most people work in teams:** production $\neq$ sum of separable worker outputs
  [definition: Alchian-Demsetz, 1972; survey evidence: Lazear and Shaw, 2007; Weidmann & Deming, 2021]

- **Organizational econ**: the firm as a mechanism for people to work together
  [Smith, 1776; Alchian-Demsetz, 1972; Marschak-Radner, 1972; Becker-Murphy, 1992; ...]

Motivation: most people work in teams – but not in macro

- **Most people work in teams:** production $\neq$ sum of separable worker outputs
  [definition: Alchian-Demsetz, 1972; survey evidence: Lazear and Shaw, 2007; Weidmann & Deming, 2021]
- **Organizational econ**: the firm as a mechanism for people to work together
  [Smith, 1776; Alchian-Demsetz, 1972; Marschak-Radner, 1972; Becker-Murphy, 1992; ...]
- **Macro/labor/firm-dynamics models:** little room for team production
  [exceptions: Garicano, 2000; Jarosch-Oberfield-RossiHansberg, 2021; Herkenhoff et al., 2024]

Motivation: most people work in teams – but not in macro

- **Most people work in teams:** production $\neq$ sum of separable worker outputs
  [definition: Alchian-Demsetz, 1972; survey evidence: Lazear and Shaw, 2007; Weidmann & Deming, 2021]
- **Organizational econ**: the firm as a mechanism for people to work together
  [Smith, 1776; Alchian-Demsetz, 1972; Marschak-Radner, 1972; Becker-Murphy, 1992; ...]
- **Macro/labor/firm-dynamics models:** little room for team production
  [exceptions: Garicano, 2000; Jarosch-Oberfield-RossiHansberg, 2021; Herkenhoff et al., 2024]
- **Questions:** How can we incorporate these ideas into a macro model?
  How to empirically discipline such a theory with micro data?
  Does team production matter for macro outcomes, and how?

This paper: conceptual framework

- **Theory**
  - production requires many **tasks**
  - workers have **het. task-specific skills**
    - talent $\sim$ absolute advantage
    - skill specificity $\sim$ *dispersion* in ind. task-specific skills
  - each firm consists of an organized collection of workers (**"team"**)
  - hiring workers involves (random) **search**

This paper: conceptual framework

- **Theory**
  - production requires many **tasks**
  - workers have **het. task-specific skills**

    → talent ∼ absolute advantage

    → skill specificity ∼ *dispersion* in ind. task-specific skills
  - each firm consists of an organized collection of workers (**“team”**)
  - hiring workers involves (random) **search**

- **Core idea:** **skill specificity** endogenously implies (1) productivity gains from team production, and (2) **coworker talent complementarities**

This paper: conceptual framework

- **Theory**
  - production requires many **tasks**
  - workers have **het. task-specific skills**

    → talent $\sim$ absolute advantage

    → skill specificity $\sim$ *dispersion* in ind. task-specific skills
  - each firm consists of an organized collection of workers (**"team"**)
  - hiring workers involves (random) **search**

- **Core idea:** **skill specificity** endogenously implies (1) productivity gains from team production, and (2) **coworker talent complementarities**

- **Implications:**
  - incentives for assortative matching → **firm-level inequality**
  - frictional coworker mismatch is costly → **agg. productivity**

This paper: theory - measurement - applications

Literature

1. **Theory**
   - **microfound task-based production fn.** $\rightarrow$ endogenous coworker complementarities
   - organizational theory $\rightarrow$ low-dim. production fn. despite high-dim. skills

This paper: theory - measurement - applications

▸ Literature

1. **Theory**
   - **microfound task-based production fn.** $\rightarrow$ endogenous coworker complementarities
   - organizational theory $\rightarrow$ low-dim. production fn. despite high-dim. skills
2. **Theory meets data**
   - **identification** with micro panel data on wages+matches $\rightarrow$ estimate & validate model
   - coworker sorting generates large firm-level dispersion in productivity+pay

This paper: theory - measurement - applications

▸ Literature

1. **Theory**
   - **microfound task-based production fn.** $\rightarrow$ endogenous coworker complementarities
   - organizational theory $\rightarrow$ low-dim. production fn. despite high-dim. skills
2. **Theory meets data**
   - **identification** with micro panel data on wages+matches $\rightarrow$ estimate & validate model
   - coworker sorting generates large firm-level dispersion in productivity+pay
3. **Applications**
   - **structural explanation for "firming up inequality":** $\uparrow$ skill specificity *[e.g. Jones, 2009; Deming, 2017]* explains 25-40% of $\uparrow$ between-firm wage inequality share in DE since '85 *[e.g. Card et al., 2013; Bloom et al., 2019]*
   - labor market frictions limit productivity gains from $\uparrow$ skill specificity

Roadmap

Theory

Theory Meets Data

Applications

Environment: task-based production & frictional matching into teams

- **Agents:** continuums of workers & firms, infinitely-lived & risk-neutral
  - **firms** are ex-ante identical, have $n \in \mathbf{Z}_+$ employees
  - **worker** i is endowed with time-invariant, task-specific skills, $\{z_i(\tau)\}_{\tau\in[0,1]}$
- **Production:** continuum of imperfectly substitutable tasks *[e.g., Acemoglu-Restrepo, 2018]*
- **Labor market matching**: workers & multi-worker firms meet through random search *[similar to Herkenhoff et al. (2024) but with high-dim. skills]*
- **Challenges:** production possibility set? how to avoid curse of dimensionality?
- **Game plan:**
  1. microfound tractable *reduced-form* firm-level production fn $f(\cdot)$
  2. given $f(\cdot)$, analyze team formation

Parametrized multi-dim. skills: "Fréchet-ing things up"

Assumption: Fréchet dist.

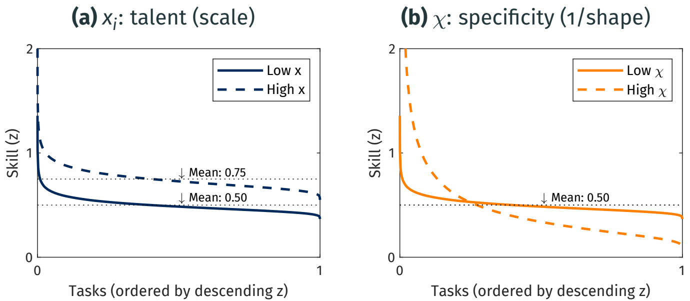
$$P\left[z_i(\tau) \leq z\right] = \exp\left(-\left(\frac{z}{\iota x_i}\right)^{-\frac{1}{\chi}}\right)$$

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

[Eaton & Kortum, 2002]

Interpretation

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



Parametrized multi-dim. skills: 2 arbitrary workers

Assumption: Multivariate Fréchet dist.

$$\Pr\left[z_1(\tau) \leq z_1, z_2(\tau) \leq z_2\right] = \exp\left[-\left(\sum_{i=1}^{n=2}\left(\left(\frac{z_i}{\iota x_i}\right)^{-\frac{1}{\chi}}\right)^{\frac{1}{\xi}}\right)^{\xi}\right]$$

with $x_i \in \mathbb{Z}_{++}$, $\chi \in [0, \infty)$, $\xi \in [0, 1]$ (Copula param).

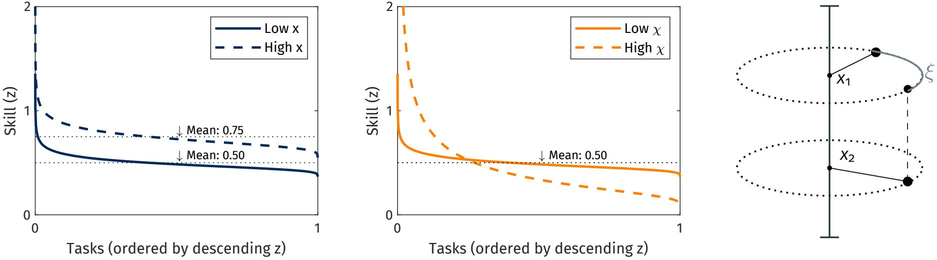
[Eaton & Kortum, 2002; Lind & Ramondo, 2023]

Interpretation

Economy-wide dist.

$$\Pr\left[z_1(\tau) \leq z_1, z_2(\tau) \leq z_2\right] = \exp\left[-\left(\sum_{i=1}^{n=2}\left(\left(\frac{z_i}{\iota x_i}\right)^{-\frac{1}{\chi}}\right)^{\frac{1}{\xi}}\right)^{\xi}\right]$$

(a) x_i: talent (scale) **(b)** χ: specificity (1/shape) **(c)** ξ: hor. distance



Production with a single team of given composition

- 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

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

Firm's optimization problem

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

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

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

Firm's optimization problem

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

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

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

- **FOCs** imply

$$\underbrace{\lambda(\tau)}_{\text{shadow cost of } \tau} = \min_i \left\{ \frac{\overbrace{\lambda_i^L}^{\text{opportunity cost of } i\text{'s time}}}{\underbrace{z_i(\tau)}_{i\text{'s skill for } \tau}} \right\}$$

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 **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

▸ Lemma

Proposition: Reduced-form production function

Under Assumption 1, talents $\mathbf{x}$ and horizontal distance ξ are sufficient statistics for team output Y given parameter χ :

$$Y = f(\mathbf{x}, \xi; \chi)$$

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

Micro-founded production function

Lemma

Proposition: Reduced-form production function

Under Assumption 1, talents **x** and horizontal distance ξ are sufficient statistics for team output Y given parameter χ :

$$Y = f(\mathbf{x}, \xi; \chi)$$

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

- **Benchmark** without division of labor: $Y = n \times \left(\frac{1}{n} \sum_{i=1}^{n} x_i\right)$

Gains from team production are increasing in skill specificity

Proposition: Reduced-form production function

$$f(\mathbf{x}, \xi; \chi) = \underbrace{n^{1+\chi\xi}}_{\text{efficiency gains}} \times \left(\frac{1}{n} \sum_{i=1}^{n} (x_i)^{\frac{1}{1+\chi\xi}} \right)^{1+\chi\xi}$$

1. Value of **team production** increasing in skill specificity (χ) ▸ Intuition
   - gains from team production realized when coworkers have differentiated expertise (ξ)

Skill specificity implies that productivity is lowered by talent dispersion

Proposition: Reduced-form production function

$$f(\mathbf{x}, \xi; \chi) = \underbrace{n^{1+\chi\xi}}_{\text{efficiency gains}} \times \underbrace{\left(\frac{1}{n} \sum_{i=1}^{n} (x_i)^{\frac{1}{1+\chi\xi}} \right)^{1+\chi\xi}}_{\text{talent complementarity}},$$

1. Value of team production increasing in skill specificity (χ) ▸ Intuition
2. Skill specificity (χ) implies **coworker talent complementarities** ▸ Intuition
   - $\frac{\partial\left(\partial^2 f(\cdot)/\partial x_1 \partial x_2\right)}{\partial \chi} > 0$

Roadmap & key takeaways

Theory

1. Under the optimal task assignment, **skill specificity** *endogenously* implies (1) gains from team production & (2) **coworker talent complementarities**
   - Fréchet + assignment theory $\rightarrow$ low-dim. production fn. despite high-dim. skills
2. Next: if production possibilities are summarized by $f(\mathbf{x}, \xi; \chi)$, what is the endogenous composition of different teams?

Endogenous team composition: frictional matching

- **Assumptions**:
  - $x \sim$ uniform; cond. on x, workers uniformly located on circle with unit circumference
  - 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

  - no OJS in baseline

- ξ is operationalized as a **match-specific shock**
  - task-specific skills perfectly observable to agents before match decision, econometrician only observes x; tractable b/c by Prop. 1, $(\mathbf{x}, \xi)$ is sufficient statistic

- **Stationary equilibrium**

▸ HJBs ▸ KFEs ▸ Definition

Surplus max. determines which teams are formed

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

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

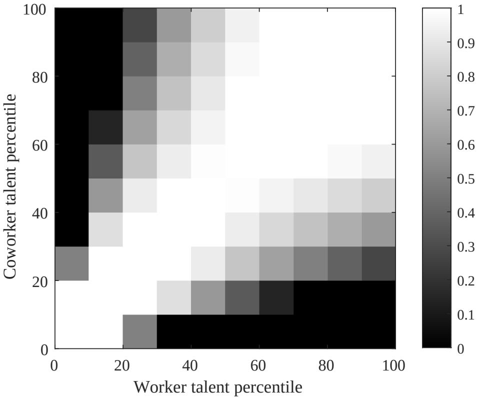
$V_u(x)$: value for unemp. worker; $V_{f.0}$: value for vacant firm; $d_u(x)$: density of unemployed workers; $u = \int d_u(x)dx$; ω: worker bargaining wgt; $\delta(x)$: sep. hazard; $\lambda_{v.u}$: hazard rate of vacancy meeting unmatched worker; H: cdf of ξ

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

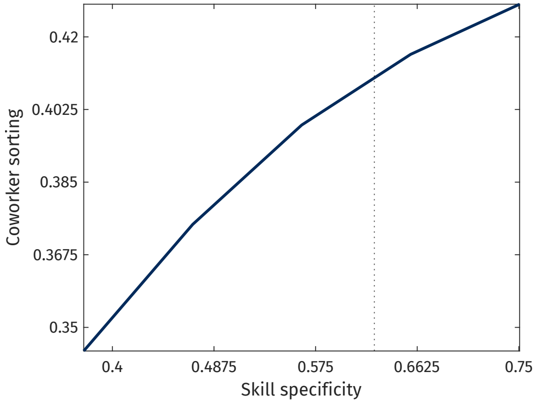
$$S(x|x',\xi)(\rho + \delta(x) + \delta(x')) = f(x,x',\xi) - \textit{outside options}$$

Equilibrium properties: conditional matching probabilities for given χ

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



Comparative 'statics': more positive assortative matching as $\chi \uparrow$



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

- Numerical solution of model with discrete talent types $\hat{x}_i \in \{1, \ldots, 10\}$
- **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
  - "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 $\hat{x} \to x$
- **Focus today:** structural identification of χ in theory & practice

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 can recover from $w(x|x')$ given x and x'
  - intuition: outside options influence *level* of w *[Eeckhout-Kircher, 2011]* but enter separably

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

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

$$\Rightarrow \frac{\partial^2 \bar{f}(x, x')}{\partial x \partial x'} \overset{\text{numerically}}{\underset{\sim}{\propto}} \boxed{\frac{\partial^2 \bar{w}(x|x')}{\partial x \partial x'}}$$

can measure this

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:** $\hat{\beta}_c$ = 0.0058

  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:** $\hat{\beta}_c$ = 0.0058

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 $\hat{\beta}_c$

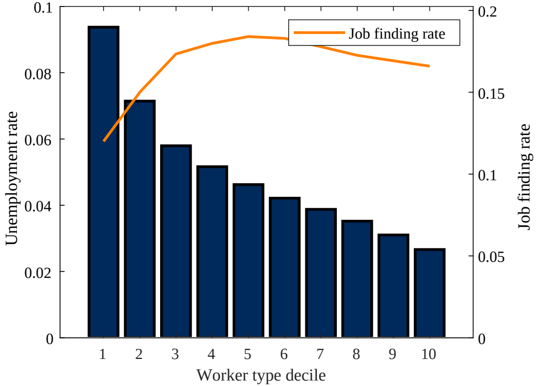
Parameter values

  - $\frac{f(x^{p80},x^{p80},1)+f(x^{p20},x^{p20},1)}{f(x^{p80},x^{p20},1)+f(x^{p80},x^{p20},1)} = 1.16$

Quantitative properties of estimated model (untargeted)

▸ Parameter values

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

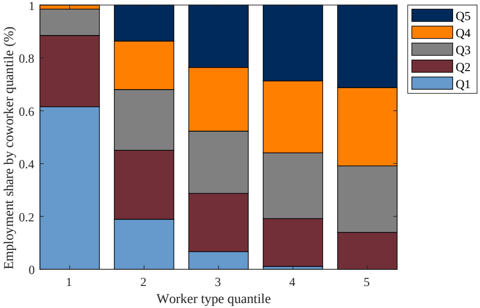


Quantitative properties of estimated model (untargeted)

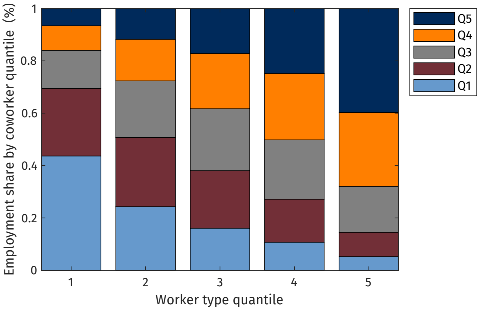
1. ✓Higher-x workers experience lower unemployment rates due to lower separation rates but job finding rates don't increase much with talent *[e.g., Cairo & Cajner, 2018]*

2. ✓**Match coworker sorting patterns**

(a) Theory



(b) Data



Quantitative properties of estimated model (untargeted)

1. ✓Higher-x workers experience lower unemployment 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**
   - $\rho_{xx} = 0.52$ (vs. 0.62 in data)

   ▸ Avg. coworker figure

3. ✓**Match 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**

Validation of core model mechanisms

- **Industry-level analysis** ▸ Details
- **Cross-sectional variation across occupations** ▸ Details

  ✓task-based proxy for $\chi \uparrow \rightarrow$ r.-f. talent complementarity $\uparrow$

  ✓r.-f. talent complementarity $\uparrow \rightarrow$ coworker talent sorting $\uparrow$

- **Direction of EE moves:** non-wage implications of complementarities ▸ Details

  ✓Δ coworker talent positively correlated with own talent

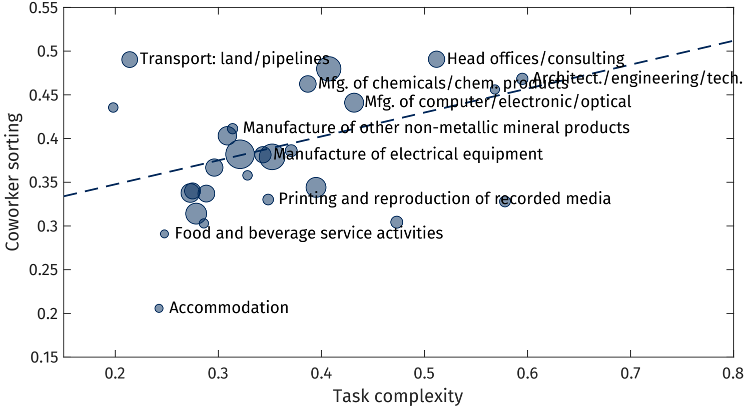
- **Heterogeneous effects of coworker deaths** *[cf. Jaeger-Heining, 2022]* ▸ Details

  ✓wage losses from coworker death *if* coworker specialized in different tasks ($\xi \uparrow$)

- **Team-production functions in science** *[cf. Ahmadpoor-Jones, 2019]* ▸ Details

  ✓talent complementarities stronger *precisely* when teamwork more valuable

Industries: more coworker sorting in industries with high task complexity

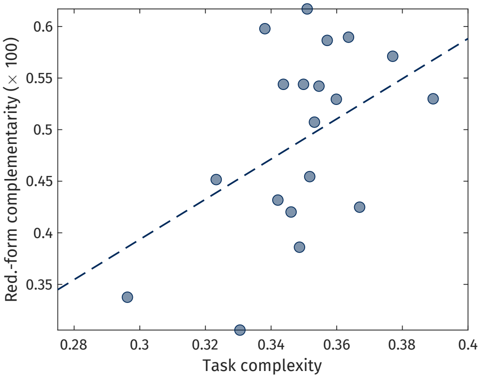


Notes. Task complexity: occupation-specific measure of the share of cognitive non-routine tasks, weighted by industry-specific occ. employment weights. Weighted linear best fit. Data: SIEED + BIBB.

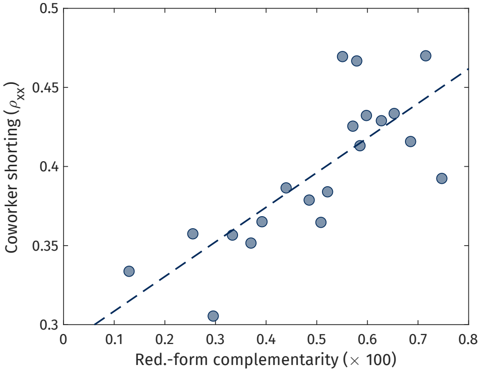
Industries: task complexity, complementarities

▸ 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

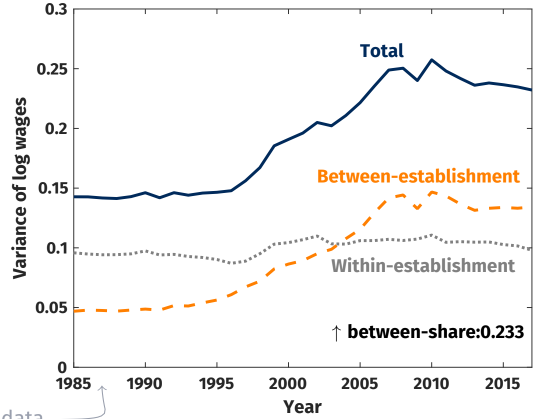
4. **Structural explanation for the “firming up of inequality”**
5. Implications for aggregate productivity

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]

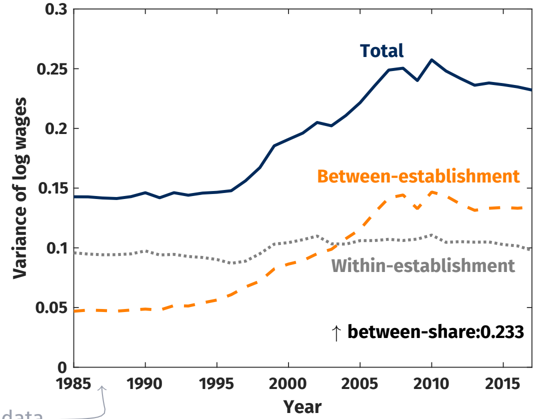


German matched employer-employee data

Applied question

▸ Details

Applied question: what are the causal driver(s)?



German matched employer-employee data

Hypothesis (1): 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

Hypothesis (2): preview of argument

1. The set of tasks any one worker can perform well has narrowed: **skill specificity** ↑
2. **Coworker 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 $\uparrow$

▸ Schooling ▸ Peer effect trends

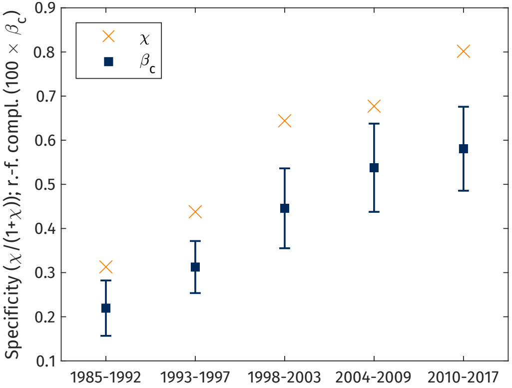
- **Method:** estimate reduced-form coefficient β_c for 5 sample periods

  $\Rightarrow$ re-estimate structural model

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

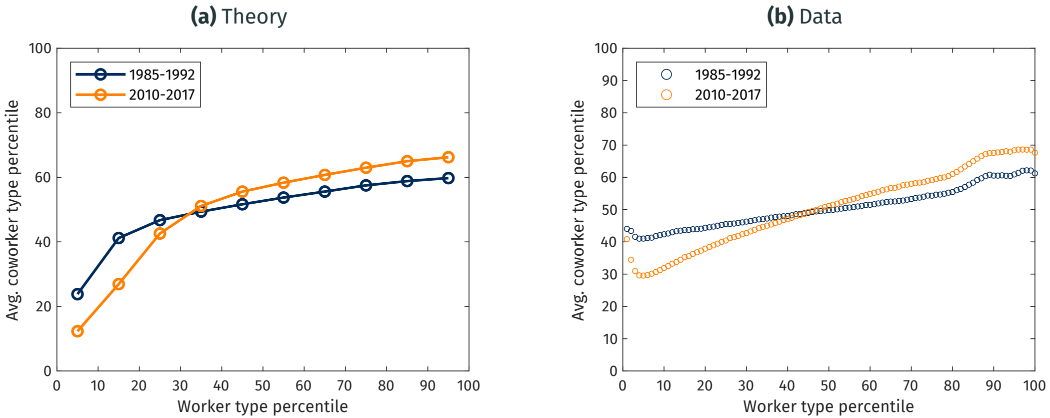
  [consistent with Grigsby's (2024) US estimates]

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



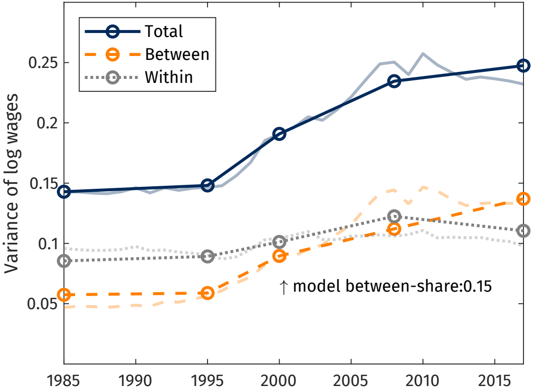
Talent sorting has intensified: theory & data

Details Model Meets Data



Model predicts increased firm-level wage inequality

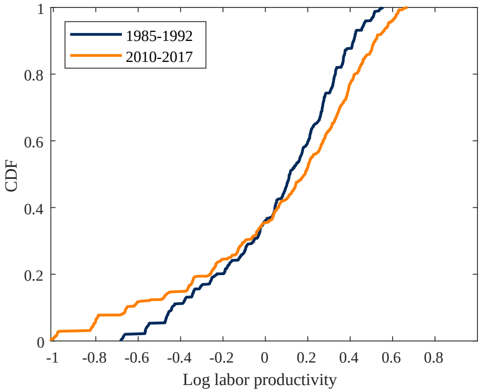
- **Model replicates untargeted rise of between-share in data**
  - ∼ >2/3 of empirical ↑ between-share ('85-'92)→('10-'17)



Productivity dispersion

Main

- 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 $\approx$ 25-40% of observed between-share $\uparrow$

- **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 58%** of model-predicted $\Delta \leftrightarrow \approx$ **38% of empirical** Δ
  - $\uparrow$ team-advantage ($n^{1+\chi\xi}$) partly counteracts effect of $\uparrow$ complementarities (CES term)
- **Robustness** exercises: 25-40%

▸ Within-industry ▸ Outsourcing

- Effect of $\downarrow$ search frictions *[e.g., Martellini-Menzio, 2021]* $\sim$ 16% 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

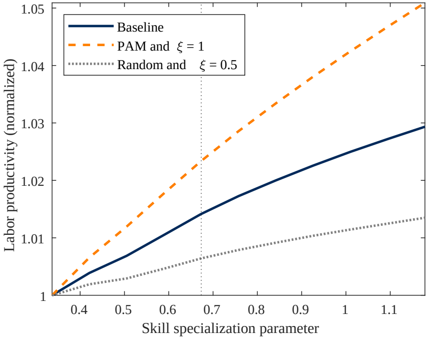
3. Estimated model endogenously generates realistic ex-post firm heterogeneity

Applications

4. **Increased skill specificity – leading to stronger complementarities and, hence, sorting – explains a substantial share of ↑ between-firm wage inequality share**
5. **Next:** brief glance at productivity application

Realizing gains from specialization requires well-matched teams

▸ Eliminating mismatch



- Gains from the division of labor are limited by the functioning of the labor market
  - microfoundation for finding that improvements in sorting constitute important amplification channel for econ development *[Bandiera-Kotia-Lindenlaub-Moser-Prat, 2024]*
  - labor market frictions may inhibit specialization *[cf. Atencio et al., 2024; Bassi et al., 2024]*

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 endogenously generates realistic ex-post firm heterogeneity

Applications

4. Increased skill specificity – leading to stronger complementarities and, hence, sorting – explains a substantial share of $\uparrow$ between-firm wage inequality share
5. **Enhanced sorting crucial to realize productivity gains from $\uparrow$ skill specialization**

Conclusion: firms form & organize teams – matters for macro

Literature

- **Classic idea absent from macro:** the firm as a "team assembly" technology
- **This paper:** develops a tractable model of team production & formation, takes it to the data, and quantitatively analyzes it
- **Takeaways:**
  1. **theory:** skill specificity $\rightarrow$ complementarities $\rightarrow$ talent sorting
  2. **measurement**: endogenous firm-level dispersion in productivity & pay
  3. **quantitative application**: $\uparrow$ skill specificity helps explain **"firming up" of inequality**

Thank You!

Relation & contributions to literature

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]*

Economy-wide distribution of skills

▸ Main

- Individual worker: Fréchet distributed task-specific skills
- Any two workers: Multivariate Fréchet distributed task-specific skills
- **Talent distribution:** $\hat{x} \sim U(0,1)$
  - interpret $\hat{x}$ as rank in talent distribution
  - model meets data: map *ordinal* types to *cardinal* talent x by parameterizing the quantile fn.
- Conditional on talent, workers are uniformly distributed on a circle with unit circumference $\rightarrow$ 'offer' distribution $\xi \sim U(0,1)$

Lemma

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 $\chi < \infty$

Low χ

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

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

HJB: surpluses

- Surplus of coalition of firm with worker x

$$(\rho+\delta)S(x) = f_1(x) - \rho\big(V_u(x) + V_{f.0}\big) + \lambda_{v.u}(1-\omega)\int \frac{d_u(\tilde{x}')}{u} S(\tilde{x}'|x,\tilde{\xi})^{+} dH(\tilde{\xi})\tilde{x}'. \quad (11)$$

- Surplus from adding x to x' with xi

$$S(x|x',\xi)\big(\rho+2\delta\big) = f_2(x,x',\xi) - \rho\big(V_u(x) + V_u(x') + V_{f.0}\big) + \delta S(x) - (\rho+\delta)S(x'). \quad (12)$$

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

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

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.

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
  - pragmatic approach: 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]
  - theory-consistent: non-param. ranking algo [Hagedorn et al., 2017] $\rightarrow$ similar ranking

$\Rightarrow$ **Worker i's talent type $\hat{x}_i$: decile rank of i's FE *within 2d-occupation***

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

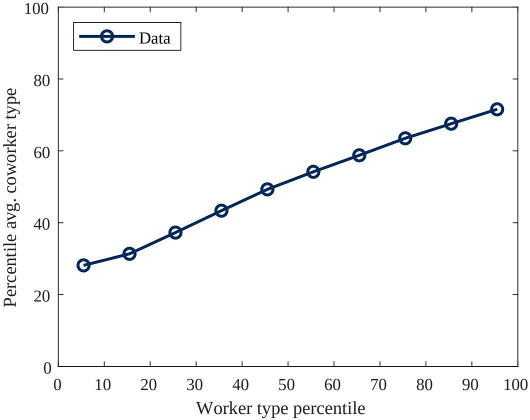
Discussion

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 are *differentiated* in their task-specific productivities
- **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 $B^\xi(x|x') = \{\xi : S(x|x',\xi) > 0\}$

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

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

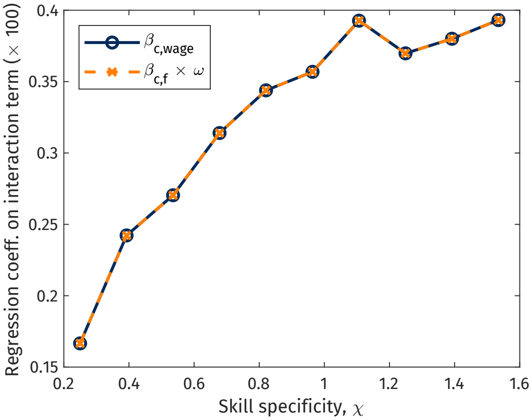
(A few fresh) thoughts on relation to Borovičková-Shimer (2024) argument

▸ Main

- Reasoning in B-S also applies to *coworker* matching: realized matches and hence wages may reflect *selection* on match-specific productivity shocks
  - model version presented today ($\neq$ JMP) explicitly microfounds & accounts for selection
- A few (fresh) thoughts
  1. Theoretical differences
     - microfoundation delivers structural interpretation of match-specific shocks ξ under which they (/their impact on f) are inherently bounded
     - χ controls both the degree of f complementarity *and* impact of ξ on output
     - wage vs *log* wage (average G wage is s. increasing and s. submodular (s. supermodular) for any s. increasing and s. *concave* (convex) G)
  2. MC study: ✓
  3. X-sectional evidence: β_c and ρ_{xx} co-move
- Alternative strategy: infer χ directly from observed equilibrium sorting

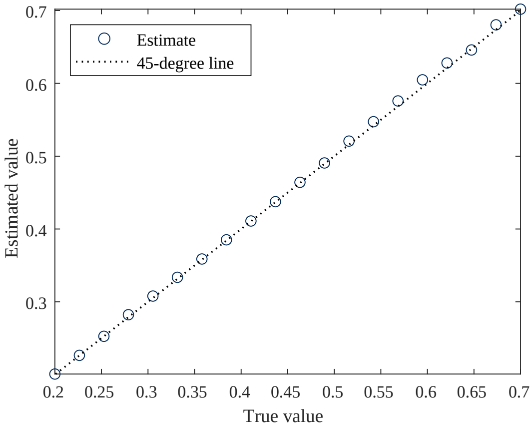
Regression coefficients co-move with χ

▸ Main



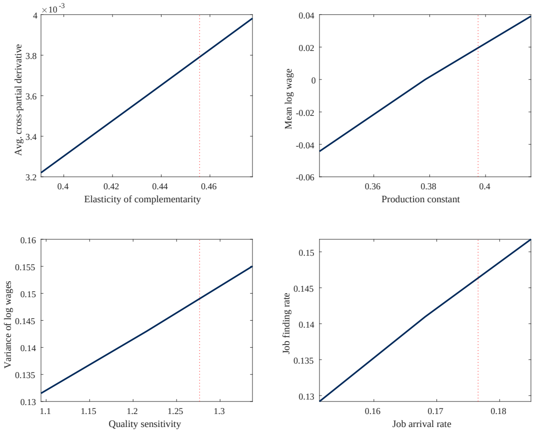
Monte Carlo study: identifying χ

Main



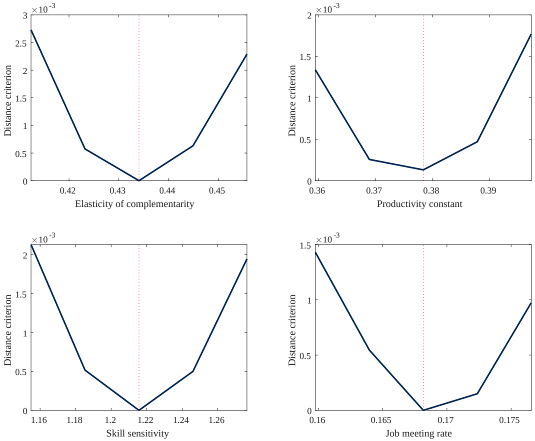
Identification validation exercise 1

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)
$\hat{\beta}_c$	0.0095***	0.0094***	0.0091***	0.0058***
	(0.00051)	(0.00039)	(0.00035)	(0.00278)
Own type controls	Yes	Yes	Yes	Yes
Coworker type controls	Yes	Yes	Yes	Yes
Employer FEs	No	Yes	Yes	Yes
Industry-year FEs	No	No	Yes	Yes
Occupation-year FEs	No	No	Yes	Yes
Type ranking	Economy	Economy	Economy	Occupation
Obs. (1000s)	4,410	4,410	4,410	4,410
Adj. R^2	0.736	0.803	0.811	0.800

Notes. Employer-clustered standard errors are given in parentheses. * 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

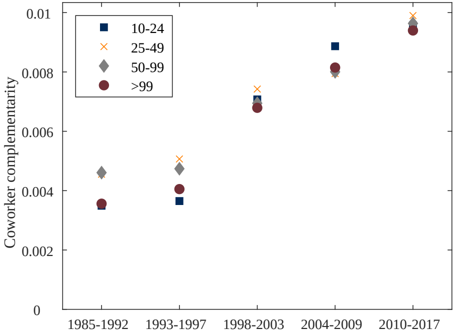
	Baseline				Within-industry avg.			
Sample Period	σ_w^2	$\sigma_{\bar{w}}^2/\sigma_w^2$	ρ_{xx}	$\hat{\beta}_c$	σ_w^2	$\sigma_{\bar{w}}^2/\sigma_w^2$	ρ_{xx}	$\hat{\beta}_c$
1	0.143	0.337	0.427	0.0036	0.125	0.249	0.333	0.00283
2	0.148	0.391	0.458	0.0042	0.125	0.288	0.351	0.00342
3	0.191	0.456	0.495	0.0070	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.

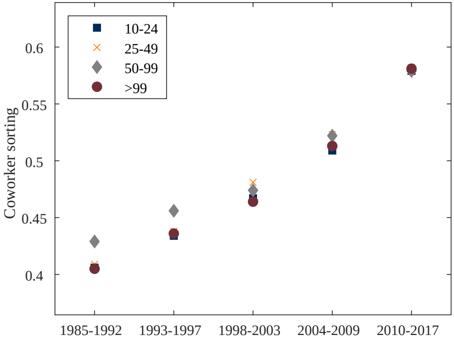
Coworker complementarity & sorting by team size

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.

Coworker complementarity: finite-differences approximation

▸ Robustness overview

- Regression approach imposes strong functional form assumptions on approximated empirical wage function $\hat{w}(x|x')$
  - ofc, mirrored inside structural model when calibrating
- Alternative: construct non-parametric $\hat{w}(x|x')$, then use finite-difference methods to compute the cross-partial derivative (but w/o FE controls)

Period	Regression	Non-parametric FD method
1985-1992	0.0036	0.0073
1993-1997	0.0042	0.0074
1998-2003	0.0070	0.0081
2004-2009	0.0082	0.0120
2010-2017	0.0091	0.0098

Estimation results (2010-2017)

Main

- Estimated value of χ implies $\frac{f(x^{p80},x^{p80},1)+f(x^{p20},x^{p20},1)}{f(x^{p80},x^{p20},1)+f(x^{p80},x^{p20},1)} = 1.16$

Parameter	Description	Target	Value	m	$\hat{m}$
χ	Skill specificity	$\hat{\beta}_c$	2.37	0.0058	0.0058
a_0	Talent mapping, constant	Avg. wage (norm.)	0.24	1	1
a_1	Talent mapping, scale	Var. log wage	1.46	0.241	0.241
b_1	Replacement rate, scale	Replacement rate	0.60	0.63	0.63
λ_u	Meeting hazard	Job finding rate	0.22	0.162	0.162
δ_0	Separation hazard, constant	Job loss rates	0.013		
δ_1	Separation hazard, scale	Job loss rates	-0.68		
ω	Worker bargaining weight	External	0.50		
$\bar{n}$	Effective team size	Avg in SIEED	20		

Model Meets Data: types and production function

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

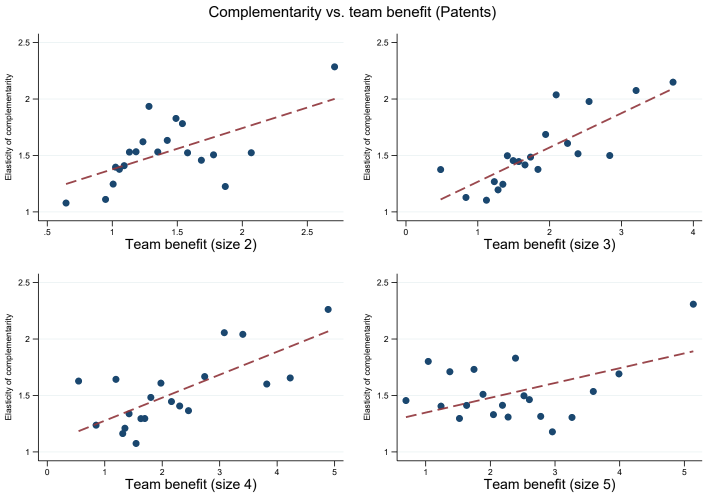
1. Estimated 'talent types' are in *ordinal* space, $\tilde{x} \in [0,1]$. Mapping $x_i = a_0 + a_1\tilde{x}_i$
   - next iteration: allow for higher-order terms
   - (a_0, a_1) captures (i) "talent-biased technological change," and (ii) Δ talent distribution
     - nb: Hakanson et al (2021) find no 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)

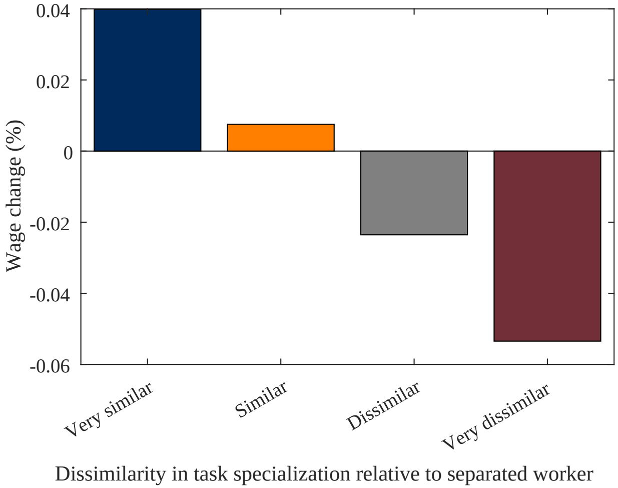
Main



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)

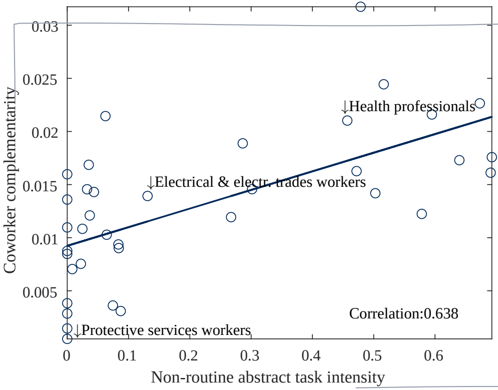
Main



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

Main

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



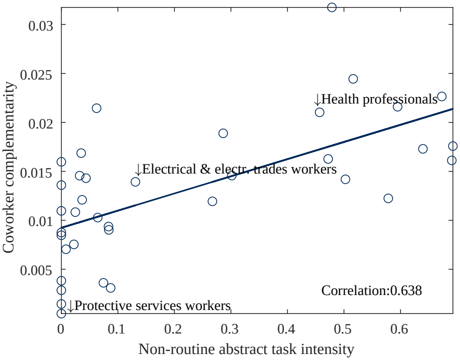
$\hat{\beta}_c$ estimated separately for each occupation

Non-routine abstract task intensity: proxy for χ

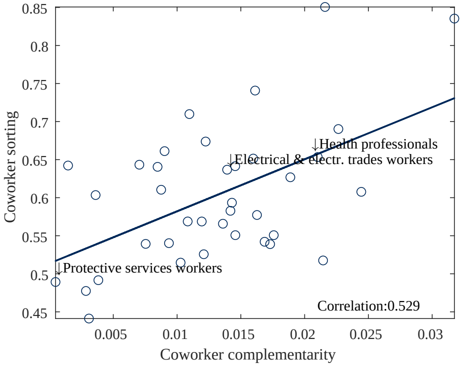
Notes. Quadros de Pessoal microdata. Analysis at ISCO-08-2d level.

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

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



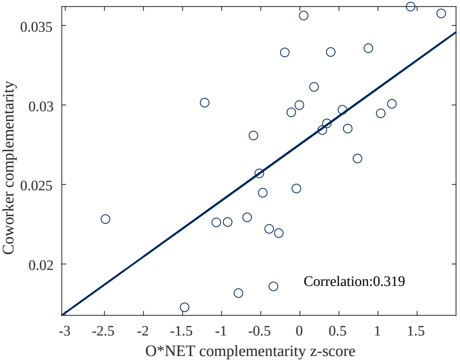
- ↑ **Coworker talent complementarity ⇒ ↑ coworker sorting**



Industries: coworker importance ⇒ complementarity ⇒ sorting

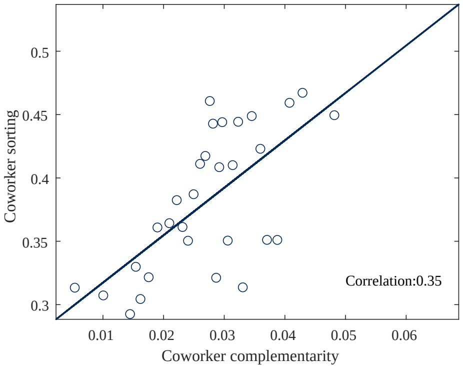
Main

- ↑ **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

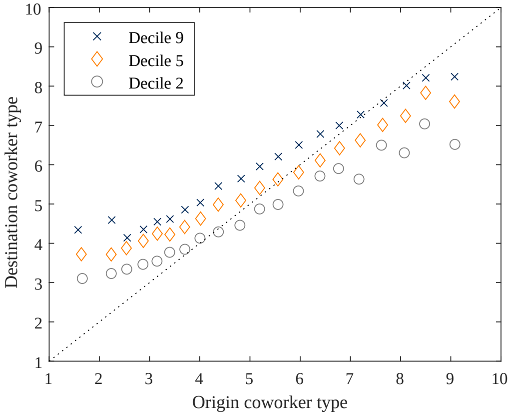
- **Theoretical prediction:** EE transitions move workers in surplus-maximizing direction $\Rightarrow \Delta\hat{x}_{-it} = \hat{x}_{-i,t} - \hat{x}_{-i,t-1}$ should be *positively* correlated with $\hat{x}_i$
  - $h_{2.1}(x, x''|x') = 1$ – worker x in a two-worker firm with coworker x'' would move 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		

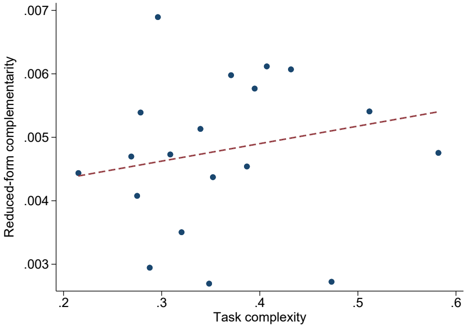
Table 1: Change in coworker type due to EE moves positively related to own type – increasingly so

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.

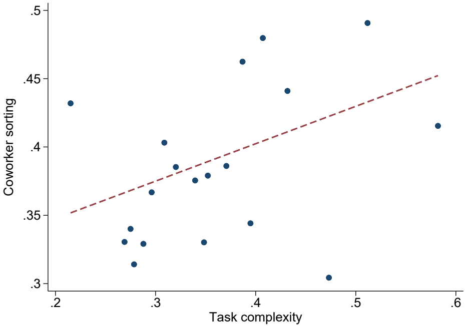
Industry-level analysis: mechanisms, w/o industry FEs

▸ Main

(a) Skill specificity → Complementarities



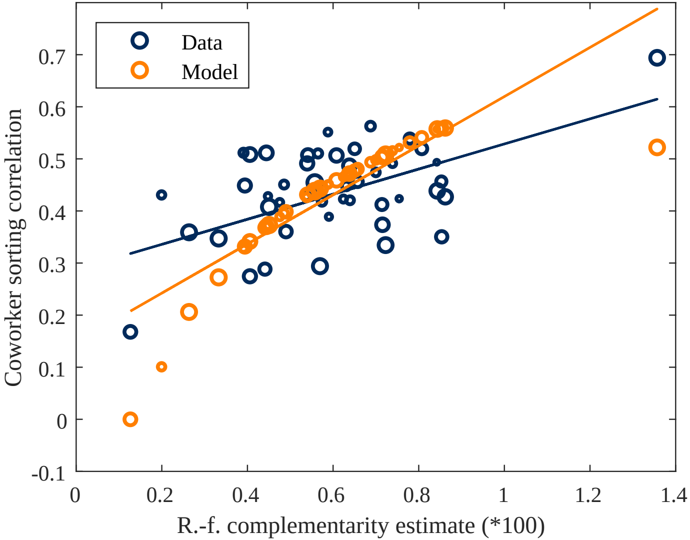
(b) Complementarities → sorting



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

Industry-level analysis: model vs. data

Main



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

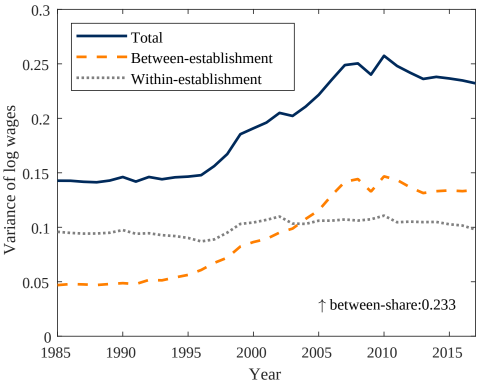
▸ Application ▸ Intro

- 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.

Fact #2: talented workers increasingly collaborate

▸ Intro ▸ Main ▸ Var. decomp. ▸ Fact #3

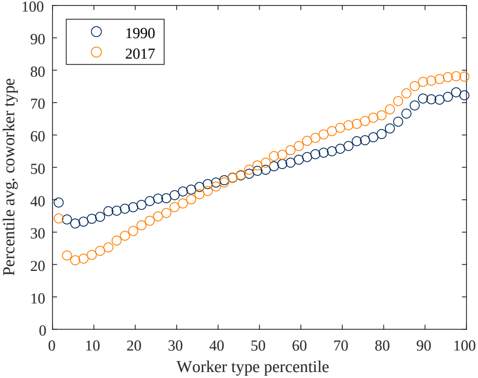
- To what extent do talented workers tend to have talented coworkers?

- **Fact 2: + assortative coworker sorting ↑**
  - ρ_{xx} = corr($\hat{x}_i$, $\hat{x}_{-it}$): 0.43 ('85-'92) ↗ 0.62 ('10-'17)

- Robust pattern

▸ Table ▸ Within-occ. nonlinear

▸ Hakanson et al. (2021)



Why might χ have increased over time: schooling argument

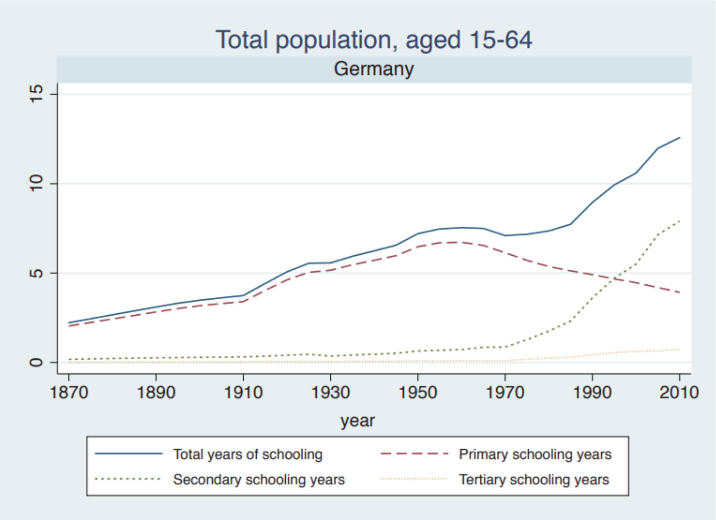
Main

- **Data:** trend toward more 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 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 $B_n = \exp\left(-\frac{b_n}{\alpha\theta_{n-1}}\right)$ 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 independent, identically distributed positive α-stable r.v.'s. Assume Z and $\{B_n\}$ are independent. Define the random variables $\{Z^{(n)}\}_{n\geq 1}$ 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 random variable with scale parameter x and shape parameter $\theta_n = \theta\alpha^n$.

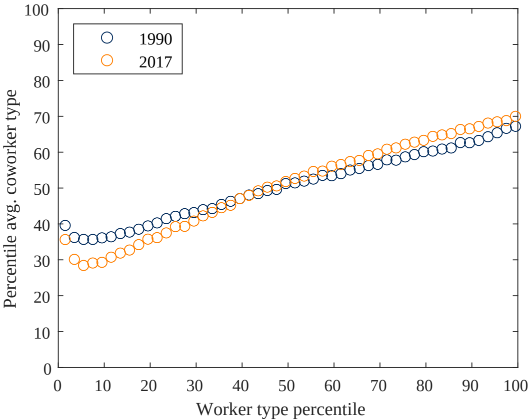
Barro Lee data for Germany



Evolution of coworker sorting: within-occupation ranking

Main

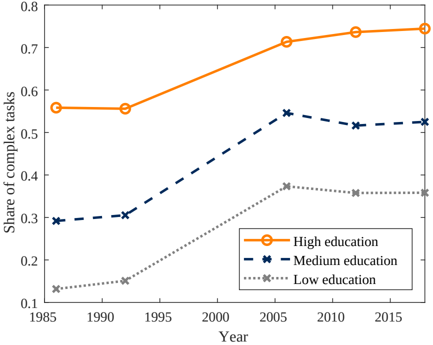
- The most talented within each occupation – the best engineer, PA, economist, manager, ... – tend to work together, and increasingly so



Task composition changes

Main

- **Task complexity $\uparrow$**: "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

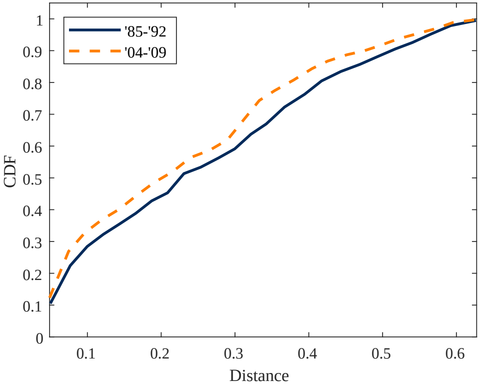
Back

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

Comparison

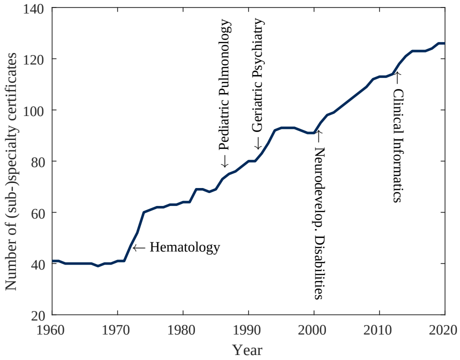
- **Q:** are workers becoming *more* likely to perform similar tasks across jobs?

- **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



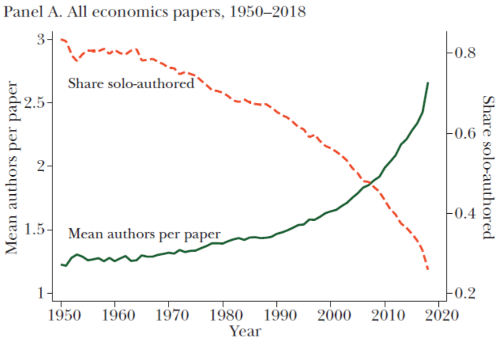
Examples: rising specialization

- **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]*

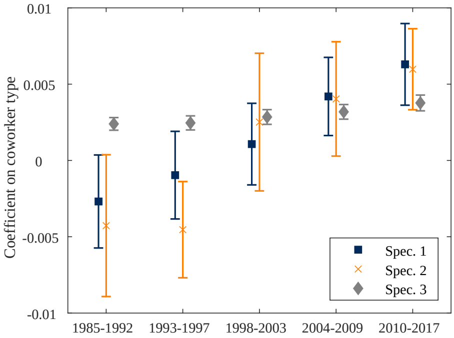


Coworker effects: log wage regression

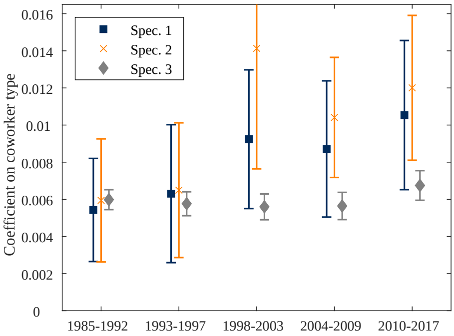
Back: cross-section Back: time series

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

(a) AKM types



(b) NP types

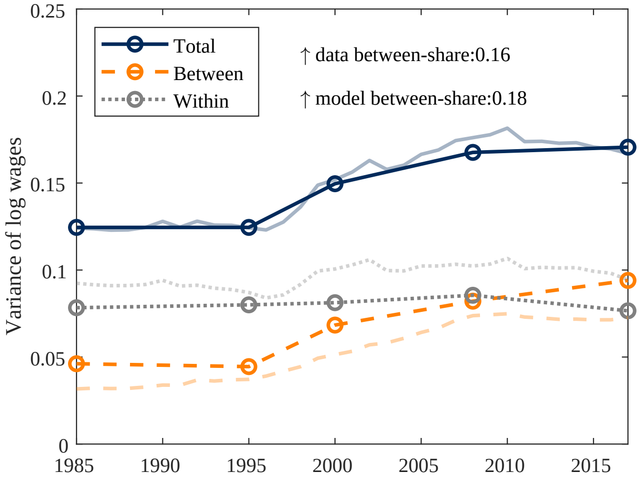


Notes. Specifications vary by ranking method – within-economy (spec. 1) vs. within-occupation (spec. 2/spec.3) and coworker group definition – establishment-year (spec. 1/spec.2) vs. establishment-occupation-year (spec.3).

Within-industry calibration: model fit & counterfactual

Main

- Counterfactual: $\chi\uparrow$ explains 83% of model-implied $\uparrow$ in between-share
- In *levels*, model over-states the between-share throughout



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 comp.	0.076	61.47

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...