Superstar Teams:
The Micro Origins and Macro Implications of Coworker Complementarities

Lukas B. Freund

University of Cambridge

22 December 2022

17^{th} Swiss Economists Abroad (SEA) Conference, Basel

Motivation

▸ Literature

- **Modern production processes**: teamwork between employees – some better, some worse – who are specialized in different job tasks
  - transformation of work: task complexity ↑ & teamwork ↑ *[Autor et al., 2003; Deming, 2017; Acemoglu-Restrepo, 2018; Atalay et al., 2021]*
- **Most macro models**: limited room for *within-firm* worker heterogeneity & coworker interdependence
  - hard to explain e.g. rise in between-firm inequality *[Card et al., 2013; Bloom et al., 2019; Criscuolo et al., 2021]*
- **This paper**: develops, tests & quantifies a parsimonious model of firms producing with teams of heterogeneous workers

Theoretical framework

- **Key concept:** coworker quality complementarities
  - $\frac{\partial^2 f(x,x')}{\partial x \partial x'} > 0$ where x and x' are the quality types of two workers

- **Steps:**

  1. **how to produce:** with heterogeneous workers

     $\Rightarrow$ derive production fn., $f(\cdot)$, looking at 1 team of *given* composition

  2. **with whom to produce:** hiring multiple workers via random search

     $\Rightarrow$ eqm. matching into teams with continuum of workers & employers *given* $f(\cdot)$

Theoretical framework: micro-founded production function

- Firm optimally assigns a continuum of tasks to N workers who differ in their *quality type*, x, and, whose efficiency varies across different tasks
  - $\chi \in [0,1]$: parameter capturing degree of worker-task specialization

Theoretical framework: micro-founded production function

- Firm optimally assigns a continuum of tasks to N workers who differ in their *quality type*, x, and, whose efficiency varies across different tasks
  - $\chi \in [0, 1]$: parameter capturing degree of worker-task specialization

Proposition (Aggregation result)

Under some existence conditions, the output of a team with workers $\{x_i\}_{i=1}^{N}$ can be written as

$$f(x_1, x_2, ..., x_N) = \underbrace{N^{1+\chi}}_{\textit{efficiency gain}} \underbrace{\left(\frac{1}{N} \sum_{i \in S} a_1 x_i^{\frac{1}{1+\chi}} \right)^{1+\chi}}_{\textit{CES: complementarities}}.$$

Theoretical framework: production function $\rightarrow$ hiring

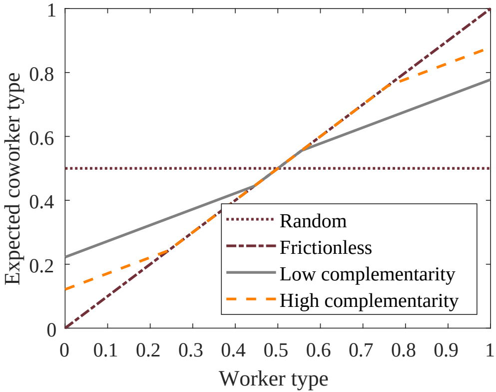
- Coworker complementarities are $\uparrow$ in the degree to which production involves task-specialization by team members

- **Question**: how do coworker complementarities shape the matching between workers and firms & the distribution of productivity & wages?
  - complementarities: coworker quality mismatch $\searrow$ output (think: closer to Leontief; Kremer (1993) O-ring)

- **Theoretical approach**: embed production block into quantitative model of hiring by multi-worker firms s.t. search frictions *[Herkenhoff et al., 2022]*
  - tradeoff: costly search vs. productivity benefit from good match
  - today: graphical illustration

Theoretical framework: production function → hiring

⇒ Coworker complementarities incentivize employers to form teams composed of workers of similar quality ⇒ ↑ coworker sorting & between-firm inequality



Notes. This figure plots the *average* type of coworkers (vertical axis) for a given worker type, for 4 different values of coworker complementarities. Closed-form results from a stylized version of the model.

Recap of main idea & roadmap

- **Model environment**
  - **how to produce:** firms assign tasks to workers who are heterogeneous in *quality* (vertical) & in which tasks they are most efficient (horizontal)
  - **with whom to produce:** multiple workers hired via random search
- **Key mechanism:** $\uparrow$ task specialization $\Rightarrow \uparrow$ efficiency gains from teamwork *but also* $\uparrow$ coworker quality complementarities $\Rightarrow \uparrow$ incentive to hire positively assorted teams $\Rightarrow \uparrow$ coworker sorting: high/high - low/low $\Rightarrow \uparrow$ between-firm inequality
- **Next:**
  1. Empirical evidence
  2. Quantitative analysis using estimated model

Empirical methods

- **Data**: German matched employer-employee panel data (LIAB)
  - paper: supplemented with universe of Portuguese worker & firm data, including information on uncensored wages & value-added per employee

- **Challenge**: how to measure coworker complementarities?
  1. global ranking of worker quality types ($\rightarrow x$'s): AKM-based or non-parametric Kemeny-Young type ranking algorithm
  2. exploit theoretically validated *proportionality*: production complementarities $\leftrightarrow$ wage complementarities

Today: evolution of complementarities over time

- paper: validation of theory in cross-section

Empirical result: complementarities over time

- **Regression** of i's wage level on covariates, incl. interaction term:

$$w_{it} = \beta_0 + \beta_1 \hat{x}_i + \beta_2 \hat{x}_{-it} + \beta_3 \left(\hat{x}_i \times \hat{x}_{-it} \right) + \psi_{j(it)} + \nu_{o(i)t} + \xi_{s(i)t} + \epsilon_{it},$$

  - $\hat{x}_{-it} = \frac{1}{|S-it|} \sum_{k \in S_{-it}} \hat{x}_k$: average type of i's coworkers in year t
  - $S_{-it} = \{k : j(kt) = j(it), k \neq i\}$: set of i's coworkers in t
  - $j(it)$: identifier of i's employer in period t
  - FEs: employer $\psi_{j(it)}$; occupation-year $\nu_{o(i)t}$; industry-year $\xi_{s(i)t}$

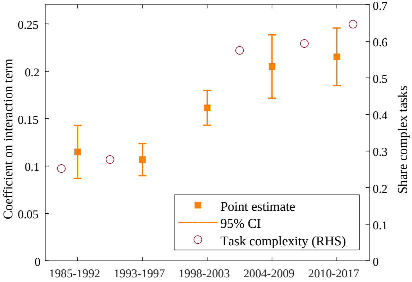
Empirical result: complementarities ↑ over time

▸ Non-parametric

- **Regression** of i's wage level on covariates, incl. interaction term:

$$w_{it} = \beta_0 + \beta_1 \hat{x}_i + \beta_2 \hat{x}_{-it} + \beta_3 \left(\hat{x}_i \times \hat{x}_{-it}\right) + \psi_{j(it)} + \nu_{o(i)t} + \xi_{s(i)t} + \epsilon_{it},$$

  - $\hat{x}_{-it} = \frac{1}{|S-it|} \sum_{k \in S_{-it}} \hat{x}_k$: average type of i's coworkers in year t
  - $S_{-it} = \{k : j(kt) = j(it), k \neq i\}$: set of i's coworkers in t
  - $j(it)$: identifier of i's employer in period t
  - FEs: employer $\psi_{j(it)}$; occupation-year $\nu_{o(i)t}$; industry-year $\xi_{s(i)t}$



Wage inequality in estimated model vs. data

▸ Estimation section

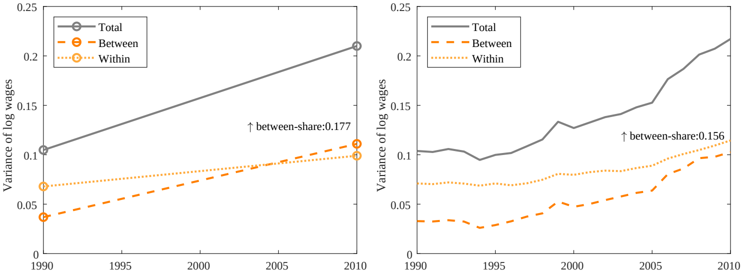
- **Next:** theory + micro data ($\rightarrow$ complementarity) $\Rightarrow$ quantitative analysis
  - focus **today:** explaining the "firming up inequality" *[Card et al., 2013; Song et al., 2019; Criscuolo et al., 2021; ...]*

- **Structural estimation** of model – for 1990 & 2010 – targeting micro-moments on wage complementarities but *not* between/within shares of wage inequality or matching patterns

Wage inequality in estimated model vs. data

▸ Bias correction ▸ Wage inequality robustness

- **Structural estimation** of model – for 1990 & 2010 – targeting micro-moments on wage complementarities but *not* between/within shares of wage inequality or matching patterns
- **Wage variance decomposition:** model vs. data



(a) Model **(b)** Data

Finding: complementarities explain $\approx \frac{1}{4} - \frac{1}{2}$ of $\uparrow$ between-share

- **Finding:** $\uparrow$ coworker complementarity explains $\approx$ half of $\uparrow$ between-establishment share of wage inequality
    - robustness: $\approx$ 1/4 to 1/2
    - skill-biased technological change underpins $\uparrow$ *overall* wage inequality & mechanically $\uparrow \nearrow$ between-share given positive sorting induced by complementarities

	Δ data	Δ model	% Δ model due to $\Delta\gamma$
Baseline	0.156	0.177	-
Scenario 1: only $\uparrow \gamma$	0.156	0.0935	52.9
Scenario 2: fix γ (1990)	0.156	0.0942	46.7

Summary

- **This paper:**
  1. proposes parsimonious **theory of teams** with heterogeneous workers
  2. **empirically validates** key model mechanisms using micro data & estimate evolution of coworker complementarities
  3. **quantifies macroeconomic implications** of coworker complementarities for aggregate productivity & evolution of wage inequality

3 takeaways:

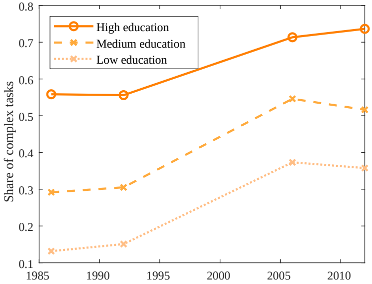
1. division of labor + worker-task specialization $\Rightarrow$ coworker quality complementarity
2. coworker complementarities + search frictions $\Rightarrow \downarrow$ agg. productivity & $\uparrow$ gaps between firms in workforce quality and pay
3. changing nature of work (esp. $\uparrow$ task complexity) $\Rightarrow \uparrow$ coworker complementarities $\Rightarrow$ "firming up" of wage inequality

Thank You!

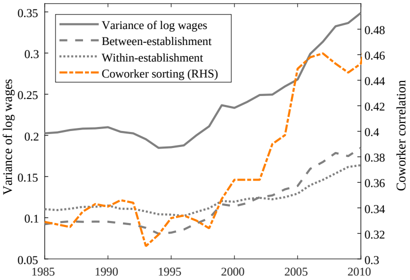
Motivation

Micro and macro trends in Germany

▸ Main



(a) Task complexity



(b) Task complexity

Contributions to literature

▶ Main

- **Firms as organisations:** division of labor within firms

  *[Lucas, 1978; Rosen, 1982; Becker & Murphy, 1992; **Kremer, 1993**; Kremer & Maskin, 1996; **Garicano, 2000**; Jarosch, Oberfield and Rossi-Hansberg, 2021; Bloesch et al., 2022]*

  ⇒ Coworker complementaries: productivity & inequality implications

- **Macro-models of labor market sorting:** allocation of workers

  *[Shimer & Smith, 2000; Eeckhout & Kircher, 2011; Hagedorn, Law, & Manovskii, 2017; de Melo, 2018; Eeckhout & Kircher, 2018; **Herkenhoff et al., 2022**; Bilal et al., 2022]*

  ⇒ Microfoundations for changes in production → sorting

- **Wage inequality:** transformation of work & "firming up inequality"

  *[Autor, Levy & Murnane, 2003; Lin, 2011; Acemoglu & Autor, 2012; **Deming, 2017**; Acemoglu & Restrepo, 2018; Alon, 2018; **Neffke, 2019**; Jones, 2021; Atalay et al., 2021]*

  *[**Card et al., 2013**; Barth et al., 2016; Alvarez et al., 2018; **Bloom et al., 2019**; Aeppli & Wilmers, 2021; **Criscuolo et al. 2021**; Hakanson et al., 2021; Sorkin & Wallskog, 2021]*

  ⇒ Structural shifts within & between firms are connected

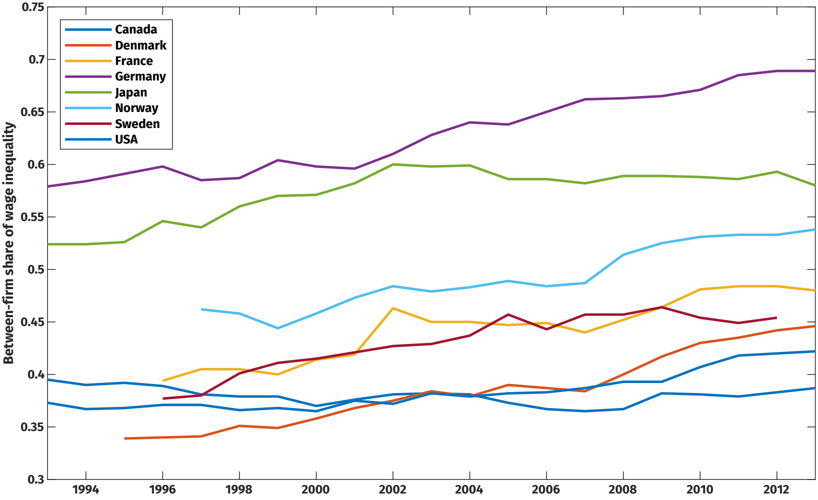
Einstein on “inevitable” specialization

“[K]nowledge has become vastly more profound in every department of science. But the assimilative power of the human intellect is and remains strictly limited. Hence it was inevitable that the activity of the individual investigator should be confined to a smaller and smaller section.”

[Einstein, 1941; quoted in Jones, 2021]

Firming up inequality: cross-country evidence (1)

▸ Main



Notes. Data from Tomaskovic-Devey et al. (2020).

Firming up inequality: different sample window

Main

- *Compression* in wage distribution in DE & PRT since $\approx$ GFC
- Re-compute role of between-firm component from start year to *respective peak* in wage dispersion
- Between-component accounts for $\approx$ 71% of $\uparrow$ to peak in DE; 55% in PRT

	Start year	Total	Between	Peak year	Total	Between	Total $\uparrow$	Share $\uparrow$ due to between
Germany	1993	0.177	0.078	**2011**	0.314	0.175	0.138	**0.709**
Portugal	1993	0.248	0.136	**2009**	0.306	0.167	0.058	**0.550**

Notes. Earnings measure varies across countries: daily wage (Germany), hourly wage (Portugal), annual labor earnings (USA). Employment unit is the establishment in Germany, the firm in Portugal and the USA. Results based on subsamples for full-time employed working-age men. Imputation of top-coded wages in Germany following Dustmann et al. (2009) and Card et al. (2013).
Source: own computations based on LIAB for Germany and Quadro de Pessoal for Portugal (preliminary, with Criscuolo & Gal), respectively; Song et al. (2019) for the USA.

Conceptual framework for decomposing wage dispersion: AKM

- AKM wage equation for log wage of a worker i employed at firm j in period t:

$$w_t^{i,j} = \alpha^i + \psi^j + \epsilon_t^{ij},$$

  - α^i: worker FE; ψ^j: firm FE; ϵ_t^{ij}: transitory earnings fluctuations
  - abstract from time-varying worker characteristics for exposition

- Variance decomposition following Song et al. (2021)

$$\begin{aligned}\mathrm{Var}(w_t^{i,j}) = & \underbrace{\mathrm{Var}(\alpha^i - \bar{\alpha}^j) + \mathrm{Var}(\epsilon_{i,j})}_{\text{within-component}} \\ & + \underbrace{\mathrm{Var}(\psi^j) + 2\mathrm{Cov}(\bar{\alpha}^j, \psi^j) + \mathrm{Var}(\bar{\alpha}^j)}_{\text{between-component}}\end{aligned}$$

  - $\mathrm{Var}(\psi^j)$: firm-specific pay premia
  - $\mathrm{Cov}(\bar{\alpha}^j, \psi^j)$: (worker-firm) sorting
  - $\mathrm{Var}(\bar{\alpha}^j)$: (worker-worker) segregation

PRT: Worker-firm sorting based on AKM & variations

Main

Detailed description of implementation available.

Evolution of the German task structure

Main

- Employment Surveys (ES) carried out by the German Federal Institute for Vocational Training (BIBB)
  - detailed information on tasks performed at work
  - individual-level, with consistent occupation codes
  - repeated cross-sections ranging from 1985/86 to 2018
  - large sample sizes (20,000-30,000 per wave)
- Methodology to study evolution of task content of production follows Spitz-Oener (2006), Antonczyk et al. (2009), Rohrbach-Schmidt & Tiemann (2013)
  - task classification
  - sample harmonization (West Germany, aged 15 to 65, employed)

Task classification

- Focus on Δ in usage of abstract/complex (non-routine, non-manual) tasks vs. "rest" (manual & routine)
  [Autor and Handel, 2009; Acemoglu & Autor, 2011; Rohrbach-Schmidt & Tiemann, 2013]
- Index capturing the usage of abstract/complex tasks for worker i in period t [Antonczyk et al., 2009]

$$T_{it}^{\text{complex}} = \frac{\text{number of activities performed by } i \text{ in task category "complex" in sample year } t}{\text{total number of activities performed by } i \text{ in sample year } t}$$

Task classification	Task name	Description
Complex	investigating	gathering information, investigating, documenting
	organizing	organizing, making plans, working out operations, decision making
	researching	researching, evaluating, developing, constructing
	programming	working with computers, programming
	teaching	teaching, training, educating
	consulting	consulting, advising
	promoting	promoting, consulting, advising
Other	repairing, buying, accommodating, caring, cleaning, protecting, measuring, operating, manufacturing, storing, writing, calculating	

Increase in aggregate task complexity, driven by within-occupation ↑

- Aggregate task intensity & decompose period-by-period change into:
  1. between component: Δ occupational employment shares
  2. within component: Δ task content within occupations

	Total	Between	Within	Within-share
1986 level	0.252			
1986-1992	0.025	0.002	0.022	0.906
1992-2006	0.298	0.057	0.241	0.809
2006-2012	0.019	0.002	0.017	0.890
2012-2018	0.053	0.028	0.025	0.476
Total change	0.395	0.089	0.306	0.775

Notes. Decompose changes in the usage of abstract tasks between periods t and $t-1$ according to $\Delta \bar{T}_t^{\text{abstract}} = \sum_o \omega_{o,t-1}(\bar{T}_{t,o}^{\text{abstract}} - \bar{T}_{t-1,o}^{\text{abstract}}) + \sum_o(\omega_{o,t} - \omega_{o,t-1})\bar{T}_{t,o}^{\text{abstract}}$ where $\bar{T}_{t,o}^{\text{abstract}}$ measures the average usage of abstract tasks by members of occupation o in period t and $\omega_{o,t}$ is the period- t employment share of occupation.

Large variation in task complexity across occupations

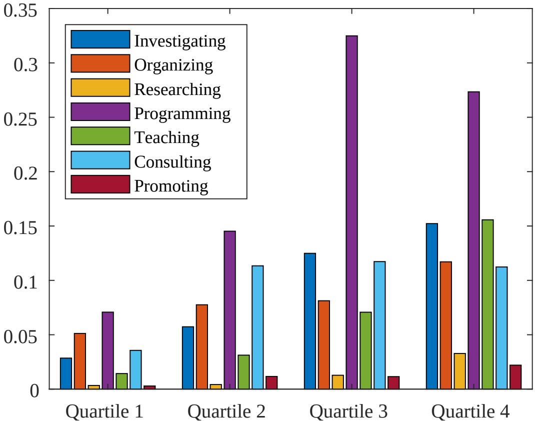
- Aggregate individual responses to 2-d occupation level, using 2012 & 2018 waves
  - 2012 & 2018: $\leftarrow$ ISCO-08 codes available

Occupation	$\bar{T}_o^{\text{complex}}$
Business and administration professionals	0.859
Legal, social and cultural professionals	0.830
Business and administration associate professionals	0.827
Administrative and commercial managers	0.820
Teaching professionals	0.807
...	...
Drivers and mobile plant operators	0.214
Agricultural, forestry and fishery labourers	0.193
Market-oriented skilled forestry, fishery and hunting workers	0.168
Food preparation assistants	0.131
Cleaners and helpers	0.124

Notes. Top-5 and bottom-5 ISCO-08 2-digit occupations when ranked by $\bar{T}_o^{\text{abstract}}$ in pooled 2012 & 2018 waves.

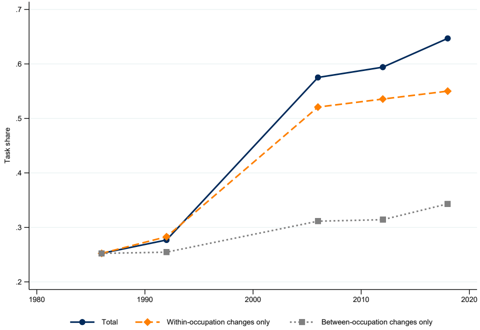
Time allocation

- **Potential concern:** miss “intensive” task usage margin
- **Partial answer:** Supplemental Survey from 2012 details the amount of time a subset of workers spent on the different tasks on a given day



Increase in aggregate task complexity, driven by within-occupation ↑ (graph)

- Aggregate task intensity & decompose period-by-period change into:
  1. between component: Δ occupational employment shares
  2. within component: Δ task content within occupations



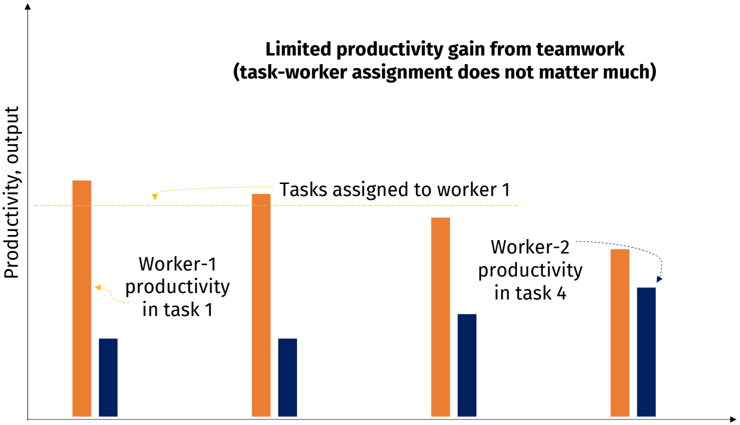
Notes. Decompose changes in the usage of abstract tasks between periods t and $t-1$ according to $\Delta\bar{T}_t^{\text{abstract}} = \sum_o \omega_{o,t-1}(\bar{T}_{t,o}^{\text{abstract}} - \bar{T}_{t-1,o}^{\text{abstract}}) + \sum_o(\omega_{o,t} - \omega_{o,t-1})\bar{T}_{t,o}^{\text{abstract}}$ where $\bar{T}_{t,o}^{\text{abstract}}$ measures the average usage of abstract tasks by members of occupation o in period t and $\omega_{o,t}$ is the period- t employment share of occupation.

Team production

Intuitive illustration: task assignment & complementarity

Main

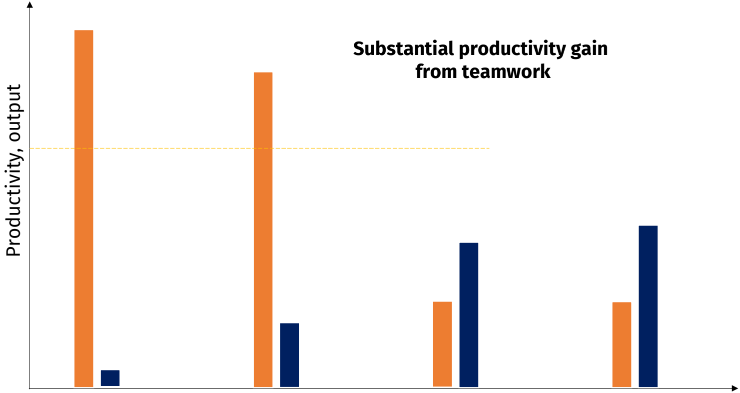
- Low specialization: good worker paired with bad coworker



Notes. Tasks are ordered by decreasing worker-1 comparative advantage, $\frac{z_1(\tau)/\lambda_1^L}{z_2(\tau)/\lambda_2^L}$.

Intuitive illustration: task assignment & complementarity

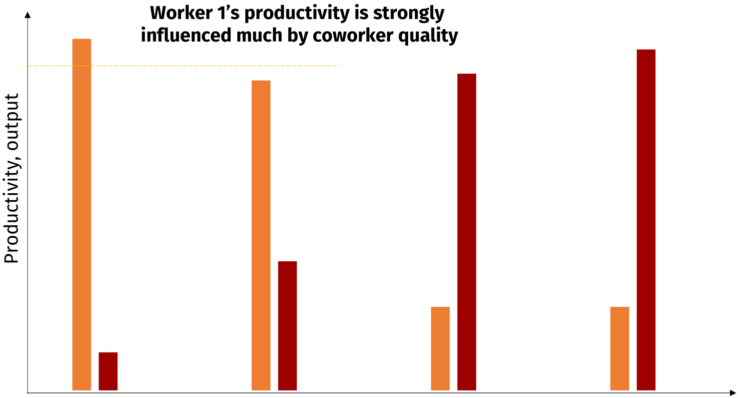
- High specialization: good worker paired with bad coworker



Notes. Tasks are ordered by decreasing worker-1 comparative advantage.

Intuitive illustration: task assignment & complementarity

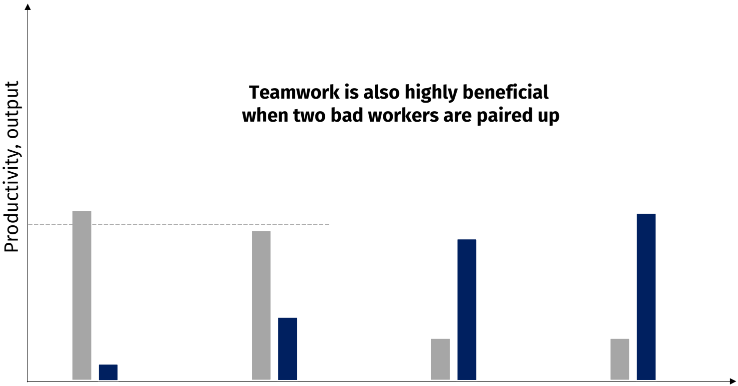
- High specialization: good worker paired with good coworker



Notes. Tasks are ordered by decreasing worker-1 comparative advantage.

Intuitive illustration: task assignment & complementarity

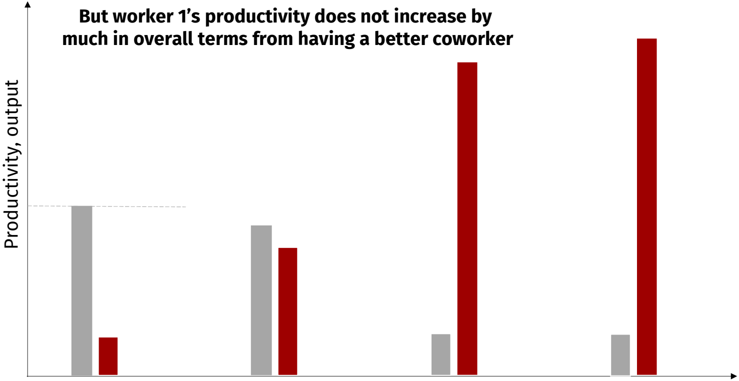
- High specialization: bad worker paired with bad coworker



Notes. Tasks are ordered by decreasing worker-1 comparative advantage.

Intuitive illustration: task assignment & complementarity

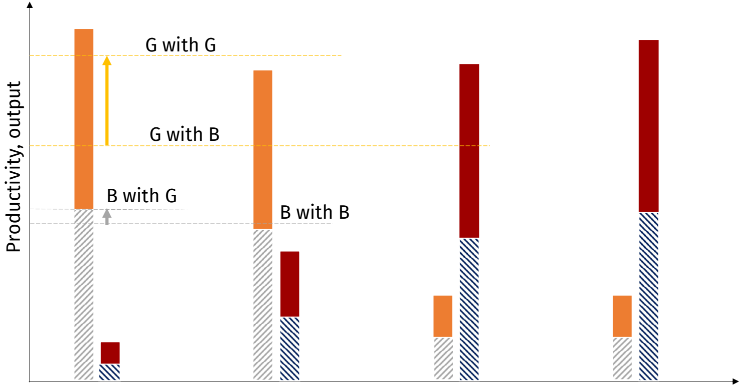
- High specialization: bad worker paired with good coworker



Notes. Tasks are ordered by decreasing worker-1 comparative advantage.

Intuitive illustration: task assignment & complementarity

⇒ Skill complementarity in production setting w/ specialization



Notes. Tasks are ordered by decreasing worker-1 comparative advantage.

No-division-of-labor benchmark

Main

Remark (No-division-of-labor benchmark)

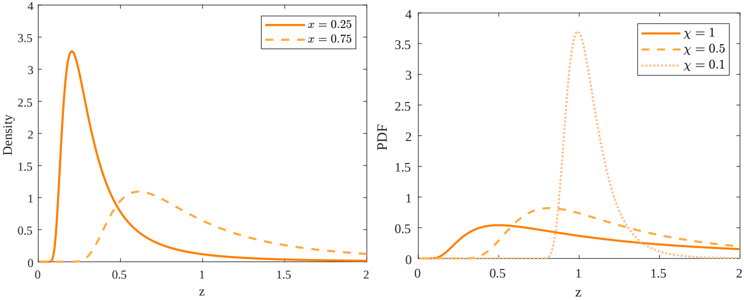
Suppose each team member produces the final output alone using the worker-specific, optimal combination of tasks given their respective, task-specific productivity draws.

Then irrespective of the value of χ (provided that $\frac{1}{\chi} - \eta + 1 > 0$), total output

$$Q = \sum_{i=1}^{N} x_i.$$

Each task is produced by every worker – $q_i(\tau) > 0 \; \forall i = 1, ..., N$ and $\forall \; \tau \in \mathcal{T}$ – and worker types are perfectly substitutable in production.

Visualizing the Fréchet distribution



(c) Scale parameter **(d)** Shape parameter

Figure 1: Illustration of properties of Fréchet distribution

Notes. This figure illustrates how the probability density function (PDF) of the Freéchet distribution varies with the scale and (inverse) shape parameter. The left panel assumes $\chi = \frac{1}{2}$. The right panel assumes $x = 1$. The illustration abstracts from the normalization factor ι, which also depends on η.

Corollary: output effects of $\chi\uparrow$

- $+$ "Efficiency gains": potential benefit from team production $\uparrow$ in χ
- $-$ "Complementarities": output more sensitive to low-quality members' types as $\chi\uparrow$
  - power mean of the underlying x_i
  - even high-x workers don't know how to efficiently perform some tasks

Corollary (Team production: output effects of changes in χ)

Up to 2^{nd}-order an increase in χ strictly increases team output provided that skill dispersion is sufficiently low relative to the average skill level;
that is, if

$$\frac{\sigma_x^2}{\bar{x}} < \frac{2\ln(N)N^{\chi+1}\bar{x}}{N^{\chi+1}\frac{1}{\chi^2(\chi+1)^2}+\ln(N)N^{\chi+1}(1-\frac{1}{\chi+1})},$$

where $\bar{x}=\frac{1}{N}\sum_{i=1}^{N}x_i$ *and* $\sigma_x^2=\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2$.

Taylor approximation to CES

Back to team production Back to team hiring

- Analytically tractable version of the hiring block:

$$f(x_1, x_2) = x_1 + x_2 - \xi(x_1 - x_2)^2$$

  where parameter ξ controls the degree of complementarity

- Justification: CES and link to team production model
  - in the $\kappa = 1$ special case, ξ maps onto $\frac{\chi}{\chi+1}$ (up to scale)

Remark (Second-order Taylor approximation to CES)

The second-order Taylor approximation to $f(x_1, x_2) = (\frac{1}{2}x_1^\gamma + \frac{1}{2}x_2^\gamma)^{1/\gamma}$ *around* $(\bar{x}, \bar{x})$ *with* $\bar{x} = \frac{x_1+x_2}{2}$ *is*

$$\bar{x} - \frac{1}{2} \underbrace{(1-\gamma)}_{\approx \xi} \frac{\sigma_x^2}{\bar{x}},$$

where $\sigma_x^2 = (\frac{x_1-x_2}{2})^2$.

Extension: team production with communication costs

Details

- Until now we assumed that firm organization is optimal $\approx$ the division of labor incurred no output losses due to coordination frictions
- But implementing the division of labor may $\downarrow$ time available for task production because of *communication* requirements

  [Becker & Murphy, 1992; Deming, 2017]
- **Extension** allowing for such coordination costs shows:
  1. qualitative link between technology & coworker complementarities exists *unless* division of labor is prohibitively costly
  2. $\chi \uparrow \Rightarrow$ importance of organizational quality for productivity $\uparrow$
- Ongoing research
  - interpretation of ex-ante firm heterogeneity: differences in organizational/management quality
  - organizational quality as a bottleneck in a knowledge economy
  - org diagram + network communication data from Fortune-100 company

Team production with communication costs...

Main

- Baseline: firm organization is optimal $\approx$ the division of labor incurred no output losses due to coordination frictions

- We'll treat this as "efficient benchmark" $\approx$ firm type $y = 1$

- But with **suboptimal team organization**, division of labor may $\downarrow$ time available for task production because of *communication* requirements

$$1 = \int_{\mathcal{T}} l_i(\tau) d\tau + c_i,$$

where c_i is time communicating with coworkers, with $\mathcal{S} = \{1, ..., N\}$:

$$c_i = \sum_{n \in \mathcal{S} \setminus i} c_{in},$$

- Next: stylized (!) extension to team production w/ communication $\Rightarrow$ characterization of firm heterogeneity as "organizational quality"

...and firm heterogeneity

Assignment problem

- Each worker produces Q_i from tasks $\{\tilde{q}_i(\tau)\}_{\tau \in \mathcal{T}}$ (with CES agg.)
- The total quantity of task τ used by some worker n is

$$\tilde{q}_n(\tau) = \sum_{i \in \mathcal{S}} q_{in}(\tau),$$

  where $q_{in}(\tau)$ is the quantity of task τ produced by worker i and used for production by worker n, s.t. an adding-up constraint

$$\sum_{n \in S} q_{in}(\tau) \leq q_i(\tau).$$

- The extent to which the division of labor absorbs communication time depends on how well the team is organized $\approx$ firm type y

$$c_{in} = \int_{\mathcal{T}} \frac{q_{in}(\tau)}{z_i(\tau)} \times \left(\frac{1}{y} - 1\right) d\tau$$

  if $i \neq n$; naturally, $c_{ii} = 0 \quad \forall i \in \mathcal{S}$
  - time spent by i communicating with n is proportional to the time spent by i solving problems for n

Key result in tractable special case

! Task complexity $(\chi) \uparrow \Rightarrow \uparrow$ importance of organizational quality for output

Proposition (Team production: frictionless labor markets)

Suppose that $x_i = x \ \forall i \in \mathcal{S}$. Then

1. *team output as a function of workforce type x, team size N, and organizational quality y is*

$$Q = Nx\left(1 + (N-1)y^{1/\chi}\right)^{\chi};$$

2. *the fraction of tasks necessary for production produced by any given worker is*

$$\frac{1}{1 + (N-1)y^{1/\chi}};$$

3. *and the elasticity of labor productivity Q/N w.r.t y is*

$$\frac{(N-1)y^{1/\chi}}{1 + (N-1)y^{1/\chi}} \geq 0.$$

Team hiring

Environment: demographics & preferences & production technology

Main

- **Time**: continuous
- **Agents:** workers & firms
  - unit mass of workers, types uniformly distributed $x \in \mathcal{X} = [0, 1]$
  - m_f mass of firms, type $y \in \mathcal{Y}$, potentially ex-ante heterogeneous
  - agents indexed by *ranks* of their respective productivity distribution, hence uniform type distribution [Hagedorn et al., 2017]
  - all agents are infinitely-lived, risk-neutral, share a common discount rate ρ, max. the present value of payoffs
- **Production technology**: firms are vacant or have 1 or 2 workers
  - normalize team size to max. $N = 2 \leftarrow$ key is "existing workforce" & "potential hire"
  - convention: from x's perspective, let x' index *coworker*
  - team production: $f(x, y, x') \leftarrow$ step 1; strictly increasing in each argument, symmetric in worker types
  - 1-worker: $f(x, y)$, short for $f(x, y, \emptyset)$

Environment: random search & wage bargaining

- **Timing** within dt-intervals
  1. exogenous separation: Poisson rate δ
  2. random search & matching
  3. production & surplus sharing
- **Meeting process**: unemployed meet *some* firm at Poisson rate M_u
  - probability for a firm to be contacted by an(y) unemployed: $M_f = M_u U$
  - no on-the-job search
  - for simplicity: today don't allow 2-worker firms to fire existing employee and replace with better matching worker
- **Matching** decisions based on joint surplus b/w firm & worker(s): privately efficient *[cf. Bilal-Engbom-Mongey-Violante, 2021]*
- **Surplus sharing**: firm bargains with potential new hire, taking into account coworker complementarities; worker bargaining power ω
  - continuous renegotiation, as if each worker is marginal (i.e., outside option: unemployment)

Key hiring decision

- **Notation:**
  - Value functions for unemployed worker, $V_u(x)$; worker x employed at y with coworker x', $V_e(x,(y,x'))$
  - Value functions for vacant firm, $V_v(y)$; firm y producing with x and x': $V_p(x,y,x')$
  - $d_u(x)$: density of unemployed workers of type x
- **Key question:** which type(s) of workers is a firm that already has one worker willing to hire – trading off match quality vs. cost of searching
- HJB for the **joint value** of a firm y with worker x, $\Omega(x,y)$

$$\rho\Omega(x,y) = f(x,y) + \delta\big[-\Omega(x,y) + V_u(x) + V_v(y)\big]$$
$$+ M_f \int \frac{d_u(\tilde{x}')}{U} \underbrace{\max\{-\Omega(x,y) + V_e(x,(y,\tilde{x}')) + V_p(x,y,\tilde{x}'),0\}}_{=(1-\omega)S(\tilde{x}',(y,x))^+} d\tilde{x}'$$

$\Rightarrow$ hiring decision based on *marginal surplus*

Stationary equilibrium

▸ Equilibrium equations

- Remainder of setup is fairly straightforward but lengthy
- Formally, after defining (i.) HJBs for unemployed & vacant & surplus values, and (ii.) Kolmogorov Forward Equations (KFEs) describing the evolution of the distribution of agents across states:

Definition

A stationary search equilibrium is a tuple of value functions together with a stationary distribution of agents across states such that (i.) the value functions satisfy the HJB Equations given the distribution; and (ii.) the distribution satisfies the KFEs given the policy functions implied by the value functions.

- Needs to be computed **numerically**
  - agents' expectations & decisions must conform w/ population dynamics to which they give rise; as distribution evolves, so do agents' expectations
  - mean-field game

Environment: firm & worker states

- Distribution across states for a **worker** type x:

$$d_w(x) = d_u(x) + \int d_m(x, \tilde{y}) d\tilde{y} + \int \int d_m(x, \tilde{y}, \tilde{x}') d\tilde{y} d\tilde{x}'$$

  - $d_u(x)$: 'density' of unemployed of type x
  - $d_m(x, y)$, shorthand for $d_m(x, y, \emptyset)$: 'density' of matches b/w x and y types
  - $d_m(x, y, x')$: 'density' of "team matches" b/w x and y with coworker x'

- Distribution across states for a **firm** type y:

$$d_f(y) = d_v(y) + \int d_m(\tilde{x}, y) d\tilde{x} + \frac{1}{2} \int \int d_m(\tilde{x}, y, \tilde{x}') d\tilde{x} d\tilde{x}'$$

  - $\frac{1}{2}$: account for 1 firm having 2 workers

- **Aggregates** can be backed out, e.g.

$$U = 1 - \left(\int d_m(\tilde{x}, \tilde{y}) d\tilde{y} + \int \int d_m(\tilde{x}, \tilde{y}, \tilde{x}') d\tilde{y} d\tilde{x}' \right) d\tilde{x}$$

Environment: surplus sharing

- **Desiderata:**
  1. hiring decisions are based on the joint value to all parties affected and, thus, constrained Pareto efficiency in allocations
  2. order of hiring does not matter
  3. wages ware continuously renegotiated

- **One-worker firm:**

$$(1-\omega)\Big(V_e(x,y) - V_u(x)\Big) = \omega\Big(V_p(x,y) - V_v(y)\Big) \tag{1}$$

Two-worker firm

$$\begin{aligned}&(1-\omega)\Big(V_e(x,(y,x')) - V_u(x)\Big) \\ &= \omega\Big(V_p(x,y,x') + V_e(x',(y,x)) - V_p(x',y) - V_e(x',y)\Big)\end{aligned} \tag{2}$$

HJB: unmatched

▸ Main

- Unmatched firm:

$$\rho V_v(y) = (1-\omega) M_f \int \frac{d_u(\tilde{x})}{U} S(\tilde{x}, y)^+ d\tilde{x} \tag{3}$$

- Unmatched worker:

$$\rho V_u(x) = b(x) + M_u \omega \left[\int \frac{d_v(\tilde{y})}{V} S(x, \tilde{y})^+ + \int \frac{d_m(\tilde{x}', \tilde{y})}{V} S(x, (\tilde{y}, \tilde{x}'))^+ d\tilde{x}' \right] d\tilde{y} \tag{4}$$

HJB: surpluses

- Surplus of coalition (x, y)

$$(\rho + \delta)S(x, y) = f(x, y) - \rho\big(V_u(x) + V_v(y)\big) + M_f(1-\omega)\int \frac{d_u(\tilde{x}')}{U} S(\tilde{x}', (y, x))^+ d\tilde{x}'. \quad (5)$$

- Marginal surplus from adding x to (y, x')

$$(\rho + 2\delta)S(x, (y, x')) = f(x, y, x') - \rho\big(V_u(x) + V_u(x') + V_v(y)\big) + \delta S(x, y) - (\rho + \delta)S(x', y) \quad (6)$$

KFE: unemployed

$$\delta\Big(\int d_m(x,\tilde{y})d\tilde{y} + \int d_m(x,\tilde{y},\tilde{x}')d\tilde{x}'d\tilde{y}\Big) = d_u(x)M_u \Big(\int \frac{d_v(\tilde{y})}{V} h(x,\tilde{y}) + \int \frac{d_m(\tilde{x}',\tilde{y})}{V} h(x,(\tilde{y},\tilde{x}'))d\tilde{x}'d\tilde{y}\Big) \tag{7}$$

KFE: one-worker matches

$$d_m(x,y)\left(\delta + M_f \int \frac{d_u(\tilde{x}')}{U} h(\tilde{x}',(y,x))d\tilde{x}'\right) = d_u(x) M_u \frac{d_v(y)}{V} h(x,y) + \delta \int d_m(x,y,\tilde{x}')d\tilde{x}' \quad (8)$$

KFE: two-worker matches

$$2\delta d_m(x,y,x') = d_u(x) M_u \frac{d_m(x',y)}{V} h(x,(y,x')) + d_u(x') M_u \frac{d_m(x,y)}{V} h(x',(y,x)) \tag{9}$$

Lemma: monotonicity of unemployment value and wage in x

Main

Lemma

The value of unemployment $V_u(x)$ is monotonically increasing in x, and so is the wage function $w(x,(y,x'))$.

Proof.

See paper appendix. Key: surplus representation. $\square$

Stylized

Frictionless matching: assignment and payoffs

Main

- Working backwards, let's first pin down the frictionless payoffs that determine the outside option

- The equilibrium of the frictionless model can be derived in many ways

- Equilibrium assignment and payoffs:
  - PAM: $\mu(x) = x$ *given supermodular* $f(x_1, x_2)$
  - wage schedule obtained from integrating over FOC

    $$w^*(x) = \int_0^x f_x(\tilde{x}, \mu(\tilde{x}))d\tilde{x}$$

    where integration constant is zero due to $f(0,0) = 0$
  - firm payoffs in this formulation are 0

- Given $f(x_1, x_2) = x_1 + x_2 - \gamma(x_1 - x_2)^2$, with $\gamma > 0$, we have

$$\mu(x) = x$$

$$w^*(x) = x \quad \text{and } v^* = 0$$

Characterization results: conditional distribution

Main

Lemma (Conditional type distribution)

Given a threshold distance s, the conditional distribution of coworkers for $x \in \mathcal{X}$ is

$$\Phi(x'|x) = \begin{cases} 0 & \text{for } x' < \sup\{0, x-s\} \\ \frac{x - \sup\{0, x-s\}}{\inf\{x+s,1\} - \sup\{0, x+s\}} & \text{for } x' \in [\sup\{0, x-s,\}, \inf\{x+s, 1\}] \\ 1 & \text{for } x' > \inf\{x+s, 1\} \end{cases}$$

Characterization results: between-share of wage variance

▸ Main

Corollary

Given a threshold distance s and a value of γ, the between-firm share of the variance of wages is equal to

$$\frac{-\frac{13\,\gamma^2\,s^5}{2400}+\frac{\gamma^2\,s^4}{80}+\frac{5\,s^3}{36}-\frac{s^2}{6}+\frac{1}{12}}{\frac{\gamma^2\,s^4}{45}-\frac{4897\,\gamma^2\,s^5}{10800}-\frac{\gamma^2\,s^6}{324}+\frac{19\,\gamma^2\,s^5\,\ln(2)}{30}+\frac{1}{12}}.$$

Characterization using a stylized model: setup

Next: intuition for how coworker complementarities shape matching can be gained from a stylized model $\rightarrow$ closed-form solutions

- **Simplified setup:**
  - no ex-ante firm heterogeneity; mass of firms $m_f = \frac{1}{2}$
  - production fn.: $f(x, x') = x + x' - \xi(x - x')^2$, where ξ controls complementarity (cost of mismatch)

    ▸ Approx. of microfounded team prod. fn.
  - no production with 1 employee & abstract from team production benefits
  - firm has no bargaining power, workers each receive their outside option plus half the surplus
- **Explicit search costs**: no discounting, guaranteed match ($M_u = M_f = 1$); but type-invariant worker search costs c
  - supermodularity in f suffices for PAM *[Atakan, 2006]*
- **Stylized** timing & specification of final stage $\downarrow$ *[cf. Eeckhout & Kircher, 2011]*

Characterization using a stylized model: timing

▸ Frictionless stage-2 outcome

0. Each firm is randomly paired with one worker $x' \in \mathcal{X}$
   - remaining: mass $\frac{1}{2}$ of uniformly distributed workers; mass $\frac{1}{2}$ of firms with one employee

1. Each (firm + x') unit is randomly paired with a worker $x \in \mathcal{X} \to$ decision:
   a. match: form a team
      + produce & share production value
      – no further actions and zero payoff in stage 2

      or

   b. search: don't form a team
      – workers pay search cost c
      + all have opportunity to re-match in stage 2, s.t.

2. frictionless matching b/w unmatched firms & workers; production
   - pure PAM: x works with x ($\leftarrow$ deterministic coupling $\mu(x) = x$)
   - payoffs *given* pure PAM: $w^*(x) = x$ and $v^* = 0$

Characterization using a stylized model: stage-1 matching decision

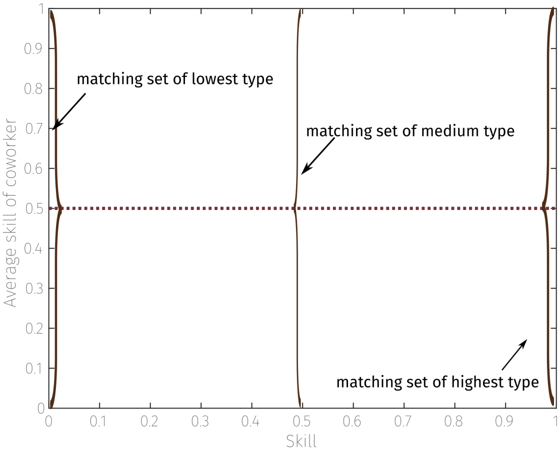
- A firm with employee x' that meets a worker of type $x \in \mathcal{X}$ decides to hire her, i.e. $h(x, x') = 1$, if

$$\underbrace{\overbrace{f(x,x')}^{\text{match}} - \overbrace{\left[w^*(x) + w^*(x') + v^* - 2c^w\right]}^{\text{search}}}_{\equiv S(x,x')} > 0$$

- Threshold distance s^* s.t. $h(x, x') = 1 \Leftrightarrow |x' - x| < s^*$
- Threshold distance satifies: $s^* = \sqrt{2c/\xi}$
  - greater complementarities (ξ) render the matching set *narrower*
  - greater search costs (c) render the matching set *wider*

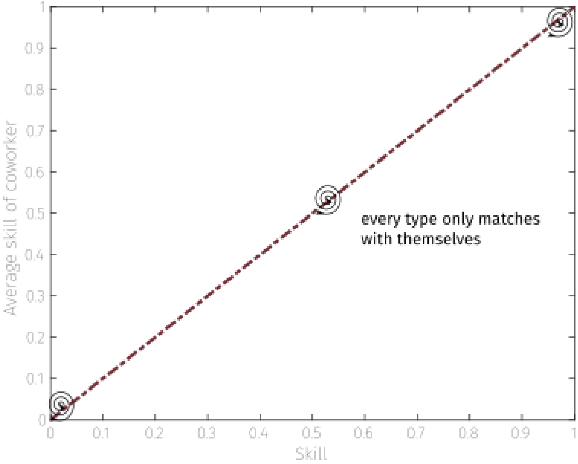
Characterization using a stylized model: matching decisions

- Random matching ($s^* = 1$)



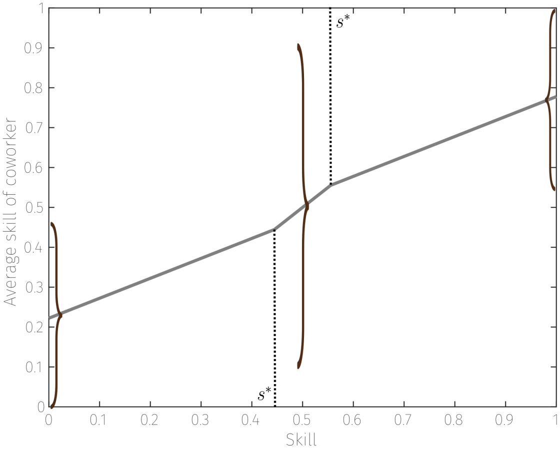
Characterization using a stylized model: matching decisions

- Perfectly assortative matching (PAM; $s^* = 0$)



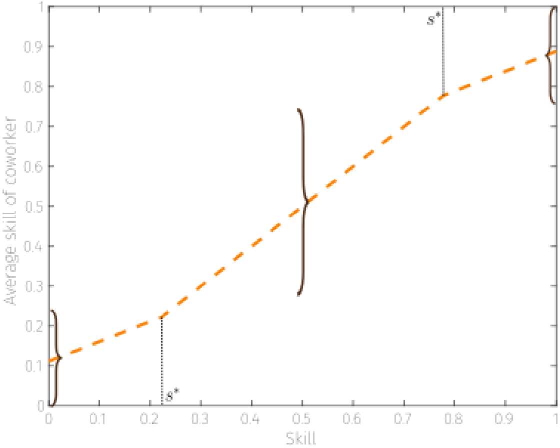
Characterization using a stylized model: matching decisions

- Matching under weak complementarities ($s^* = 0.45$)



Characterization using a stylized model: matching decisions

- Matching under stronger complementarities ($s^* = 0.225$)



Characterization using a stylized model: corollary

Main

Corollary (Stylized model)

For a given threshold s, which is decreasing in χ:

1. *the coworker correlation is:* $\rho_{xx} = (2s+1)(s^2-1)^2$;
2. *the average coworker type is*

$$\hat{\mu}(x) = \begin{cases} \frac{x+s}{2} & \text{for} \quad x \in [0, x+s) \\ x & \text{for} \quad x \in [x-s, x+s] \\ \frac{1+x-s}{2} & \text{for} \quad x \in (x-s, 1]. \end{cases}$$

3. *the between-firm share of the variance of wages is decreasing in* s

Estimation

Overview: estimation

Main

- **Review:** micro-origins and macro-distributional implications of coworker complementarities through the lens of a model that has empirical support
- **Goal:** estimate the model to facilitate quantitative analysis
- **This section:**
    1. methodology
    2. validation
    3. results for 1990 & 2010

Methodology: overview

- **Baseline economy:** Germany, $\approx$ 2010, monthly frequency

- Strategy
  1. functional-form assumptions + externally calibrate some parameters: ρ, team-benefit factor, bargaining power ω, mass of firms m_f, unemployment flow benefit
  2. offline estimation of job destruction rate, δ
  3. online estimation of meeting rate & production parameters: **target micro-evidence on wage complementarities** + avg. wage level + variance of log wages

- Macro moments of interest (e.g., coworker sorting, between-firm share): untargeted!

Methodology: distributional & functional form assumptions

- **Focus on worker heterogeneity:** no ex-ante heterogeneity in firm productivity
- **Microfoundations** imply labor *productivity*, $\gamma = \frac{\chi}{1+\chi}$

$$f(\{x_i\}_{i=1}^N)/N = N^{\frac{1}{1-\gamma}-1}\left(\frac{1}{N}\sum_{i=1}^N x_i^{1-\gamma}\right)^{\frac{1}{1-\gamma}},$$

- **3 generalizations**:
  - allow for productivity constant a_0 so $f(x,\emptyset) = a_0 + a_1 x$
  - lower bound on γ given $\chi \in [0,1]$: $\frac{1}{2}$; here allow for more flexibility
  - teamwork efficiency gain: controlled by a_2, not moving with χ
- **Team production:**

$$f(x,x') = 2a_2\left[a_0 + a_1\left(\frac{1}{2}(x)^{1-\gamma} + \frac{1}{2}(x')^{1-\gamma}\right)^{\frac{1}{1-\gamma}}\right]$$

Methodology: indirect inference

▸ Identification validation exercises

- Externally calibrate ρ, ω, b_1, a_2; and directly target δ to match EU rate
- Jointly estimate parameter vector $\psi = \{\gamma, a_0, a_1, \lambda_u\}$ that minimizes

$$\mathcal{G}(\psi) = \sum_{k=1}^{4} \left(\frac{\hat{m}_k - m_k(\psi)}{\frac{1}{2}|\hat{m}_k| + \frac{1}{2}|m_k(\psi)|} \right)^2,$$

  where $\hat{m}_k$: empirical moment; $m_k(\psi)$: model counterpart

- **Targets & main parameter(s)**
  - unemployment rate: $M_u \downarrow$
  - average log wage: $a_0 \uparrow$ & $a_1 \uparrow$
  - variance of log wages: $a_0 \downarrow$ & $a_1 \uparrow$
  - avg. cross-partial derivative $\widehat{\frac{\partial w^2(x,x')}{\partial x \partial x'}}$: $\chi \uparrow$ & $a_1 \uparrow$
- Repeat this for **2010** *and* **1990**

Parameterization: 2010

Parameter	Description	Value	Source
ρ	Discount rate	0.01	External
ω	Worker bargaining weight	0.50	External
b_1	Home production proportionality	0.7	External
a_2	Team benefit	1.1	External
δ	Separation rate	0.01	Offline estimation
γ	Elasticity of complementarity	0.62	Internal estimation
a_0	Production constant	3.61	Internal estimation
a_1	Skill sensitivity	19.54	Internal estimation
M_u	Meeting rate	0.30	Internal estimation

Table 1: Model parameters for 2010

Notes. This table summarizes the baseline calibration of the model, including parameter estimated internally to match empirical moments for 2010.

Estimation: model fit

▸ Untargeted: sorting ▸ Untargeted: wage functions

- **Untargeted**: good fit to coworker sorting patterns & wage functions

		2010			1990		
Parameter	Targeted	Value	m	$\hat{m}$	Value	m	$\hat{m}$
γ	$\widehat{\frac{\partial w^2(x,x')}{\partial x \partial x'}}$	0.62	0.134	0.134	0.54	0.067	0.067
a_0	Avg. log wage	3.61	2.58	2.58	6.62	2.67	2.67
a_1	Var. log wage	19.54	0.20	0.20	15.15	0.09	0.09
M_u	Unemp. rate	0.30	0.05	0.05	0.28	0.05	0.05

Table 2: Estimated parameters and targeted moments

Notes. This table lists for each of the estimated parameters the targeted moment and, separately for 2010 and 1990, the estimated value, and moments in data and model. The model-implied moment is m, while $\hat{m}$ is the data counterpart. All model-implied moments are constructed based on the subset of workers employed in teams. All wage-related moments are reported in terms of hourly wages (using a scaling factor of 7.5 working hours per day).

Evidence

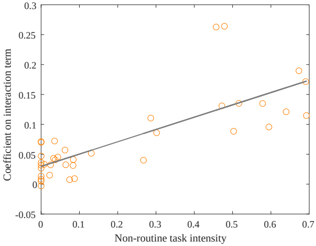
Overview of empirical validation tests

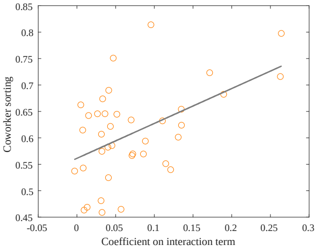
▸ Main

1. **Prediction #1: coworker complementarities $\nearrow$ task complexity**
   - ☑occupations performing more complex exhibit stronger coworker wage complementarities
   - ☑coworker wage complementarities are (weakly) monotonically increasing in the layer of a firm's internal hierarchy
2. **Prediction #2: $\uparrow$ coworker complementarities $\Rightarrow \uparrow$ positively assortative matching of coworkers**
   - ☑coworker sorting is more pronounced in occupations exhibiting stronger coworker wage complementarities
   - ☑industries with production processes featuring relatively high skill complementarity also tend to be characterized by greater coworker sorting
3. **Premise: differences across firms in the quality of their workforce are a key source of between-firm inequality in productivity and pay**
   - extensive literature *[Card et al., 2013; Berlingieri et al., 2017; Song et al., 2019; Hakanson et al., 2021; Sorkin-Wallskog, 2021]*
   - ☑measures of between-firm inequality in productivity and wages are increasing in the degree of coworker sorting

Result 2: complexity $\Rightarrow$ complementarities $\Rightarrow$ sorting (occupations)

- Occupations performing ↑ non-routine abstract tasks exhibit ↑ coworker wage complementarities
    - proxy for χ: occupation's reliance on non-routine, abstract (NRA) tasks *[Deming, 2017]*, measured using task indices from Mihaylov and Tidens (2019)
- Occupations exhibiting ↑ coworker wage complementarities feature ↑ coworker sorting
    - coworker wage complementarities & sorting measured at the ISCO-08 2-digit occupation level



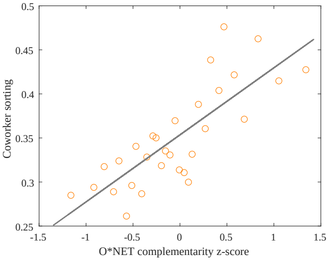


Notes. Analysis at the level of ISCO-08 2-digit occupations.

Result 3: complementarities ⇒ sorting & inequality (industries)

▸ Prod. disp.

⇒ Industries which according to a literature-based proxy feature relatively ↑ skill complementarity also tend to be characterized by ↑ coworker sorting and between-firm inequality in productivity & pay



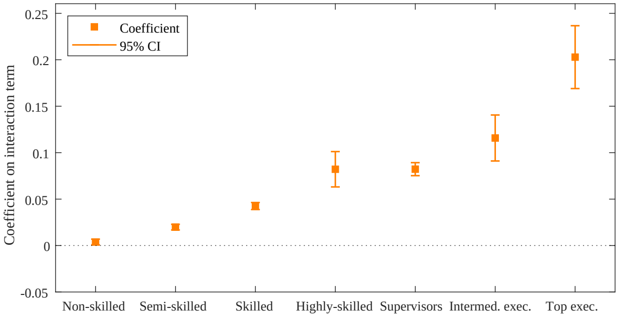
(a) Coworker sorting **(b)** Between-share of wage inequality

Notes. Analysis at the NACE-4-digit industry level. The x-axis measures an industry-level proxy for complementarity: the weighted mean score of an occupation-level index constructed from O*NET data measuring the importance of: teamwork, impact on coworker output, communication, and contact *[Bombardini et al., 2012]*.

Result 4: complexity $\Rightarrow$ complementarities (hierarchical layers)

Approx. for DE

$\Rightarrow$ Coworker wage complementarities are (weakly) monotonically $\uparrow$ in the layer of a firm's hierarchy



Notes. This figure reports the point estimate for the coefficient β_3, alongside confidence intervals, when the wage regression is estimated separately for each of 7 hierarchically ordered layers *[cf. Caliendo et al., 2020]*, with coworkers defined by membership in same layer at a given employer.

Additional results

- Coworker wage complementarities also ↑ over time, as in LIAB ▸ Jump
- Industries that score high on the O*NET-based complementarity index also feature high dispersion in *firm productivity* ▸ Jump
- Measures of between-firm inequality in productivity and pay are increasing in the degree of coworker sorting ▸ Jump
- Similar ↑ in worker segregation as in DE but between-firm share flat due to falling firm-specific pay premia ▸ Jump
- Non-parametric measure of coworker sorting: close to model and ↑ over time ▸ Jump

Data sources: short description of main datasets

▸ Main ▸ Imputation procedure

- Germany: Linked-Employer-Employee-Data of the IAB (LIAB), 1993-2017
  - interviews from IAB Establishment Panel + related *establishment* and individual data generated in administrative processes
  - built up from a 2% sample of establishments, but includes comprehensive employment biographies of individuals employed at these establishments
  - worker info includes: (real) **daily wage**, occupation and education; establishment info not yet utilized
  - main drawbacks: (i) large but not universe and no representative worker panel; (ii) **top coding** (affects >50% of university-educated men in regular full-time employment) $\rightarrow$ adopt standard imputation methods (Dustmann et al., 2009; CHK, 2013); (iii) establishment productivity data is survey based

- Portugal: Quadros de Pessoal & Relatório Único, 1986-2017
  - $\approx$ universe of private sector firms and workers employed by them
  - annual panel
  - worker information includes: detailed earnings measures (base wage, regular benefits, irregular benefits (performance-pay, bonuses, etc.), overtime pay); no top-coding; also includes information on hours worked within the month (regular and overtime) $\Rightarrow$ (real) **total hourly wage**
  - firm information includes balance sheet data from 2004 onward
  - main drawbacks: (i) poor quality of personal identifiers up to 2002; (ii) break in firm identifiers from 2010; (iii) changes in classifications of occupations and industries with limited official crosswalks

Sample restrictions

- Data cleaning $\Rightarrow$ broadly harmonized samples
- Main restrictions
  - men $\leftarrow$ TBD
  - age 20-60
  - full-time employed
  - drop agriculture, public sector, utilities industries
  - firms (and their employees) with at least 10 employees
- DE: West Germany
- PRT: at least. 80% of official minimum wage

Imputation procedure for Germany

Main

- Follow imputation approach in CHK2013, building on Gartner et al. (2005) and Dustmann et al. (2009)
  1. fit a series of Tobit models to log daily wages
  2. then impute an uncensored value for each censored observation using the estimated parameters of these models and a random draw from the associated (left-censored) distribution

- Currently I fit 16 Tobit models (4 age groups, 4 education groups) *after* having restricted the sample (to include West German men only, in particular) and I follow CHK in the specification of controls by including not only age, firm size, firm size squared and a dummy for firms with more than ten employees, but also the mean log wage of co-workers and fraction of co-workers with censored wages. Finally, following Dauth & Eppselheimer (2020) I limit imputed wages at 10×99th percentile.

Minceriana returns to education: Mundlak decomposition

▸ Main

- Series of cross-sectional regressions of log wages (w_{it}) on years of schooling (S_{it}) of the form $w_{it} = a_t + b_t S_{it} + u_{it}$, where the error u_{it} includes an establishment-level component: $u_{it} = \mu_{J(it),t} + v_{it}$.
- Ignoring this group component in estimation would yield a population regression coefficient $b_t^{OLS} = \frac{\text{Cov}(w_{it}, S_i)}{Var(S_i)} = b_t + \lambda_t c_t$
- So, the standard (Mincerian) return to schooling in each year consists of 3 distinct components:
  1. b_t: return to schooling conditional on job quality
  2. $\lambda_t = \frac{\text{Cov}(\mu_{J(i,t),t}, \bar{S}_{J(i,t),t})}{var(\bar{S}_{J(i,t)})}$: between-establishment return to schooling/return to mean-establishment schooling
  3. $c_t = \frac{\text{Cov}(S_i, \bar{S}_{J(i,t),t}))}{Var(S_i)}$: Kremer-Maskin index of sorting of schooling
- Using the results from Mundlak (1978) we may directly estimate each of these components by fitting regressions of the form:

$$w_{it} = a_t + b_t S_i + \lambda_t \bar{S}_{J(i,t),t} + e_{it}$$

and we can estimate the sorting index from the regression

$$\bar{S}_{J(i,t),t)} = d_t + c_t S_i + \eta_{i,t}$$

Cluster-based methodology: motivation

Main

- **Standard AKM approach** estimates large number of firm-specific parameters, identified solely off worker mobility $\Rightarrow$ incidental parameters problem $\approx$ limited mobility bias $\Rightarrow var(\psi) \uparrow$ & $cov(\psi, \alpha) \downarrow$
- **Bonhomme, Lamadon, and Malresa (2019**, Ecma): two-step grouped fixed-effects estimation
  1. Recover firm classes using k-means clustering, based on the similarity of their earnings distribution
  2. Estimate parameters of correlated random effects model by maximum likelihood, conditional on the estimated firm classes
- **Potential advantages**
  1. mitigate limited mobility bias
     – sufficient number of workers who move between any given cluster to identify the cluster fixed effects
  2. allows relaxing sample restrictions (n-connected set restriction when estimating group-specific firm/cluster FEs)
  3. *if* also take step 2, can estimate match complementarities between firms and workers (given estimated worker types from step 2)
- Limitation: loss of information by imposing homogeneity within

Cluster-based methodology: clustering step

- Obtain clusters by solving **weighted k-means problem**

$$\min_{k(1),\ldots,k(J),H_1,\ldots,H_K} \sum_{j=1}^{J} n_j \int \left(\hat{F}_j(w) - H_{K_j}(w)\right)^2 d\mu(w), \tag{10}$$

  - $k(1), ..., k(J)$: partition of firms into K known classes
  - $\hat{F}_j$: empirical cdf of log-wages in firm j
  - n_j: average number of workers of firm j over sample period
  - $H_1, ..., H_K$: generic cdf's

- Implementation here:
  - baseline value of $K = 10$, as in BLM, but experiment with $K = 20$ and $K = 100$
  - I use firms' cdf's over entire sample period on a grid of 20 percentiles

Cluster-based methodology: half-BLM

- "Half-BLM": take step (1), impute class to each worker-year observation, and then estimate two-way fixed effect wage regression using cluster effects instead of firm effects:

$$w_{it} = \alpha_i + \sum_{k=1}^{K} \psi_k^{g(i)} \mathbb{I}(J(i,t) = k) + \beta^{g(i)} X'_{it} + r_{it}$$

  - $\mathbb{I}(J(i,t) = k)$ are dummies indicating which cluster k firm $J(i,t)$ has been assigned to
  - superscript g in $\psi_k^{g(i)}$ (and β^g) indicates that cluster FE may be allowed to vary by group

- **Implementation**: estimation on rolling 2-year intervals
  - similar to Palladino et al. (2021), but also Lachowska et al. (2021) who estimate AKM models separately for successive two-year intervals and then apply the leave-out methodology of Kline et al. (2020) to correct for biases in the estimated variance components
  - experimented also with 5-year intervals to facilitate greater comparability with standard AKM approach, including FEs provided by IAB-FDZ

Evidence from the literature: coworker effects

- **Model**: teamwork in production environments involving specialization generates complementarities between coworkers' abilities
- Neffke (2019):
  - detailed Swedish micro data: occupation & educational specialization
  - **coworker effects are important in many knowledge-intensive jobs** (e.g. health care, engineering) + professional occupations (e.g., lawyers); and in skill-intensive (e.g. R&D) + crafts-based industries (e.g. construction)
- Bloesch, Larsen, and Taska (2022):
  - occupation-level measure of within-firm task differentiation across positions, using Burning Glass Technologies US vacancy data
  - **managerial + professional jobs are the most differentiated**, followed by technicians; service + manual jobs have the lowest differentiation scores
- Peer effects literature: mixed *[Mas & Moretti, 2009; Azoulay et al., 2010; Jaravel et al., 2018; Nix, 2020; Cornelissen et al., 2017; Cardoso et al., 2018]*

Evidence from the literature: sorting

- **Model**: workforce composition & reallocation as a key channel behind between-firm inequality
- Hakanson et al. (2021): use *direct* measures of skills and $\uparrow$ between-firm skill inequality is a key driver behind $\uparrow$ between-firm wage inequality (SWE)

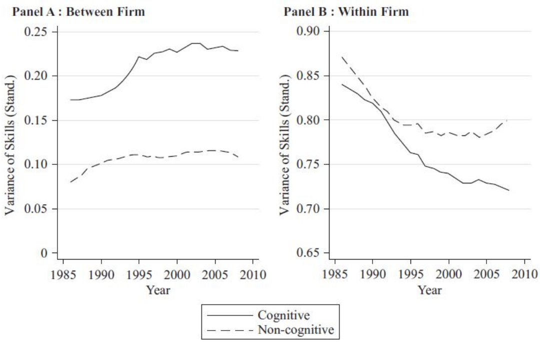
Details

- Studies of $\uparrow$ between-firm wage inequality: key role played by widening gap in **workforce composition** *[CHK, 2013; Song et al., 2019]*
  - typically, evidence based on AKM-style regressions
- Productivity dispersion, between-firm earnings inequality, and coworker sorting are successively higher for more recent **cohorts** of firms in the U.S. *[Sorkin and Wallskog, 2021]*
- Polarization in productivity and labor market outcomes are correlated across **sectors** *[Berlingieri et al., 2017]*

Evidence from the literature (2): Hakanson et al. (2021)

▸ Main

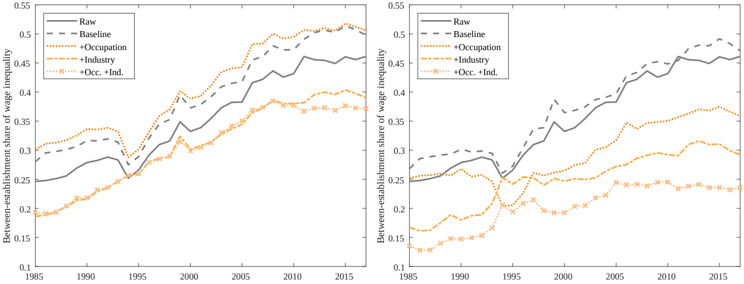
- *Direct* measures of cognitive and non-cognitive skills across Swedish firms during 1986–2008, using test data from military enlistment
- ↑ in sorting can account for 45% of ↑ in between-firm wage dispersion



Notes. The sample includes men 30–35 years old employed at private firms with at least ten employees.

Robustness: between-establishment wage inequality

Main

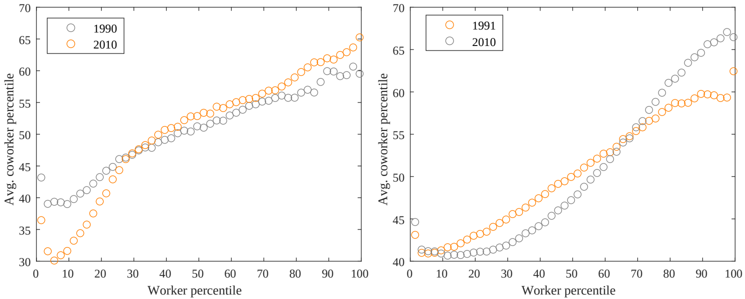


(a) With worker fixed effects **(b)** Without worker fixed effects

Notes. This figure depicts the share of the variance of log wages occurring between as opposed to within establishments for different measures of wages, as described in the main text. The relevant regressions are performed for the entire sample, then the decomposition is performed year-by-year.

Nonlinear coworker sorting

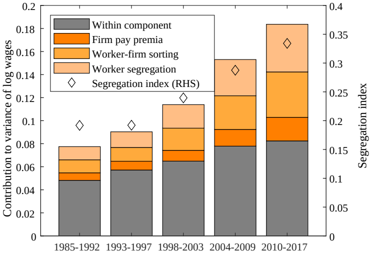
Overview



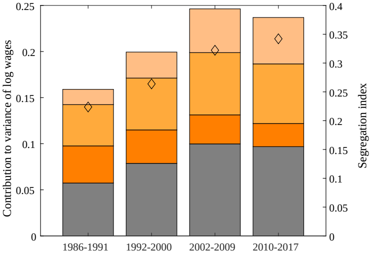
(c) Germany **(d)** Portugal

Notes. This figure plots, for any percentile of the worker fixed-effect distribution, the average quality of coworkers. To enhance visual clarity, workers are first binned into 50 cells and then the coworker quality is computed for each cell.

AKM and Song et al. (2019) decompositions



(a) Germany



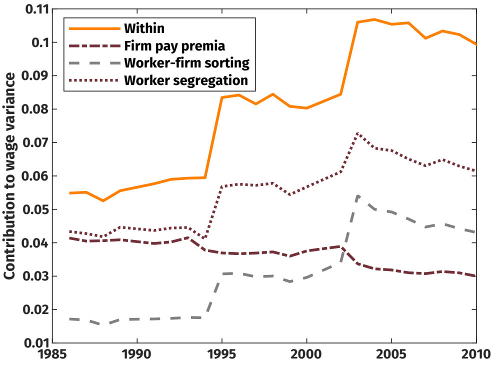
(b) Portugal

PRT: Bloom et al. (2019) style decomposition

Evidence overview
Method comparison

- The reason that in PRT the between-firm inequality component hasn't ↑ (significantly) is a compression in firm-specific pay premia

  [consistent with evidence in Silva, Leitao, and Montana (2022)]



PRT: definitions of layers

- Definitions underlying the PRT 'layer' classification
- Some classification changes in 2010, which seems to affect esp. higher-skilled professionals vs. supervisors/team leaders

Table A.1: Classification of Workers According to Hierarchical Levels

Level	Tasks	Skills
1. Top executives (top management)	Definition of the firm general policy or consulting on the organization of the firm; strategic planning; creation or adaptation of technical, scientific and administrative methods or processes	Knowledge of management and coordination of firms fundamental activities; knowledge of management and coordination of the fundamental activities in the field to which the individual is assigned and that requires the study and research of high responsibility and technical level problems
2. Intermediary executives (middle management)	Organization and adaptation of the guidelines established by the superiors and directly linked with the executive work	Technical and professional qualifications directed to executive, research, and management work
3. Supervisors, team leaders	Orientation of teams, as directed by the superiors, but requiring the knowledge of action processes	Complete professional qualification with a specialization
4. Higher-skilled professionals	Tasks requiring a high technical value and defined in general terms by the superiors	Complete professional qualification with a specialization adding to theoretical and applied knowledge
5. Skilled professionals	Complex or delicate tasks, usually not repetitive, and defined by the superiors	Complete professional qualification implying theoretical and applied knowledge
6. Semi-skilled professionals	Well defined tasks, mainly manual or mechanical (no intellectual work) with low complexity, usually routine and sometimes repetitive	Professional qualification in a limited field or practical and elementary professional knowledge
7. Non-skilled professionals	Simple tasks and totally determined	Practical knowledge and easily acquired in a short time
8. Apprentices, interns, trainees	Apprenticeship	

Notes: Hierarchical levels defined according to Decreto Lei 121/78 of July 2nd (Lima and Pereira, 2003).

Main result: non-parametric approximation

▸ Regression evidence

- Baseline measure suggest $\approx 2\times\uparrow$ in the magnitude of the coefficient
- Also report a numerical approximation to the avg. cross-partial derivative, $\widehat{\frac{\partial w^2(x,x')}{\partial x \partial x'}}$
  - closer to structural model

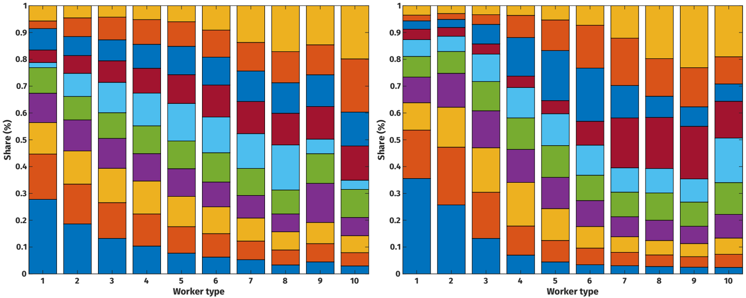
Sample period	$\hat{\beta}_3$	$\widehat{\frac{\partial w^2(x,x')}{\partial x \partial x'}}$
1985-1992	0.1351	0.0727
1993-1998	0.1499	0.060
1998-2003	0.2015	0.010
2004-2009	0.2535	0.148
2010-2017	0.2694	0.158

Notes. Reported in *hourly* wages for better comparison with PRT, assuming 7.5 working hours per day.
The results shown are based on ranking workers by their AKM-FE & non-parametric wage functions on slightly outdated specification.

DE: preliminary conditional coworker distributions

▸ Overview

- DE LIAB data, sample period 1 (1985-1994); worker skill for now based on AKM worker effect; individuals with top-coded wages (→) still in sample ⇐

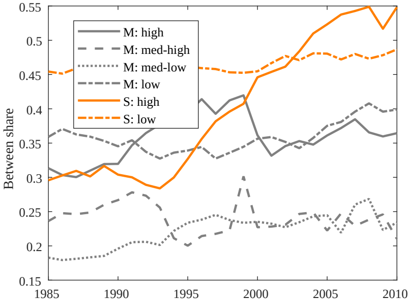


(a) 1985-1994 **(b)** 2003-2010

DE: wage decomposition across sectors

Overview FUI-style decomp.

- Eurostat classification of industries into 6 groups Eurostat def.
- Similar results when running AKM estimation & decomposition for each industry group, where between-components comprise: firm-specific pay premia, worker-firm sorting and worker-worker sorting
- **Driver: knowledge-intensive services** ("S: high")

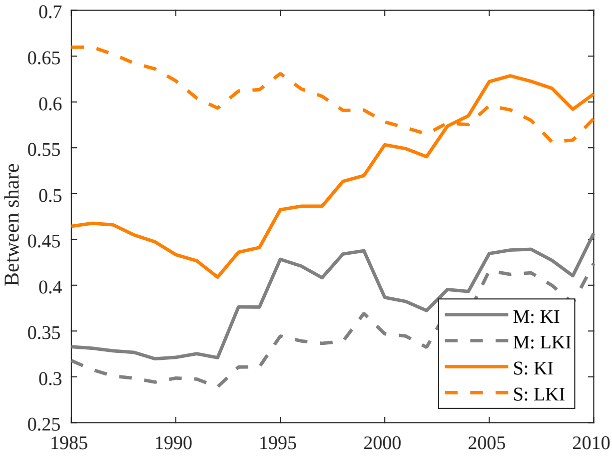


Notes. DE, full sample, at least 10 employees, yearly decomposition.

DE: AKM-decomposition based

Main

- FUI-style decomposition after estimating AKM model using pre-clustering a la Bonhomme et al. (2019)



Notes. Based on Eurostat classification

Interpretability: grouping industries

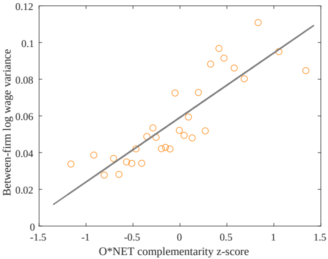
▸ Main

- **Objective**: meaningful interpretation of cross-industry results in terms of knowledge intensity (KI)
- Classify 2-digit industries into 4-6 groups following classification by Eurostat
  - based on NACE rev. 2 so feasible for both DE and PRT, but for latter only from 1994
  - manufacturing classified by level of R&D intensity; services based on share of tertiary educated persons
- Examples:
  - KI manufacturing: pharmaceuticals; computer, electronic and optical products
  - LKI manufacturing: food products; textile
  - KI services: professional services like legal; research and development
  - LKI services: accommodation and food services; real estate activities
- Alternative: continuous score (e.g. Bahar, 2019)

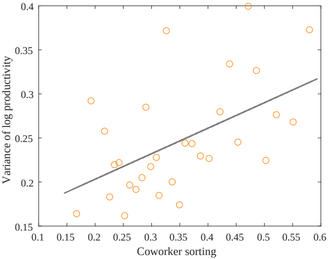
Result 2 (continued)

▸ Main

$\Rightarrow$ Industries with production processes featuring relatively high skill complementarity also tend to be characterized by greater coworker sorting and between-firm inequality in productivity & pay.



(a) Between-firm wage inequality



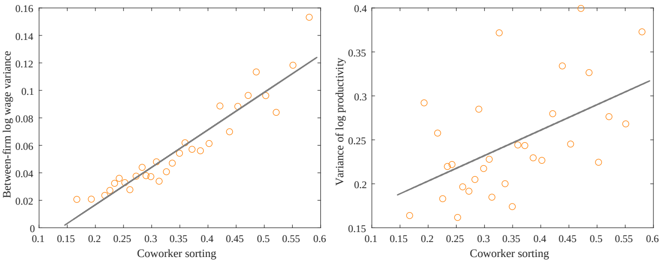
(b) Productivity dispersion

Notes. From joint analysis with Criscuolo & Gal.
The x-axis measures an industry-level proxy for complementarity at the 4-digit level. Employment-weighted mean score of an occupation-level index constructed from O*NET data measuring the importance of: teamwork, impact on coworker output, communication, and contact *[Bombardini et al., 2012]*.

Industry level analysis: coworker sorting ⇒ between-firm inequality

▸ Main

⇒ Measures of between-firm inequality in productivity and pay are increasing in the degree of coworker sorting.



(a) Between-share of wage inequality **(b)** Productivity dispersion

Notes. From joint analysis with Criscuolo & Gal.

DE: complexity $\Rightarrow$ complementarities (occ. groups)

Main

- The magnitude of the regression coefficient on the interaction term is also increasing in the degree of task complexity across 4 broad occupation groups in the LIAB data

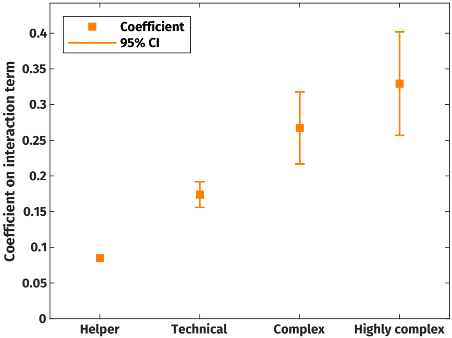


Figure 2: Complementarities across qualification groups

Notes. Workers are grouped by KldB-2010 5-digit occupation group.

Quantification

Between-share adjustment procedure (1)

▸ Main

- **Problem.** For any degree of coworker sorting less than unity, i.e. $\rho_{xx} < 1$, the level of the between-share in a model with team size $m = 2$ will be biased upward relative to the case of $m > 2$ and, in particular, $m \to \infty$ (LLN...)
  - implication 1: upward bias in level
  - implication 2: downward bias in estimated $\uparrow$ between-share as sorting increases

- **Propose** statistical adjustment method

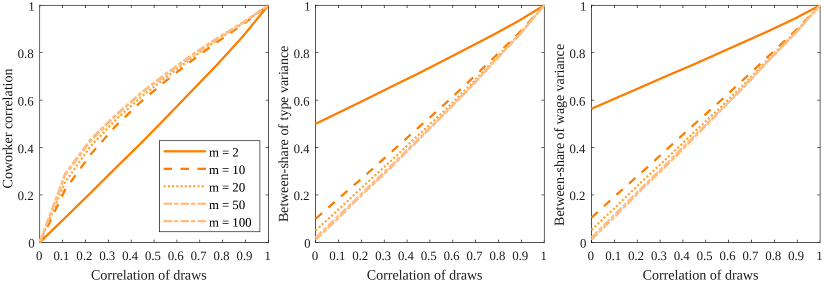
- Consider a random vector $X = (X_1, X_2, \ldots, X_m)'$ whose distribution is described by a Gaussian copula over the unit hypercube $[0, 1]^m$, with an $m \times m$ dimensional correlation matrix $\Sigma(\rho^c)$, which contains ones on the diagonal and the off-diagonal elements are all equal to ρ_c

- Interpretation. Each vector of observations drawn from the distributions of X, $x_j = (x_{1j}, x_{2j}, \ldots, x_{mj})'$, describes the types of workers in team of size m, indexed by j

Between-share adjustment procedure (2)

- Since the marginals of the Gaussian copula are simply continuous uniforms defined over the unit interval, the variance of the union of all draws is just $\frac{1}{12}$
- The mean of the elements of X is itself a random variable, $\bar{X}$. That is, for some realization x_j, we can define $\bar{x}_j = \frac{1}{m}\sum_{i=1}^{m} x_{ij}$
- The variance of $\bar{X}$ will be $\frac{1}{m^2}\left(\frac{m}{12} + m(m-1)(\frac{\rho_c}{12})\right)$
- So $\sigma^2_{x,\text{between-share}}(\rho_c, m) = \frac{1}{m}\left(1 + (m-1)\rho_c\right)$
- Correction-factor $= \frac{1}{2}\left(1 + \rho_c\right) - \frac{1}{\hat{m}}\left(1 + (\hat{m}-1)\rho_c\right)$ where the empirical average size is $\hat{m}$

Between-share adjustment procedure (3)



(a) Coworker sorting **(b)** Between-share of type var. **(c)** Between-share of wage var.

Non-parametric approximation of the cross-partial derivative

- Recall structural relationship b/w *production* & *wage* complementarities:

$$\omega \frac{\partial^2 f(x,y,x')}{\partial x \partial x'} = \frac{\partial^2 w(x,(y,x'))}{\partial x \partial x'}$$

$\Rightarrow$ **Objective:** empirical estimate of $\widehat{\frac{\partial w^2(x,x')}{\partial x \partial x'}}$

- **Sketch of method:** for sample period t
  1. given global ranking of workers, bin similarly ranked workers and coworkers (n_x quantiles)
  2. $n_x \times n_x$ matrix with typical element w_{kl}: avg. wage of a worker in quantile k of the type distribution whose average coworker is in quantile l of the coworker type distribution, with $k = 1, ..., n_x$ and $l = 1, \ldots, n_x$
  3. fit cubic polynomial for each worker type $x \Rightarrow \tilde{w}_{k,l}$.
  4. FD approx. of cross-partial derivative at each grid point $\Rightarrow \left[\widehat{\frac{\partial w^2(x,x')}{\partial x \partial x'}}\right]_{kl}$
  5. compute average, weighting by observed match density

$$\frac{\partial \widehat{w^2(x,x')}}{\partial x \partial x'} = \sum_{k=1}^{n_x} \sum_{l=1}^{n_x} \frac{d_m(x_k, x_l)}{\sum_{k=1}^{n_x} \sum_{l=1}^{n_x} d_m(x_k, x_l)} \times \left[\frac{\partial \widehat{w^2(x,x')}}{\partial x \partial x'}\right]_{kl}$$

Validation: average coworker in data vs. model

▸ Main

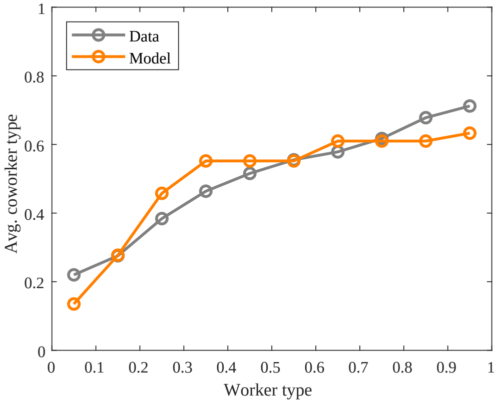
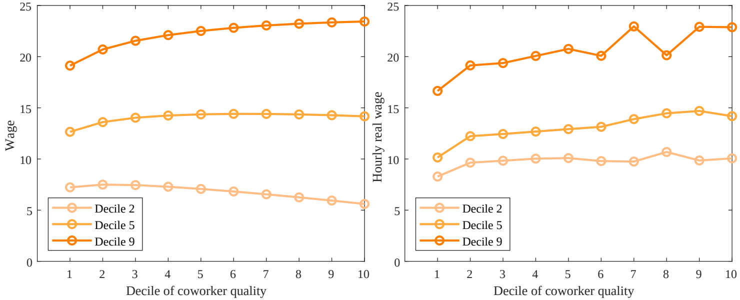


Figure 3: Average coworker: data vs. model (untargeted)

Notes. This figure plots, for each worker type, the average coworker type, in both model and data. Types are binned to deciles. All model-implied moments are constructed based on the subset of workers employed in teams.

Validation: wage functions

Main



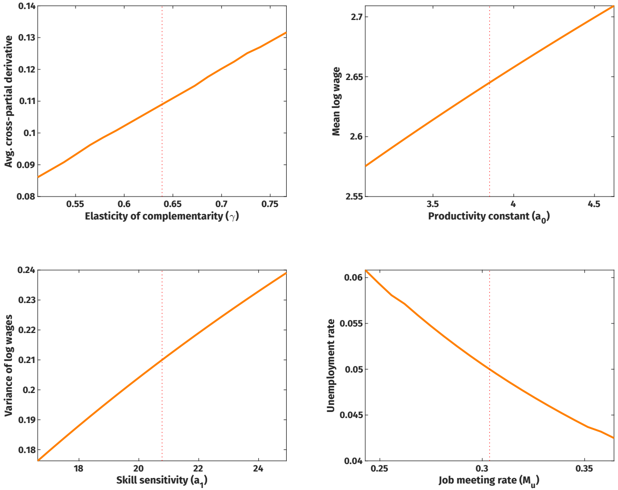
(a) Model **(b)** Data (2010)

Figure 4: Non-parametric wage functions: model vs. data (untargeted)

Notes. The left panel plots the model-implied wage function, for three different values of x (represented by different lines), against x' on the horizontal axis. The right panel depicts the non-parametric wage function estimated in the data. All model-implied moments are constructed based on the subset of workers employed in teams.

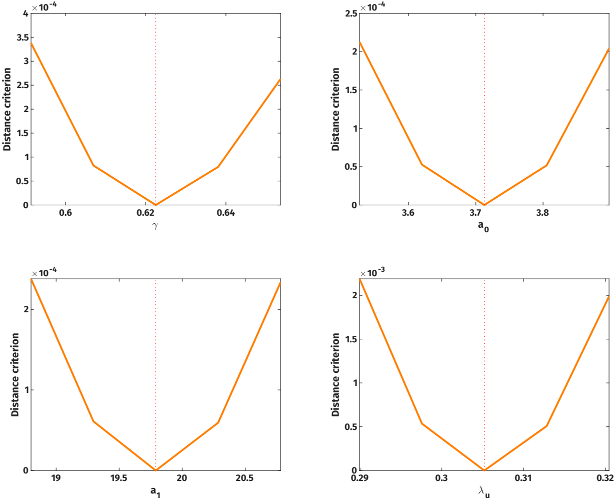
Identification validation exercise 1

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 $\mathcal{G}$.