Superstar Teams

Lukas Freund

University of Cambridge

May 24, 2024

Bristol Macroeconomics Workshop

Motivation: opening the "black box" of firms in macro

- **Most production processes involve groups of heterogeneous workers**
  - too complex for any individual to perform all tasks well
    $\rightarrow$ **firms coordinate** groups of **workers with specialized skills**
  - individuals also vary in their *overall* productivity ("**talent**")

Motivation: opening the "black box" of firms in macro

- **Most production processes involve groups of heterogeneous workers**
  - too complex for any individual to perform all tasks well
    $\rightarrow$ **firms coordinate** groups of **workers with specialized skills**
  - individuals also vary in their *overall* productivity ("**talent**")

- **Q:** Who should work with whom? Market allocation? Implications?
  - classic themes in literature *[Becker-Murphy, 1992; Kremer, 1993; Garicano, 2000]*
  - but theoretically & micro oriented

Motivation: opening the "black box" of firms in macro

- **Most production processes involve groups of heterogeneous workers**
  - too complex for any individual to perform all tasks well
    $\rightarrow$ **firms coordinate** groups of **workers with specialized skills**
  - individuals also vary in their *overall* productivity ("**talent**")
- **Q:** Who should work with whom? Market allocation? Implications?
  - classic themes in literature *[Becker-Murphy, 1992; Kremer, 1993; Garicano, 2000]*
  - but theoretically & micro oriented
- **This paper:** tractable model for empirical & quantitative analysis that highlights *macro* implications: agg. productivity, trends in labor market inequality

  (1) theory (2) measurement (3) application(s)

Modeling firms as team assemblies & core mechanism

- **What's a firm:** an organized collection of workers with task-specific skills (team)
- **Key idea:** specialization $\Rightarrow$ coworker talent complementarities $\Rightarrow$ sorting
- **Implications:**
  - $\uparrow$ specialization $\Rightarrow$ talent sorting $\uparrow$
  - input/labor market frictions lower output b/c teams are formed inefficiently

The paper in a nutshell

- **Develop tractable theory of the firm centered on team-production & formation**
  1. **task-based** production
  2. **skill heterogeneity:** task-specificity & variable talent
  3. **team production**
  4. assembling a team takes time due to **search** frictions

The paper in a nutshell

- **Develop tractable theory of the firm centered on team-production & formation**
  1. **task-based** production
  2. **skill heterogeneity:** task-specificity & variable talent
  3. **team production**
  4. assembling a team takes time due to **search** frictions

microfound tractable firm-level production fn. with *endogenous* complementarities

The paper in a nutshell

- **Develop tractable theory of the firm centered on team-production & formation**
  1. **task-based** production
  2. **skill heterogeneity:** task-specificity & variable talent
  3. **team production**
  4. assembling a team takes time due to **search** frictions

  (1–3) **microfound tractable firm-level production fn. with *endogenous* complementarities**

- **Confront theory with data:** structurally identify production fn. with micro panel data on wages+matches $\rightarrow$ estimate & test model

The paper in a nutshell

- **Develop tractable theory of the firm centered on team-production & formation**
  1. **task-based** production
  2. **skill heterogeneity:** task-specificity & variable talent
  3. **team production**
  4. assembling a team takes time due to **search** frictions

  microfound tractable firm-level production fn. with *endogenous* complementarities

- **Confront theory with data:** structurally identify production fn. with micro panel data on wages+matches → estimate & test model

- **Main quantitative application:** deepening specialization explains ≈ 25% of ↑ between-firm wage inequality share in DE since 1985 *[cf. Card et al., 2013; Bloom et al., 2019]*

The paper in a nutshell

- **Develop tractable theory of the firm centered on team-production & formation**
  1. **task-based** production
  2. **skill heterogeneity:** task-specificity & variable talent
  3. **team production**
  4. assembling a team takes time due to **search** frictions

  $\}$ **microfound tractable firm-level production fn. with *endogenous* complementarities** (items 1–3)

- **Confront theory with data:** structurally identify production fn. with micro panel data on wages+matches $\rightarrow$ estimate & test model

- **Main quantitative application:** deepening specialization explains $\approx$ 25% of $\uparrow$ between-firm wage inequality share in DE since 1985 *[cf. Card et al., 2013; Bloom et al., 2019]*
  - theory has implications for agg. productivity, productivity dispersion, job ladders, ...

Relation & contributions to literature

- **Firm organization: parsimonious team-model for quantitative applications**

  Firms: *Lucas, 1978; Rosen, 1982; Becker & Murphy, 1992; Hopenhayn & Rogerson, 1993;* ***Kremer, 1993****; Kremer & Maskin, 1996;* ***Garicano, 2000****; Klette & Kortum, 2004;* ***Garicano & Rossi-Hansberg, 2006****; Kohlhepp, 2022; Kuhn et al., 2022; Minni, 2022; Bassi et al., 2023*

  Task assignment: *Costinot & Vogel, 2010;* ***Acemoglu & Restrepo, 2018****; Ocampo, 2021; Adenbaum, 2022*

  Teams: *Akcigit et al., 2018; Chade & Eeckhout, 2020;* ***Jarosch et al., 2021****; Herkenhoff et al., 2022*

- **Frictional labor market sorting: endogenize & measure complementarities**

  Shimer & Smith, 2000; Cahuc et al., 2006; Eeckhout & Kircher, 2011/2018; Hagedorn et al., 2017; de Melo, 2018; ***Herkenhoff et al., 2022****; Lindenlaub & Postel-Vinay, 2023*

- **Wage inequality: structural model of ↑ firm-level inequality due to technological △**

  Technology: *Katz & Murphy, 1992; Krusell et al., 2000; Autor, Levy & Murnane, 2003;* ***Jones, 2009****; Deming, 2017; Acemoglu & Restrepo, 2018; Alon, 2018; Neffke, 2019; Jones, 2021; Atalay et al., 2021*

  Firms: ***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; Kleinman, 2022*

Theory

Overview of environment

- **Ex-ante homogeneous firms** assemble teams of **heterogeneous workers** who vary in task specialization (*what* you're skilled at) & talent (*how* skilled you are)
  - task-based production
  - hire workers in search-frictional labor market

- **Challenge:** how to keep this tractable?

- **Analyze economy-wide organization of production in 2 steps:**
  1. team production: how much does a team produce if tasks are optimally allocated?
  2. team formation: who is matched to work with whom?

Step 1: production in a single team of given composition

- Firm: **1 team of $n \in \mathbb{Z}_{++}$ workers** produces output from **unit continuum of tasks $\mathcal{T}$**

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

- **Task aggregation:**

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

- **Workers $\rightarrow$ tasks:** i supplies 1 time unit, task-specific productivities $\mathbf{z}_i = \{z_i(\tau)\}_{\tau \in \mathcal{T}}$

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

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

Firm's optimization problem

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

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

shadow cost of τ

opportunity cost of i's time

productivity of i for τ

Tractability: leverage insight from trade literature

Assumption: Multivariate Fréchet distribution of worker-task skills

$$\Pr\left[Z_1(\tau) \leq z_1, \ldots, Z_n(\tau) \leq z_n\right] = \exp\left[-\left(\sum_{i=1}^{n}\left(\left(\frac{z_i}{\iota x_i}\right)^{-\frac{1}{\chi}}\right)^{\frac{1}{\xi}}\right)^{\xi}\right]$$

- **Payoff #1:** low-dim. representation of worker-task skill dist.
  - x_i: scale term $\sim$ **"talent" type**
  - χ: *within*-worker skill dispersion $\sim$ **specialization**
  - $\xi \in (0, 1]$: correlation in z's *across* workers $\sim$ are team members specialized in the same or different tasks

Illustration: talent types

- **"Talent" type** x_i: *how* skilled is *i* conditional on whatever task they're specialized in (absolute advantage)

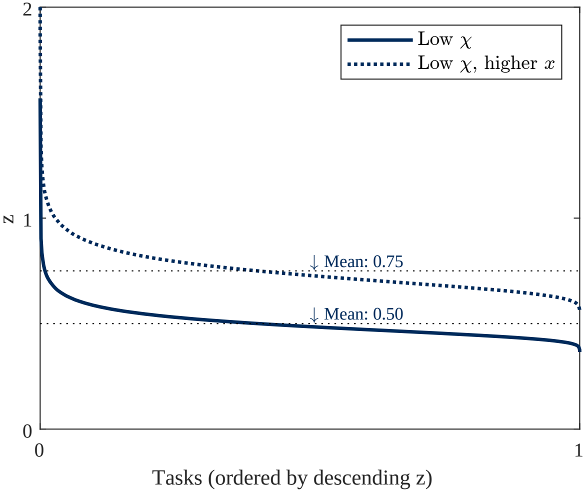
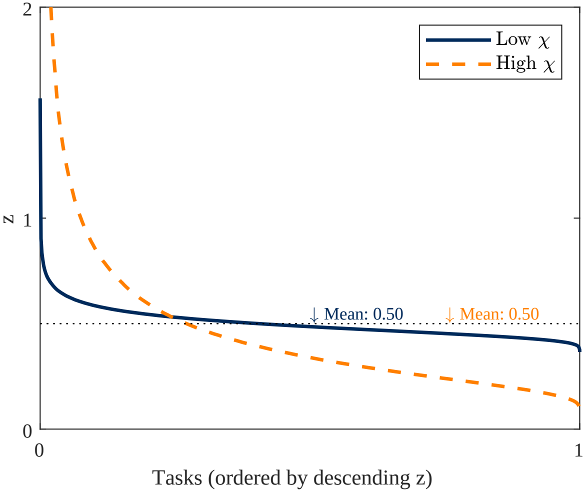


Illustration: specialization parameter

▸ Illustration: correlation

- **Parameter χ:** *within*-worker skill dispersion across tasks $\sim$ **specialization**



Tractability: leverage insight from trade literature

Assumption: Multivariate Fréchet distribution of worker-task skills

$$\Pr\left[Z_1(\tau) \leq z_1, \ldots, Z_n(\tau) \leq z_n\right] = \exp\left[-\left(\sum_{i=1}^{n}\left(\left(\frac{z_i}{\iota x_i}\right)^{-\frac{1}{\chi}}\right)^{\frac{1}{\xi}}\right)^{\xi}\right]$$

- Payoff #1: low-dim. representation of worker-task skill dist.
- **Payoff #2:** max-stability property allows closed-form characterization of key objects

Micro-founded production function

Lemma

Proposition: Reduced-form production function

Talent types $(x_1, \ldots, x_n)$, χ and ξ are sufficient statistic for team output Y:

$$Y = f(x_1, \cdots, x_n; \chi, \xi)$$

- Proof sketch

  Details

  - solving firm's mini-planner problem implies task-demand and assignment of tasks by comparative advantage with complete division of labor
  - Fréchet distribution allows closed-form characterization of distribution of shadow costs $\lambda(\tau)$ conditional on this assignment
  - integrate over tasks and workers, normalizing $\lambda = 1$

Micro-founded production function: characterization

Proposition: Reduced-form production function

$$f(x_1, \cdots, x_n; \chi, \xi) = n^{1+\chi\xi} \times \left(\frac{1}{n} \sum_{i=1}^{n} (x_i)^{\frac{1}{1+\chi\xi}} \right)^{1+\chi\xi}$$

vs. no-division-of-labor: $f(x_1, ..., x_n) = n \times (\frac{1}{n} \sum_i^n x_i)$

1. **Output depends on team composition** – f is not additively separable

The whole is greater than the sum of its parts...

Proposition: Reduced-form production function

$$f(x_1, \cdots, x_n; \chi, \xi) = \underbrace{n^{1+\chi\xi}}_{\text{efficiency gains}} \times \left(\frac{1}{n} \sum_{i=1}^{n} (x_i)^{\frac{1}{1+\chi\xi}} \right)^{1+\chi\xi}$$

2. **Efficiency gains from teamwork – large when workers have skills that are *specialized* (χ) in *different* (ξ) tasks**
   - no-division-of-labor: $f(\cdot) = n \times (\frac{1}{n} \sum_i^n x_i)$
   - intuition: gains from tasks being assigned by comparative advantage

▸ task assignment

...but output is lowered by talent dispersion

Proposition: Reduced-form production function

$$f(x_1, \cdots, x_n; \chi, \xi) = \underbrace{n^{1+\chi\xi}}_{\text{efficiency gains}} \times \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i)^{\frac{1}{1+\chi\xi}}\right)^{1+\chi\xi}}_{\text{talent complementarity}},$$

3. **Coworker talent complementarities – strong when workers have skills that are *specialized* (χ) in *different* (ξ) tasks**
   - reduced-form elasticity of complementarity $\gamma := \frac{\partial \ln(f_j/f_i)}{\partial \ln(x_i/x_j)} = \frac{\chi\xi}{1+\chi\xi}$
   - greater dispersion in team members' talent reduces output, other things equal
   - $\partial f(\cdot)/\partial x_i \partial x_{-i} > 0$

...but output is lowered by talent dispersion

Proposition: Reduced-form production function

$$f(x_1, \cdots, x_n; \chi, \xi) = \underbrace{n^{1+\chi\xi}}_{\text{efficiency gains}} \times \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i)^{\frac{1}{1+\chi\xi}}\right)^{1+\chi\xi}}_{\text{talent complementarity}},$$

3. **Coworker talent complementarities are strong when workers have skills that are *specialized* (χ) in *different* (ξ) tasks**
   - intuition: greater specialization means higher-x team members can't easily perform more tasks as their skill rapidly diminishes

task assignment

Key takeaways

1. $\uparrow$ **Skill specialization** *endogenously* generates $\uparrow$ **coworker talent complementarities**

Firm organization meets frictional matching

- So team composition matters for productivity, especially when χ is high
- **How is team composition determined?**
- Exploit tractability of team production fn. f to **integrate team production with dynamic eqm. matching**
    - team composition determined by tradeoff between complementarities and search costs
    - preview: dynamic setting will also be key for identification of f with panel data

Frictional matching into teams: environment

- **Random-search + multi-worker firms** *[Herkenhoff-Lise-Menzio-Phillips, 2024]*
    - teams instead of one-worker firms *[e.g., Shimer-Smith, 2000; Hagedorn-Law-Manovskii, 2017]*
- **Production technology:** $f(\cdot)$ with $n \in \{0, 1, 2\}$
- **Employment states:** unemp., employed alone, employed with $x' \in \mathcal{X}$
- **Nash wage bargaining** with continuous renegotiation
- ξ: **match-specific shock** observed by firms & workers before match decision
    - microfoundation: coworkers' task specialization matters for f, not only x
    - ξ is endogenously determined

Matching – stationary equilibrium

Details

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

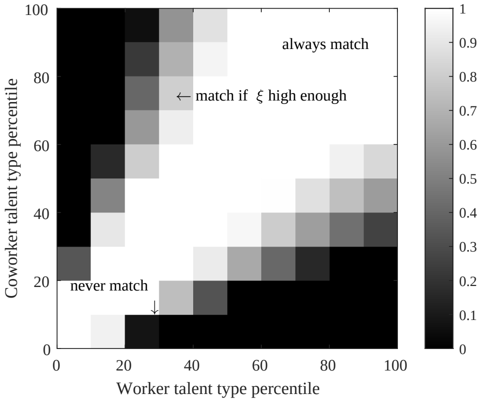
Definition: Stationary equilibrium

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

1. the value functions satisfy the HJB equations given the distribution;
2. the distribution is stationary given the policy fn's implied by the value fn's.

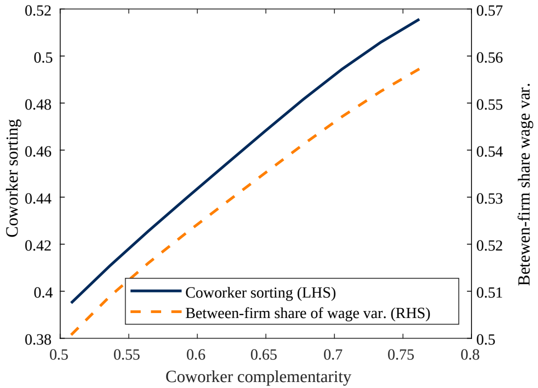
Mechanism: conditional matching probabilities for given χ

- What is the probability that a firm with a type-x worker would hire a type-x' worker conditional on meeting?



Mechanism: more positive assortative matching as $\chi \uparrow$

- What is the correlation between talent x & coworker talent x', ρ_{xx}?



Key takeaways

1. $\uparrow$ Skill specialization endogenously generates $\uparrow$ coworker talent complementarities

2. $\uparrow$ **Talent complementarities** lead to $\uparrow$ **positive assortative matching**

Model Meets Data

Taking the model to (micro) data

- **Goal:** test core mechanisms & analyze quantitative properties of the model

- **Steps:**
  1. matched employer-employee panel data
  2. map model objects to data
  3. estimate & quantitatively analyze model
  4. validate core model mechanisms *(brief today)*

Panel micro data for Germany

▸ SIEED

- **Primary data**: **SIEED matched-employer employee panel for W Germany**
  - 1.5% sample of establishments + entire biographies of associated workers; social security information on employer, daily wage, occupation, demographics
  - initially focus on 2010-2017, later extend to 1985-2017
  - selection: f-t employees aged 20-60, $\geq$ 10 observations per establishment-year
- **What do we need to know to quantitatively analyze the model?**
  - some parameters – e.g. labor market separation and meeting rates – can be $\sim$ directly read off micro data or are informed by literature (e.g., discount rate)
  - x and x' can be recovered using applied micro methods
  - key challenge: χ

Mapping theory to data: worker & coworker types

- **Theory:** wage monotonically $\uparrow$ in x, so can measure using panel dimension
- Implementation: fixed effect (FE) wage regressions
  - AKM model *[Abowd et al., 1999]*: job switching identifies time-invariant "worker types," controlling for possibility that some employers pay more to all workers
    - pre-est. k-means clustering to address limited mobility bias *[Bonhomme et al., 2019]*
  - AKM misspecified according to structural model, so also use a non-param. ranking algo *[Hagedorn et al., 2017]* that's consistent $\rightarrow$ similar worker types

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

Mapping theory to data: worker & coworker types

- **Theory:** wage monotonically $\uparrow$ in x, so can measure using panel dimension
- Implementation: fixed effect (FE) wage regressions
  - AKM model *[Abowd et al., 1999]*: job switching identifies time-invariant "worker types," controlling for possibility that some employers pay more to all workers
    - pre-est. k-means clustering to address limited mobility bias *[Bonhomme et al., 2019]*
  - AKM misspecified according to structural model, so also use a non-param. ranking algo *[Hagedorn et al., 2017]* that's consistent $\rightarrow$ similar worker types

$\Rightarrow$ **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
- Then can compute, e.g., $\hat{\rho}_{xx} = \text{corr}\left(\hat{x}_i, \hat{x}_{-it}\right)$

Measurement: a useful identification result

- Within-worker skill dispersion χ isn't easily observable
- **Theory:** Proposition 1 ties χ to $\frac{\partial^2 f(\cdot)}{\partial x \partial x'}$, which given x & x' we can identify from $w(x|x')$

Proposition: Measuring complementarities

Coworker complementarities (CC) in production are proportional to CC in wages:

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

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

can measure this

Proof sketch

Key takeaways

1. $\uparrow$ Skill specialization endogenously generates $\uparrow$ coworker talent complementarities
2. $\uparrow$ Talent complementarities lead to $\uparrow$ + assortative coworker matching
3. **Complementarities can be identified from panel micro data on wages and matches**

Mapping theory to data: coworker complementarity

- Want to approximate $\frac{\partial^2 \bar{w}(x|x')}{\partial x \partial x'}$
- **Estimating polynomial equation:**

$$\frac{w_{it}}{\bar{w}_t} = \beta_0 + \beta_1 \hat{x}_i + \beta_{11} \hat{x}_i^2 + + \beta_2 \hat{x}_{-it} + \beta_{22} \hat{x}^2_{-it} + \beta_c \left(\hat{x}_i \times \hat{x}_{-it} \right)$$
$$+ \psi_{j(it)} + \nu_{o(i)t} + \xi_{s(i)t} + \epsilon_{it}$$

  - alternative: non-parametric approximation

Evidence on coworker complementarity (2010-2017)

▸ Robustness ▸ B-o-E calc. γ ▸ Peer effects

- **Estimating equation:**

$$\frac{w_{it}}{\bar{w}_t} = [\ldots] + \overbrace{\beta_c}^{\text{coworker complementarity}} (\hat{x}_i \times \hat{x}_{-it}) + [\ldots]$$

	$\hat{\beta}_c$	Non-parametric FD method
Coworker complementarity	**0.0058***	0.0075
Obs. (1000s)	4,410	4,410

Notes. Regressions include FEs for employer; occupation-year; industry-year. Employer-clustered standard errors in parentheses. Observations weighted by the inverse employment share of the respective type and (rounded) coworker type cell. FD: finite differences.

Robustness checks: measuring 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

Quantifying the model: approach

▸ Parameter values ▸ Identification validation ▸ Within-industry

- **Calibrate** the model to the W German economy (2010-2017)
  1. externally calibrated: discount rate, $\bar{n}$, bargaining power
  2. offline estimation: job separation hazard
  3. **online estimation** (indirect inference): meeting rate, unemp. flow benefit, production
     - targets: $\hat{\beta}_c$, total wage variance, avg. wage level, replacement rate, job finding rate
- **Production function disciplined by $\hat{\beta}_c$**
  - set χ to align regression estimates for this reduced-form moment using model-generated data with empirical estimate

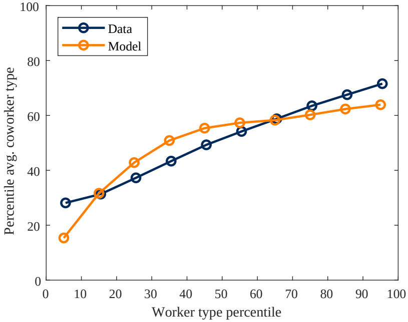
▸ Monte-Carlo

- **Macro moments of interest are untargeted**: sorting, between-firm wage inequality

Calibration results & model validation

▸ Between-share adjustment method

- ✓**Match coworker sorting patterns**
  - $\rho_{xx} = 0.43$ (vs. 0.62 in data)
- ✓**Match between-firm wage inequality**
  - between-share 0.48 (vs. 0.57 in data)
- Extensive **validation** of core mechanisms



Validation of core model mechanisms

▸ Details

- Non-wage implications of complementarities: **direction of EE moves** ▸ Details
  - Δ coworker talent positively correlated with own talent
- Cross-sectional analysis: **variation across occupations/industries** ▸ Details
  - task-based proxy for $\chi \uparrow \Rightarrow$ coworker talent complementarity $\uparrow$
  - coworker talent complementarity $\uparrow \Rightarrow$ coworker talent sorting $\uparrow$
- Work-in-progress analyses
  - $\xi \uparrow \Rightarrow$ coworker talent complementarity $\uparrow$
  - $\chi \uparrow \Rightarrow$ wage dispersion conditional on $(x, x') \uparrow$
  - wage effects induced by separation from a worker $\leftrightarrow$ Jaeger-Heining (2022)

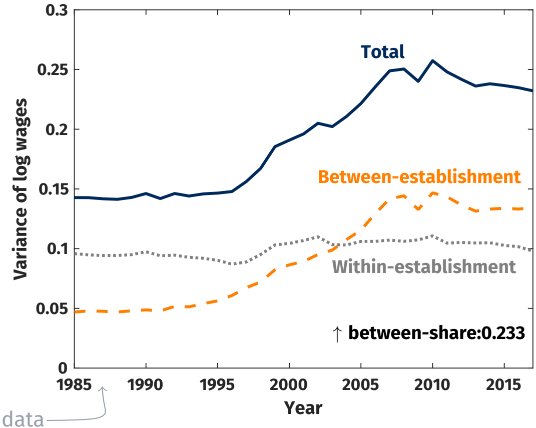
Firming Up Inequality

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]

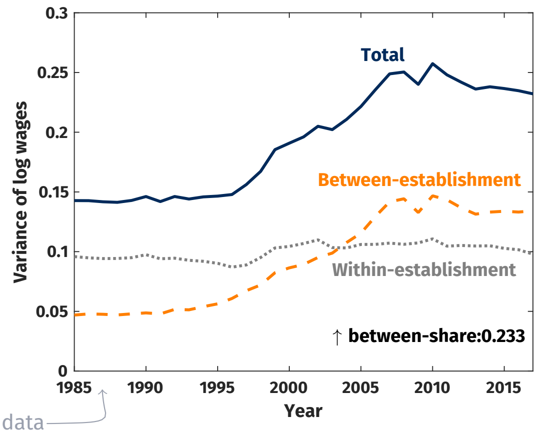


German matched employer-employee data

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

Details

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



German matched employer-employee data

Hypothesis

- **Hypothesis:**
  1. the set of tasks any one worker can perform very well has narrowed: specialization $\uparrow$
  2. coworker talent complementarities $\uparrow$
  3. individuals of similar talent increasingly work together
  4. this generates greater between-firm wage dispersion

Hypothesis

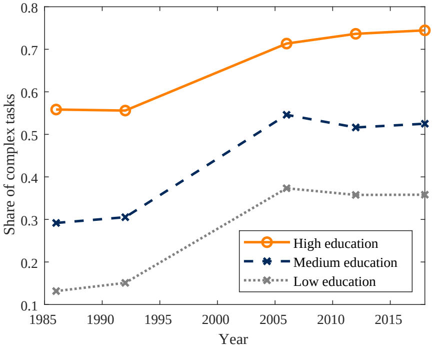
- **Hypothesis:**
  1. the set of tasks any one worker can perform very well has narrowed: specialization $\uparrow$
  2. coworker talent complementarities $\uparrow$
  3. individuals of similar talent increasingly work together
  4. this generates greater between-firm wage dispersion

- **Approach:**
  - evaluate descriptive evidence
  - use structural model for counterfactual exercises quantifying this mechanism

Qualitative evidence suggests specialization (χ) ↑

▸ Occ. movements ▸ Science

- **Task complexity ↑**: "extensive margin" of χ
  - DE longitudinal task survey ▸ BIBB
  - "complex": cognitive non-routine (e.g., organizing, researching)
- Literature points in similar direction
  - Grigsby ('23): transferability ↓ among occ w high/social/manual skills
  - Jones ('09): burden of knowledge
  - Jovanovic-Rousseau ('08): proliferation of occupation codes

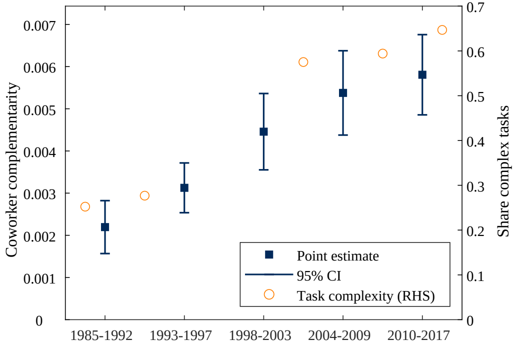


Coworker talent complementarity has strengthened

▸ Schooling ▸ Peer effect trends

- **Theory:** specialization (χ) $\uparrow$ is associated with coworker complementarity $\uparrow$

✓**Coworker complementarity has more than doubled between 1985-1992 and 2010-2017**

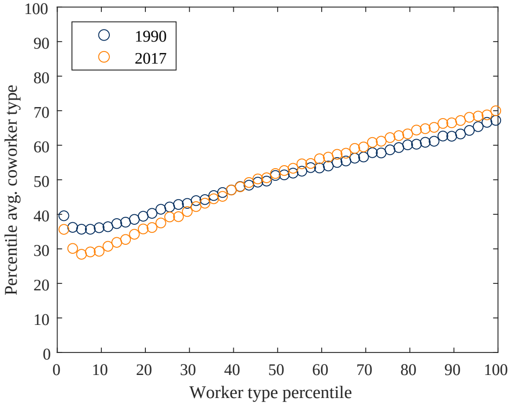


Talent sorting has intensified

▸ Details

- **Theory:** complementarity ↑ is associated with talent sorting ↑

✓ **Coworker matching has become more positively assortative**

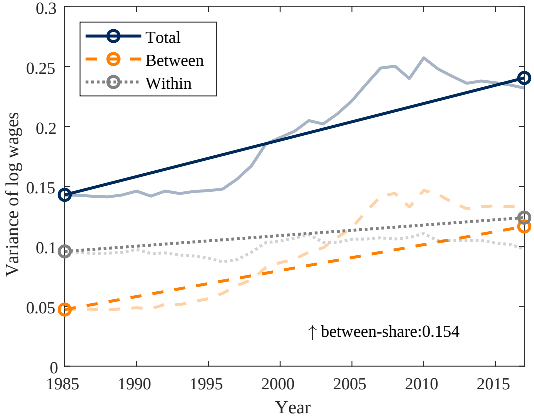


Key takeaways

1. $\uparrow$ Skill specialization endogenously generates $\uparrow$ coworker talent complementarity
2. $\uparrow$ Coworker complementarity leads to $\uparrow$ positive assortative coworker sorting
3. Complementarities can be disciplined using panel micro data on wages and matches
4. **Coworker complementarity has doubled** since 1985 & **talent sorting has intensified**

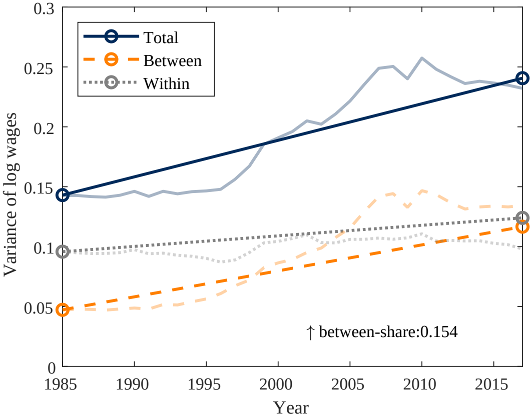
Model matches *changes* in firm-level wage distribution

- How quantitatively important are these developments in accounting for the firming up of inequality?
- **Re-calibrate model**: ’85-’92
- ✓**Model replicates untargeted rise of between-share in data**
    - 2/3 of ↑ between-share in data, (’85-’92)→(’10-’17)



Model matches changes in firm-level wage distribution – why?

- ✓Model replicates untargeted rise of between-share in data
- **Reflects several parameters changing**
  - χ: XX in '85-'92 (vs. XX in '10-'17)
  - job arrival & separation ↑
  - ...



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

- **Q:** How much of $\uparrow$ between-firm share of wage var. is due to $\uparrow$ complementarities?

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

- **Q:** How much of $\uparrow$ between-firm share of wage var. is due to $\uparrow$ complementarities?
- **Counterfactual:** between-firm share in 2010s absent $\chi \uparrow$ since '85-'92

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

- **Q:** How much of $\uparrow$ between-firm share of wage var. is due to $\uparrow$ complementarities?
- **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**
- **Robustness** exercises: 25-40%

	Δ model	Implied % Δ model due to Δ parameter
Model	0.154	-
Cf.: fix period-1 complementarity	0.065	58

Overview of model robustness checks

- Declining search frictions ▸ Jump
- Within-industry calibration ▸ Jump
- Economy-wide vs. within-occupation analysis ▸ Jump
- OJS ▸ Jump
- Increased talent dispersion ▸ Jump

Key takeaways

1. $\uparrow$ Skill specialization endogenously generates $\uparrow$ coworker talent complementarity
2. $\uparrow$ Coworker complementarity leads to $\uparrow$ positive assortative coworker sorting
3. Complementarities can be disciplined using panel micro data on wages and matches
4. Coworker complementarity has doubled since 1985 & talent sorting has intensified
5. **This explains a substantial share of $\uparrow$ between-firm wage inequality share**

Other applications

Overview of extensions & other implications

- **Agg. productivity**: complementarities + search frictions $\Rightarrow$ mismatch costs
  Jump
- **Productivity dispersion**: talent sorting $\Rightarrow$ firm-level productivity gaps Jump
- **"Coworker job ladders"**: workers, especially talented ones, tend to switch toward jobs with better coworkers Jump
- **Person-level inequality**: segregation by itself *need not* cause greater person-level inequality, but it may in interaction with, e.g., product market frictions or fairness norms Jump

Conclusion

Conclusion: firms form & organize teams – matters for macro

- **Main idea:** if individuals have specialized skills, firms assemble teams of 'complementary' coworkers, generating systematic sorting patterns
- **This paper:**
  1. **task-based microfoundation for firm-level production fn. with skill complementarities**

     $\Rightarrow$ specialization + team production $\rightarrow$ complementarities
  2. **measurement** combining reduced-form micro evidence with model structure

     $\Rightarrow$ between-firm differences w/o assuming ex-ante heterogeneity
  3. structural explanation for the **"firming up" of inequality**

     $\Rightarrow$ role of **increased complementarities**

Thank You!

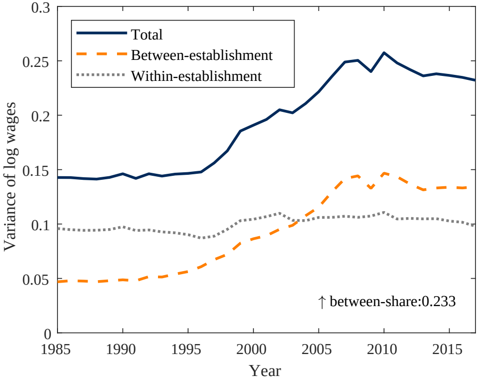
Extra Slides

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

▸ 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

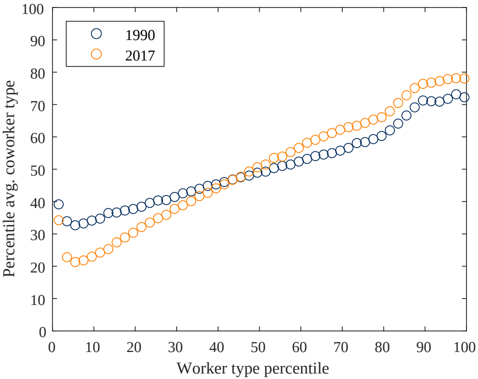
- To what extent do talented workers tend to have talented coworkers?

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

- Robust pattern

▸ Table ▸ Within-occ. nonlinear

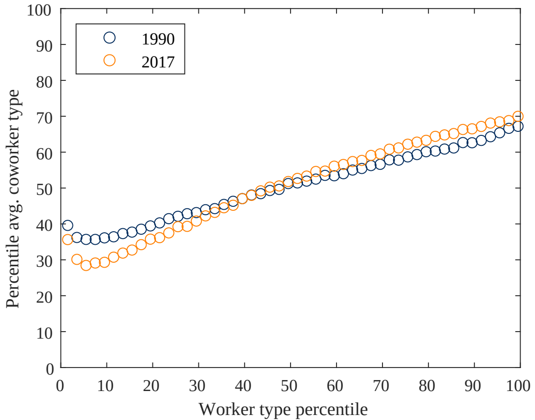
▸ Hakanson et al. (2021)



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



Fact #3: increased education premium due to workplace effects

Main

- **Fact 3: increase in return to schooling is primarily due to workplace effects**

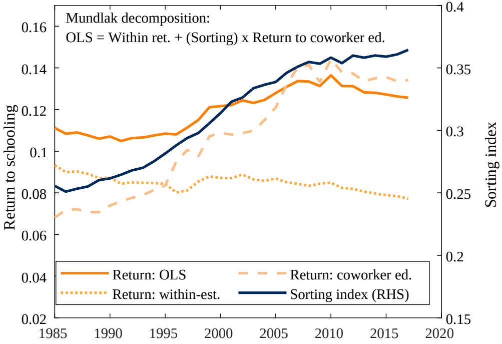
- Mundlak decomposition of year-specific OLS return to schooling:

$$\beta_t^{\text{ols}} = \beta_t^{\text{within}} + \rho_t \times \beta_t^{\text{estab.}}$$

$$\ln w_{it} = \beta_0 + \beta_t^{\text{within}} S_i + \beta_t^{\text{estab.}} \bar{S}_{j(i,t),t} + e_{it}$$

where $\bar{S}_{j(i,t),t}$ is avg. years of schooling in establishment j of worker i in year t

1. β_t^{within}: within-establishment return
2. $\beta_t^{\text{estab.}}$: return to avg. establishment schooling

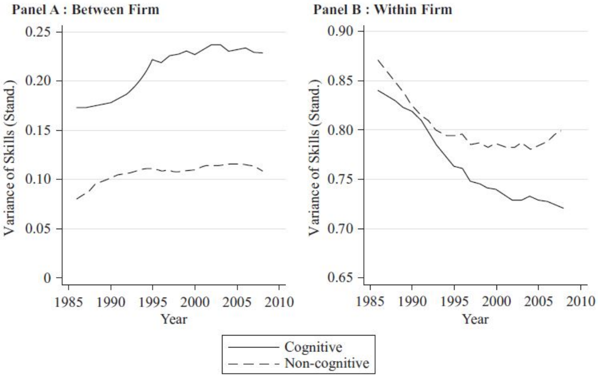


Notes. Plot of coefficients from year-by-year regressions of log wages.

Evidence from the literature: 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



Workers increasingly tend to perform similar tasks across different jobs

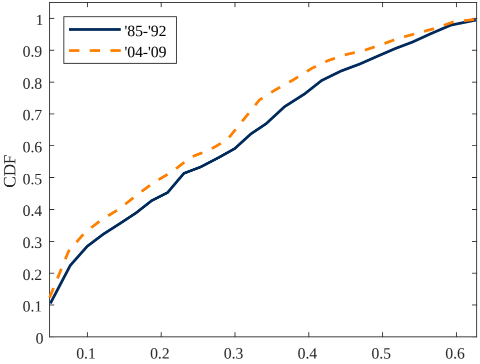
Back

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

Comparison

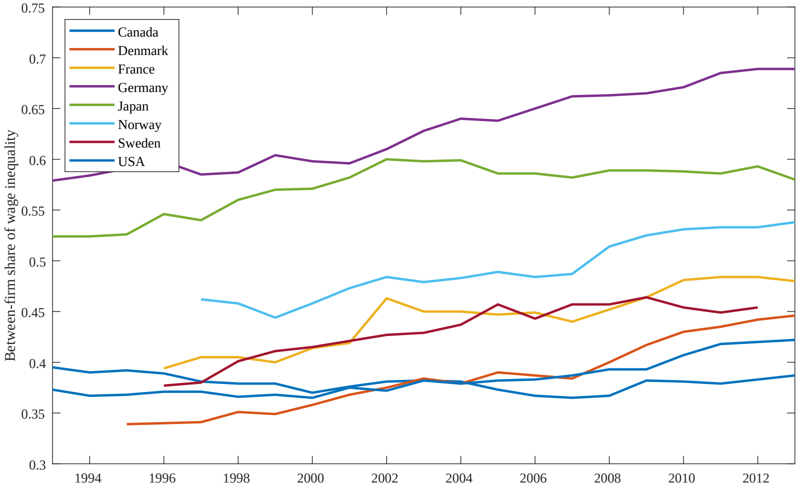
- **Q:** are workers becoming *more* likely to perform similar tasks across jobs, over time?

- **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: ✓at different quantiles



Firming up inequality: cross-country evidence

Main

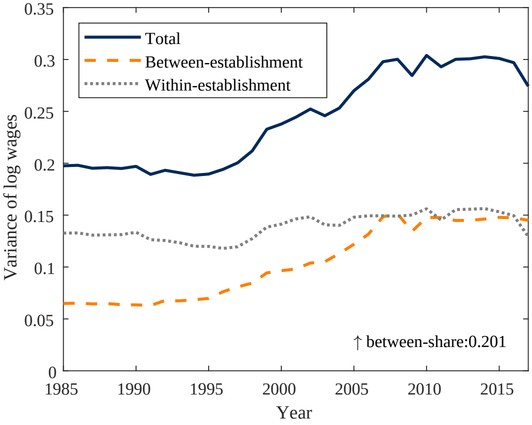


Notes. Data from Tomaskovic-Devey et al. (2020). Measures of earnings differ across countries and, for Germany, between T-D et al. and my study based on the SIEED.

Between-/within-employer wage var decomp. - panel establishments

▸ Main

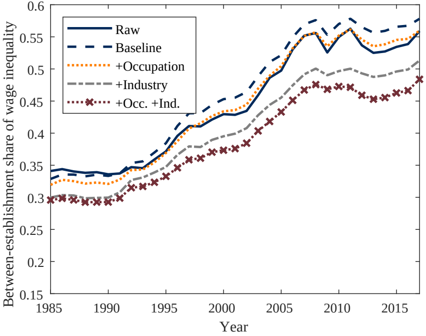
- Instead of considering *all* employers, restrict attention to “panel establishments”



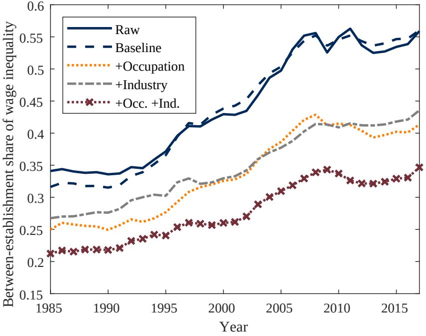
Between-/within-employer wage var. decomp. - alternative w-residuals

▸ Main

- "With worker FEs": regress $\ln \tilde{w}_{it} = \alpha_i + X'_{it}\hat{\beta} + \epsilon_{it}$, construct $\ln w_{it} = \ln(\tilde{w}_{it} - X'_{it}\hat{\beta})$.
- "Without worker FEs": regress $\ln \tilde{w}_{it} = \alpha_0 + X'_{it}\hat{\beta} + \epsilon_{it}$, and consider residuals ϵ_{it}



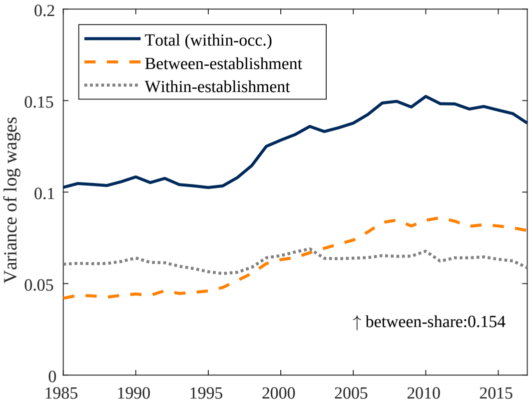
(a) With worker FEs



(b) Without worker FEs

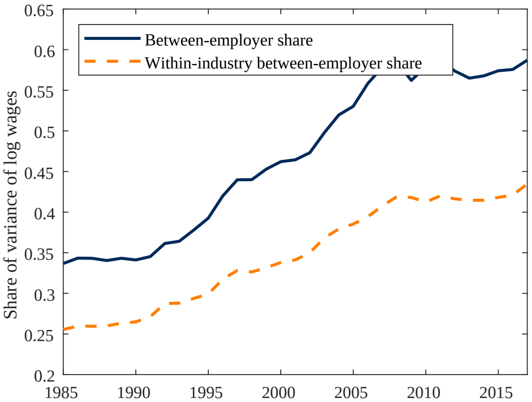
Between-/within-employer wage var. decomp. - within-Occupation

▸ Main



Between-/within-employer wage var. decomp. - within-Industry

▸ Main ▸ Other moments

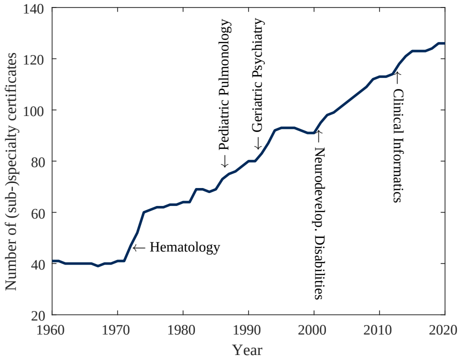


Notes. Based on 'baseline' residualized wages.

Examples: rising specialization

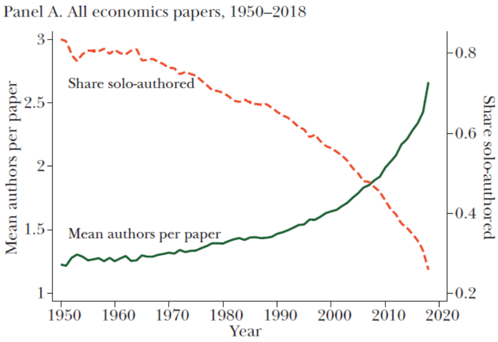
Main

- **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.

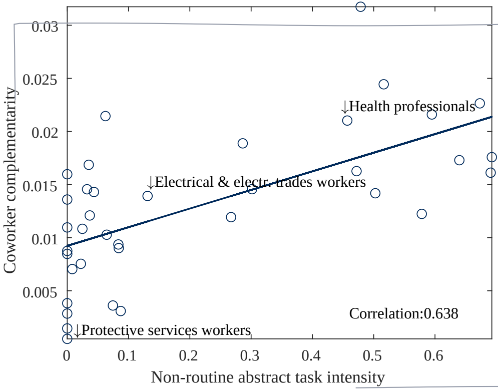
- **Rise of research teams** *[Jones, 2021]*



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

More validation

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



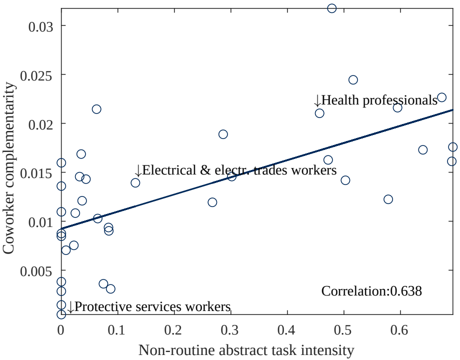
$\hat{\beta}_c$ estimated separately for each occupation

Non-routine abstract task intensity: proxy for χ

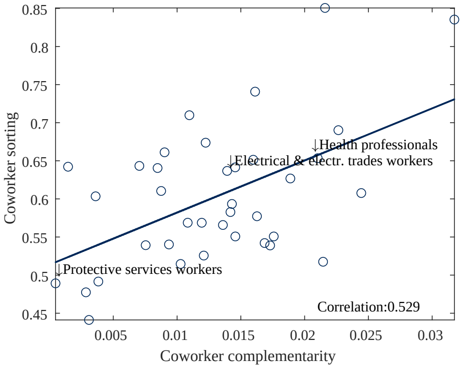
Notes. Quadros de Pessoal microdata. Analysis at ISCO-08-2d level.

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

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



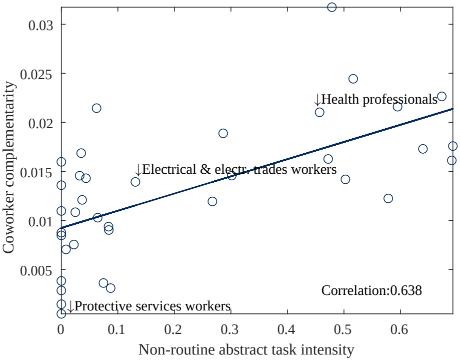
- **↑ Coworker talent complementarity ⇒ ↑ coworker sorting**



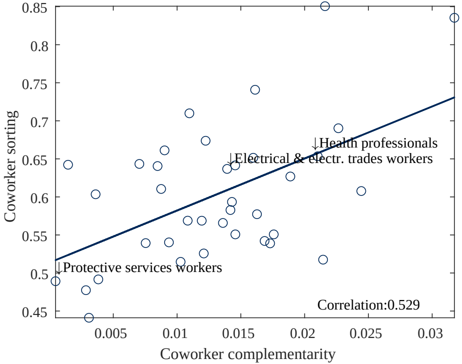
Occupations: task complexity ⇒ complementarity ⇒ sorting

▸ Main

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



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

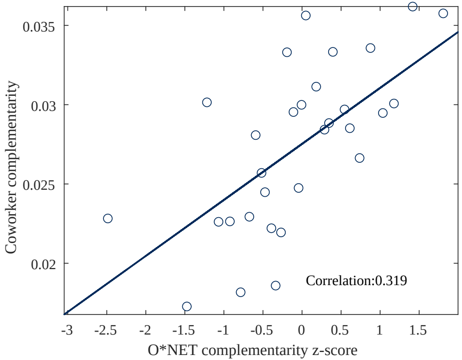


Notes. Quadros de Pessoal microdata. Horizontal axis indicates occupation's reliance on non-routine, abstract (NRA) tasks *[Mihaylov and Tidens, 2019]*.

Industries: coworker importance ⇒ complementarity ⇒ sorting

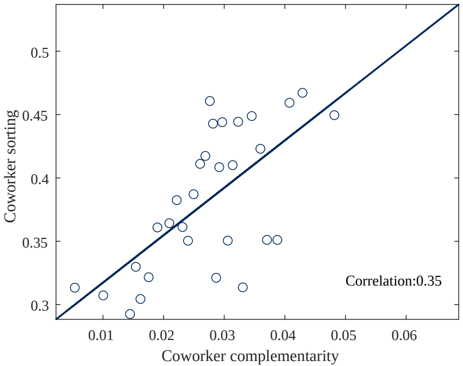
Main

- ↑ **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**

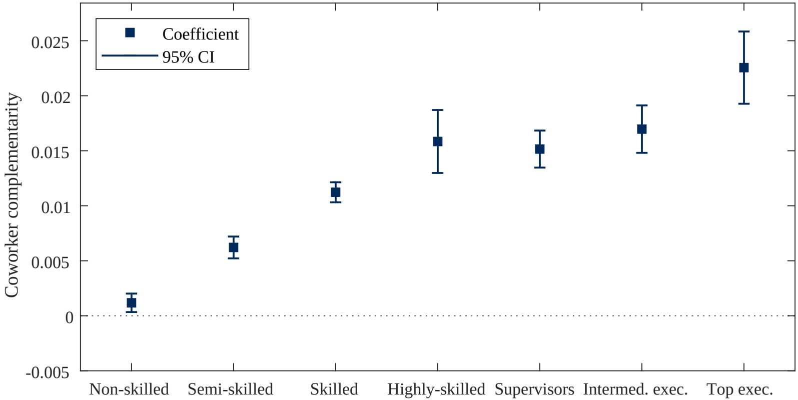


Notes. NACE-4-digit industries.

Hierarchies: complexity ⇒ complementarities

▸ Main ▸ Occupations

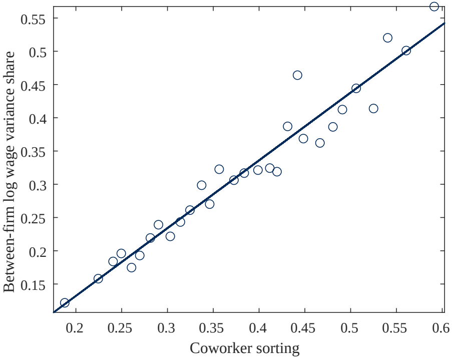
⇒ Coworker wage complementarities are (weakly) ↑ in the layer of a firm's hierarchy



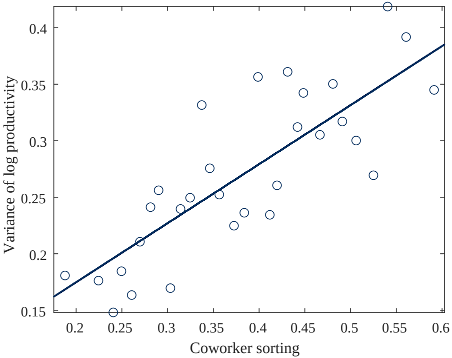
Industries: coworker sorting ⇒ between-firm inequality

Main

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



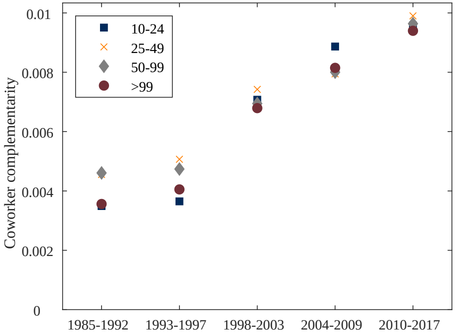
(a) Between-firm share of wage dispersion



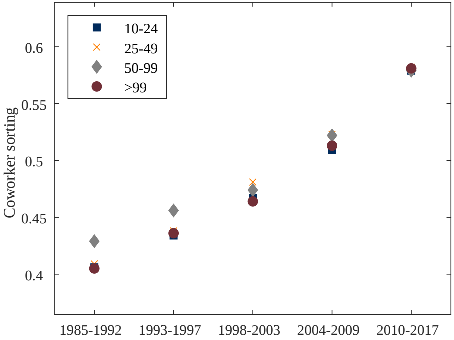
(b) Productivity dispersion

Coworker complementarity & sorting by team size

Robustness

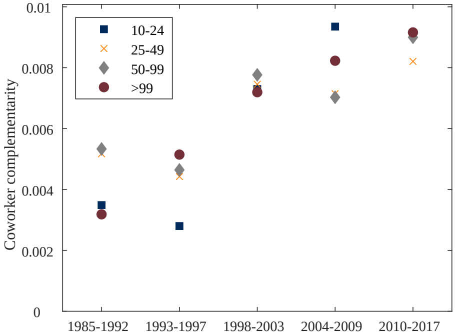


(a) Complementarity

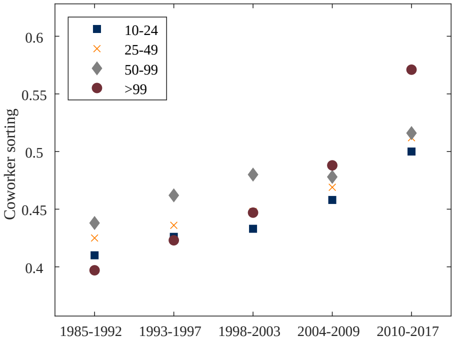


(b) Sorting

Coworker complementarity & sorting by team size – panel estab. only



(a) Complementarity

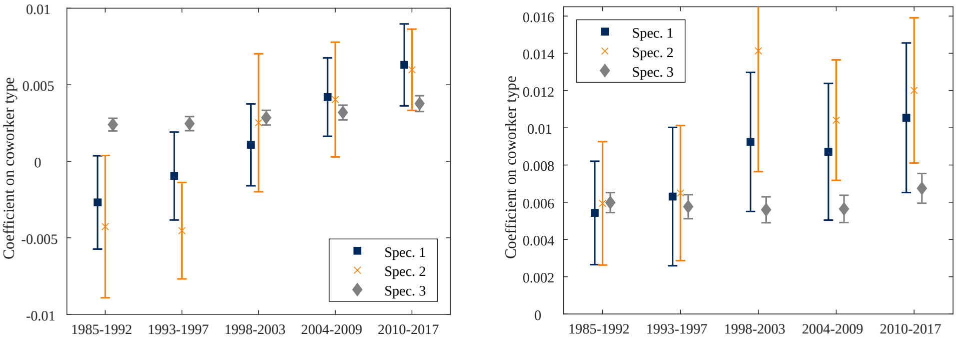


(b) Sorting

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).

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 ($t \rightarrow t+1$), and job movers with non-contiguous E spells ($t \rightarrow t+s$, $s > 1$)
- 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

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

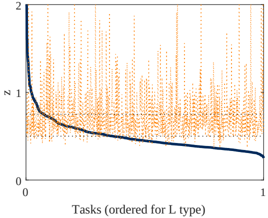
Overview: robustness
Within-industry calibration

	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

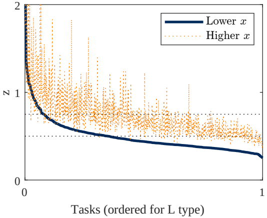
Notes. Within-industry avg. is person-year weighted average across OECD STAN-A38 (2-digit) industries.

Illustration: specialization parameter ▸ Back

- ξ: controls correlation of coworkers' task-specific productivities



(c) Independent draws ($\xi = 1$)



(d) Correlated draws ($\xi \to 0$)

Proof sketch

▸ Main

1. Derive $G(p) := \Pr\{\tilde{\lambda}(\tau) \leq p\}$ given $G_i(p) := \Pr\{\lambda_i(\tau) \leq p\}$, using FOC and max-stability property
2. Use $G(p)$ + standard CES shadow price index to solve (int. by sub.) for

$$\lambda = \left(\int_{\mathcal{T}} \tilde{\lambda}(\tau)^{1-\eta} d\tau \right)^{\frac{1}{1-\eta}}. \tag{5}$$

3. Use $G(p)$ and $G_i(p)$ to derive probability that i produces some task τ, which by LLN (continuum assumption!) is equal to share of tasks produced, π_i
4. Relate λ_i^L to value of all tasks produced by the worker, $\lambda_i^L = \pi_i \lambda Y$
5. Normalize $\lambda = 1$, then algebra yields $Y = f(x_1, ..., x_n; \chi)$

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

▸ Main ▸ Extension to communication frictions

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

1. **Complete division of labor**, with tasks assigned by **comparative advantage**
   - i's task set $\mathcal{T}_i = \left\{ \tau \in \mathcal{T} : \frac{z_i(\tau)}{\lambda_i^L} \geq \max_{k \neq i} \frac{z_k(\tau)}{\lambda_k^L} \right\}$
   - **classic source of efficiency gains**

2. **i's share of tasks $\uparrow$ in i's talent, $\downarrow$ in coworkers' talent**
   - i's task share $\pi_i = \left(x_i^{\frac{1}{1+\chi\xi}}\right) \left(\sum_{k=1}^{n} (x_k)^{\frac{1}{1+\chi\xi}}\right)^{-1}$

Intuition: comparative statics for task shares

Main

- Suppose that $x_i > x_j$. Then

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

Intuition: comparative statics for task shares

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

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

Surplus max. determines which worker types a firm w/ worker x hires

▸ Main

- Joint value of firm with worker x, $\Omega(x)$, satisfies:

$$\rho\Omega(x) = f(x) + \delta\big[-\Omega(x) + V_u(x) + V_{f.0}\big] + (1-\omega)\lambda_{v.u}\int \frac{d_u(\tilde{x}')}{u}\max\{S(\tilde{x}'|x), 0\}d\tilde{x}'$$

  - $V_u(x)$: value for unemp. worker; $V_{f.0}$: value for vacant firm; $S(x)$: surplus from zero-worker firm hiring x
  - $d_u(x)$: density of unemployed workers of type x; $u = \int d_u(x)dx$
  - ω: worker bargaining wgt; δ: sep. rate; $\lambda_{v.u}$: rate of vacancy meeting unmatched worker

Surplus max. determines which worker types a firm w/ worker x hires ▸ Main

- Joint value of firm with worker x, $\Omega(x)$, satisfies:

$$\rho\Omega(x) = f(x) + \delta\Big[-\Omega(x) + V_u(x) + V_{f.0}\Big] + (1-\omega)\lambda_{v.u}\int \frac{d_u(\tilde{x}')}{u}\max\{S(\tilde{x}'|x), 0\}d\tilde{x}'$$

  - $V_u(x)$: value for unemp. worker; $V_{f.0}$: value for vacant firm; $S(x)$: surplus from zero-worker firm hiring x
  - $d_u(x)$: density of unemployed workers of type x; $u = \int d_u(x)dx$
  - ω: worker bargaining wgt; δ: sep. rate; $\lambda_{v.u}$: rate of vacancy meeting unmatched worker

- Surplus $S(x|x')$ reflects complementarities – and hiring decisions reflect surplus

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

$$h(x|x') = \mathbf{1}\{S(x|x') > 0\}$$

HJB: unmatched

Main

- Unmatched firm:

$$\rho V_{f.o} = (1-\omega)\lambda_{v.u} \int \frac{d_u(x)}{u} S(x)^{+} dx \tag{6}$$

- Unmatched worker:

$$\rho V_u(x) = b(x) + \lambda_u \omega \left[\int \frac{d_{f.o}}{v} S(x)^{+} + \int \frac{d_{m.1}(\tilde{x}')}{v} S(x|\tilde{x}')^{+} d\tilde{x}' \right] \tag{7}$$

HJB: surpluses

- Surplus of coalition of firm with worker x

$$(\rho + \delta)S(x) = f(x) - \rho\big(V_u(x) + V_{f.0}\big) + \lambda_{v.u}(1 - \omega)\int \frac{d_u(\tilde{x}')}{u} S(\tilde{x}'|x)^+ d\tilde{x}' \tag{8}$$

- Surplus from adding x to x'

$$S(x|x')\big(\rho + 2\delta\big) = f(x, x') - \rho\big(V_u(x) + V_u(x') + V_{f.0}\big) + \delta S(x) - (\rho + \delta)S(x') \tag{9}$$

KFE: unemployed

$$\delta\left(d_{m.1}(x) + \int d_{m.2}(x, \tilde{x}')d\tilde{x}'\right) = d_u(x)\lambda_u\left(\int \frac{d_{f.0}}{v} h(x, \tilde{y}) + \int \frac{d_{m.2}(\tilde{x}')}{v} h(x|\tilde{x}')d\tilde{x}'\right). \tag{10}$$

KFE: one-worker matches

$$d_{m.1}(x)\left(\delta + \lambda_{v.u}\int \frac{d_u(\tilde{x}')}{u} h(\tilde{x}'|x)d\tilde{x}'\right) = d_u(x)\lambda_u \frac{d_{f.0}}{v} h(x) + \delta \int d_{m.2}(x,\tilde{x}')d\tilde{x}'. \tag{11}$$

KFE: two-worker matches

$$2\delta d_{m.2}(x,x') = d_u(x)\lambda_u \frac{d_{m.1}(x')}{v} h(x|x') + d_u(x')\lambda_u \frac{d_{m.1}(x)}{v} h(x'|x). \tag{12}$$

Measurement: a useful identification result

Main · Non-separable case: scatterplot

- **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'

Wage eq.

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

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*

Mapping model to data: coworker types

Main

- 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 model to data: identification strategy for χ

Main

- **Literature:** complementarities – primarily between workers and firms – usually inferred indirectly from sorting patterns
  - exception: Hagedorn-Law-Manovskii (2017)
- **This paper:** directly measure coworker complementarity in the data, recover χ structurally given $\gamma = \frac{\chi}{\chi+1}$
- Paper does *not* use microfoundation itself to measure χ, respectively γ
- Experiment: fit a (truncated) Fréchet distribution to Grigsby's (2023) non-parametric estimates of the multi-dimensional skill dist. estimated from CPS data
  - recover $\gamma = 0.84$ for 2006 but *very* noisy estimates Jump
- **Ongoing work:** use the extended microfoundation to identify χ Jump

Direct estimation of χ: proof of concept

- Grigsby (2023): only paper that provides a *cardinal* measure of skill task-specificity
- Evidence on time trends are qualitatively consistent with "specialization hypothesis": cross-type average of within-type variance across specific skills grew by nearly 50% b/w 1980s and 2000s & skill transferability has declined amongst high-skill occupations
- His operationalization of worker types and tasks does *not* directly map onto my model (no identifying assumption; coarse occupational skills; US vs DE data)
- **Proof of concept:** but *suppose* we just take those data, extract moments capturing average within-worker cross-task efficiency dispersion, fit a (truncated) Fréchet, recover $\gamma = \frac{\chi}{1+\chi}$
  $\Rightarrow$ ✓ γ **similar to structural estimation result based on evidence from wage CC**

Semi-structural back-of-envelope calculation for γ

Main

- Structurally recover $\gamma \frac{\chi}{\chi+1}$ by estimating $\frac{\partial^2 w(x|x')}{\partial x \partial x'}$ in the data, which was shown to be proportional to $\frac{\partial^2 f(x,x')}{\partial x \partial x'}$
- But how is $\frac{\partial^2 f(x,x')}{\partial x \partial x'}$ related to γ?
- Definitionally, $\gamma = (f f_{ij})/(f_i f_j)$ for any $i \neq j$
- Can we avoid full structural model? $\Rightarrow$ If have measures not only of f_{ij} but also output f and marginal products f_i
- Suppose, for any x and x', we use *wages* to back out marginal products – competitive wage determination rather bargaining! – and recover output from sum of wages divided by labor share
- Find $\gamma \approx 0.79$ – very close to structural estimate!

Overview of validation exercises: direction EE transitions & cross-section

Main

- 2 additional types of validation exercise:
  - ✓**EE transitions** reallocate workers to more + assortative matches — Details
  - do model-implied relationships also hold in **cross-section**?
    1. $\chi \uparrow \Rightarrow$ coworker complementarity $\uparrow$
    2. coworker complementarity $\uparrow \Rightarrow$ + assortative matching $\uparrow$

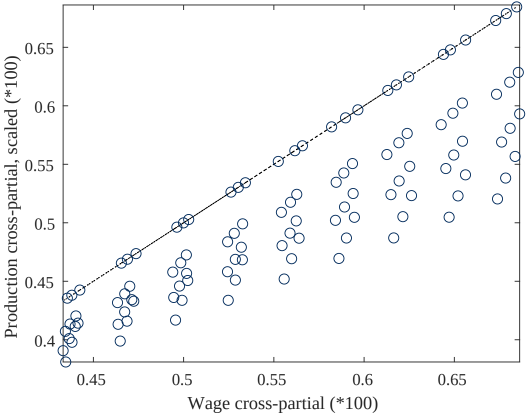
can test predictions *because* we have measures of complementarity!

- Implementation of cross-sectional exercises: rich Portuguese micro data
  - universe of private-sector actors, employer-employee data & income statements
- Cross-sectional exercises:
  - ✓ **Hierarchies** — Details
  - ✓ **Industries** — Details
  - ✓ **Occupations** — Details

Wage and production cross-partials beyond the benchmark

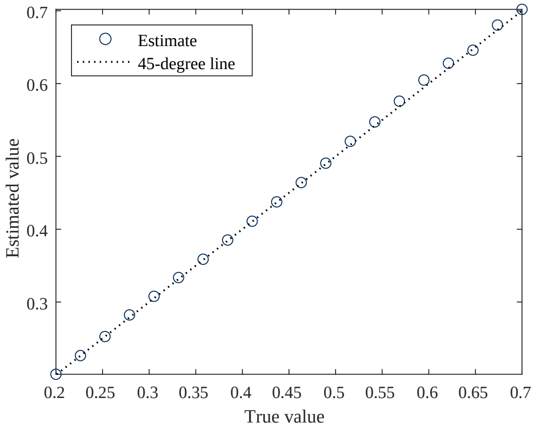
Main

- Solve model for many combinations of χ, λ_e and b
- Compare FD approx of $f_{xx'}(x, x')$, scaled by ω, and $w_{xx'}(x|x')$
- Main parameter driving wedge: λ_e



Monte Carlo study

Main



Implications for aggregate productivity

▸ Effects of χ ↑: random vs eqm

- **Production complementarities imply coworker sorting matters for agg productivity**
    - $f(x_1, \cdots, x_n) = n^{1+\chi\xi} \times \left(\frac{1}{n} \sum_{i=1}^{n} (x_i)^{\frac{1}{1+\chi\xi}} \right)^{1+\chi\xi}$

Implications for aggregate productivity

Effects of χ ↑: random vs eqm

- **Production complementarities imply coworker sorting matters for agg productivity**
  - $f(x_1, \cdots, x_n) = n^{1+\chi\xi} \times \left(\frac{1}{n} \sum_{i=1}^{n} (x_i)^{\frac{1}{1+\chi\xi}} \right)^{1+\chi\xi}$
- **Search** frictions induce misallocation $\sim$ coworker **mismatch**

Implications for aggregate productivity

▸ Effects of χ ↑: random vs eqm

- **Production complementarities imply coworker sorting matters for agg productivity**
  - $f(x_1, \cdots, x_n) = n^{1+\chi\xi} \times \left(\frac{1}{n} \sum_{i=1}^{n} (x_i)^{\frac{1}{1+\chi\xi}} \right)^{1+\chi\xi}$
- **Search** frictions induce misallocation $\sim$ coworker **mismatch**
- **Quantify** mismatch costs: compare eqm outcome vs to productivity under pure PAM
  $\Rightarrow$ 2010s gap: 2.05%, similar for earlier periods

Implications for aggregate productivity

Effects of χ $\uparrow$: random vs eqm

- **Production complementarities imply coworker sorting matters for agg productivity**
  - $f(x_1, \cdots, x_n) = n^{1+\chi\xi} \times \left(\frac{1}{n} \sum_{i=1}^{n} (x_i)^{\frac{1}{1+\chi\xi}} \right)^{1+\chi\xi}$
- **Search** frictions induce misallocation $\sim$ coworker **mismatch**
- **Quantify** mismatch costs: compare eqm outcome vs to productivity under pure PAM
  $\Rightarrow$ 2010s gap: 2.05%, similar for earlier periods
- **Trends:** $\uparrow$ talent sorting limited $\uparrow$ in mismatch costs given $\chi \uparrow$

  $\Rightarrow$ no-reallocation counterfactual: productivity gap 4.65%

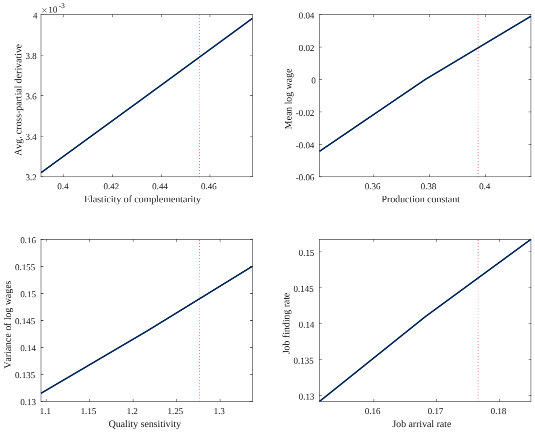
Parameterization, including estimation results (2010s)

▸ Main

Parameter	Description	Targeted moment	Value	m	$\hat{m}$
χ	Specialization	$\hat{\beta}_c$	0.67	0.0058	0.0058
a_0	Production, constant	Avg. wage (norm.)	0.29	1	1
a_1	Production, scale	Var. log wage	1.71	0.241	0.241
b_1	Replacement rate, scale	Replacement rate	0.60	0.63	0.63
δ	Separation hazard	Job loss rate	0.008	0.008	0.008
λ_u	Meeting hazard	Job finding rate	0.22	0.162	0.162
ρ	Discount rate	External	0.008		
ω	Worker bargaining weight	External	0.50		
$\bar{n}$	Effective team size	External	25		

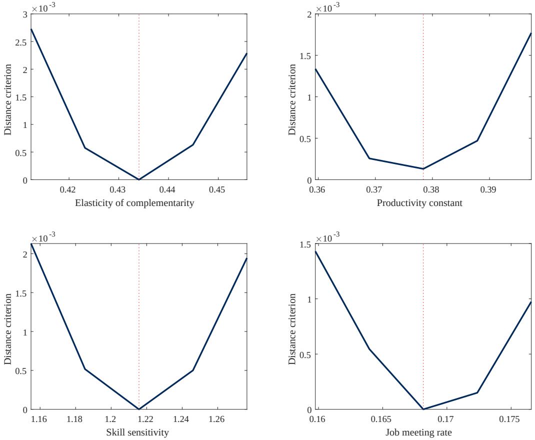
Identification validation exercise 1

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 $\mathcal{G}$.

What about shifts in the talent distribution?

Robustness overview

- **Q:** could $\uparrow$ economy-wide x-dispersion, σ_x, also explain $\uparrow$ firm-level inequality?
  - Kremer-Maskin (1996): as the dispersion of skills (talent) $\uparrow$ in the economy, the relative dispersion of talent within firms $\downarrow$
- **Obs. #1:** Hakanson et al. (2021): direct skill measures (from military enlistment tests) point to Flynn effect but not not to increased dispersion (in cognitive test scores)
- **Obs. #2:** model exercise
  - method: instead of rank interpretation, i.e. $x \sim U$, we separately parameterize σ_x by assuming a $\mathcal{N}$ distribution
  - finding: $\sigma_x \uparrow \Rightarrow \rho_{xx} \uparrow$, $\sigma_w \uparrow$, and $\sigma_{\bar{w}}/\sigma_w \uparrow$ – *but* no measured increase in $\hat{\beta}_c$
- **Conclusion:** empirically unclear whether $\sigma_x \uparrow$, and if so, this would not explain the observed increase in coworker complementarity, i.e. latter is a distinct channel

Training policies in a team production context (w-i-p)

▸ Overview

- "Training policies" $\sim$ non-parametric perturbations of the talent distribution
  - left-tail intervention: give everyone in 1st decile productivity of those in 2nd decile
  - right-tail intervention: give everyone in 9th decile productivity of those in 10th decile
- Team production: effect of "training policies" partially via coworker spillovers!
- Relative effectiveness of left-tail vs right-tail intervention:
  1. the stronger are coworker complementarities, the relatively greater are the realized productivity gains from a left-tail intervention, b/c low-x tend to be weak links
  2. but raising the productivity of coworkers of workers with high productive potential generates greater gains – and with sorting, those coworkers are themselves high-x
- Equilibrium: relative effectiveness of left-/right-tail training depends on both forces
  - tentative quantitative finding: right-tail intervention boosts average productivity by more *but* left-tail training also lowers inequality

The effect of declining search frictions

▸ Main

- $\downarrow$ search frictions could also explain $\uparrow$ coworker sorting
- Job arrival & separation rates estimated to $\uparrow$ from p1 to p2
- **Counterfactual analysis:** explains 6% of model-implied $\uparrow$ in between-employer share of wage variance

	Δ model	Implied % Δ model due to Δ parameter
Model 1: baseline	0.159	-
Cf. a: fix period-1 complementarity	0.065	59
Cf. b: fix period-1 search frictions	0.150	6

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 ↑)
  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

Robustness: model with OJS - brief overview

Main

- **Baseline:** no job-to-job transitions – but is main story robust when workers can switch to better job after accepting job out of unemployment?
  - two opposing effects from increased complementarities
- **Extension:** employed workers also meet vacancies at Poisson rate λ_e
- **Main findings:**
  - better fit to empirical sorting patterns in cross-section Comparison
  - contribution of $\uparrow$ complementarities to $\uparrow$ firm-level wage inequality slightly smaller, more attributed to $\uparrow$ labor market transitions

    – conservative estimates (endogenous search effort, forward-looking wage specification)
- Opens door to thinking about coworker complementarities and **job ladders** Jump

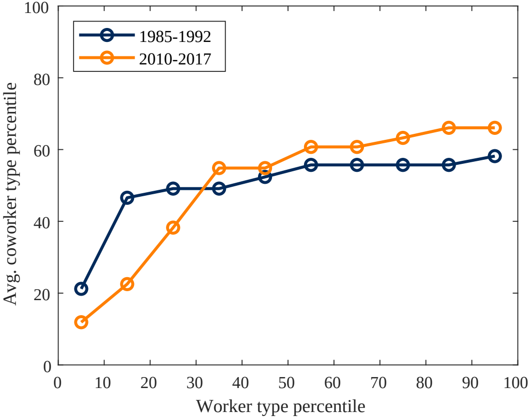
Extension: model with OJS

Main

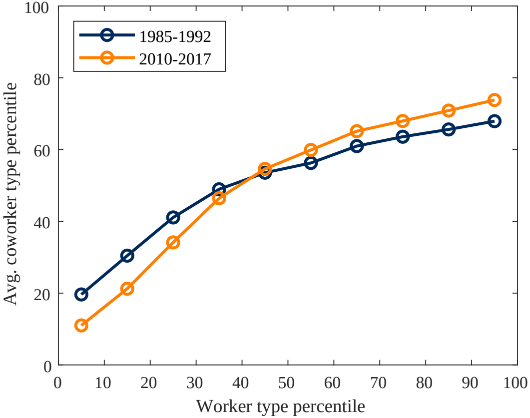
- Baseline model abstracted from OJS
  - transparent trade-off, connection to analytical results
- **Consider extension to OJS:** employed worker meet vacancies at Poisson rate λ_e
  - wages both off and on the job are continuously renegotiated under Nash bargaining, with unemployment serving as the outside option *[cf. di Addario et al., 2021]*
  - re-estimate, with empirical labor market flows disciplining λ_e
- Qualitative question: is coworker sorting outcome robust, even if workers can switch to better job after accepting job out of unemployment?
- **Analyses:**
  1. coworker sorting patterns & changes
  2. additional model validation: direction of EE flows in model & data

Model-implied coworker sorting patterns: without and with OJS

Main



(a) Baseline



(b) OJS

EE transitions in theory and data

Main

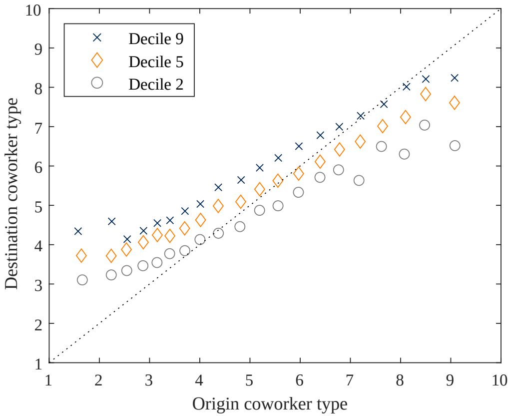
- **Theoretical prediction:** EE transitions move workers in surplus-maximizing direction $\Rightarrow \Delta\hat{x}_{-it} = \hat{x}_{-i,t} - \hat{x}_{-i,t-1}$ should be *positively* correlated with $\hat{x}_i$
  - $h_{2.1}(x, x''|x') = 1$ – worker x in a two-worker firm with coworker x'' would move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

Main

- **EE transitions push toward greater coworker sorting:** for any given origin, higher *x*-workers move to workplaces with better coworkers than lower-*x* workers do
- *But* in data 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		

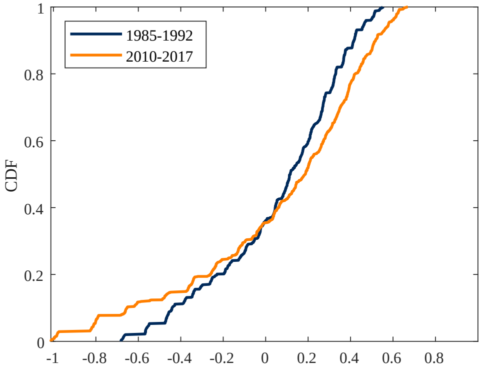
Table 1: Change in coworker type due to EE moves positively related to own type – increasingly so

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.

Productivity dispersion

Main

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

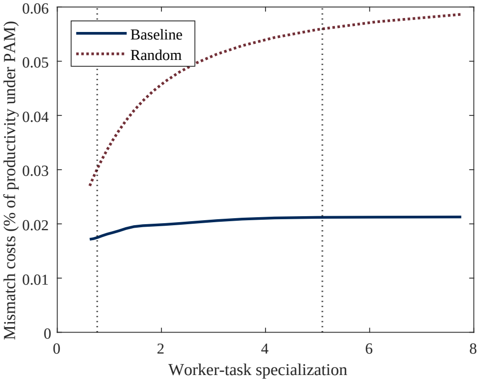


Productivity costs of complementarity & labor market functioning

▸ Main

"The benefits of the division of labor are limited by the functioning of the labor market"

- Microfoundation: $\uparrow \chi \Rightarrow \uparrow$ efficiency *benefit* from teamwork *but* also $\uparrow$ mismatch *costs*
- **Q:** how does the gap to potential vary depending on labor market structure?
- **A:** under random sorting, productivity gap due to misallocation $\uparrow$ more sharply as $\chi \uparrow$
- Outside model: severe labor mkt frictions (e.g., dev'ing countries *[Donovan et al., 2023]*) may inhibit specialization *[cf. Atencio et al., 2023; Bassi et al., 2023]*



Implications for overall inequality?

Main

- **Coworker complementarities do not necessarily ↑ variance of person-level wages**
  - (un-)surprising? Varriance decomposition perspective vs. common intuition *[Kremer, 1993]*
  - (i) **reallocation effect**, (ii) valuation effect, (iii) outside option effect

- Several mechanisms though through which ↑ sorting could ↑ person-level inequality
  1. regulation or norms that lead to within-firm wage compression *[Akerlof-Yellen, 1990]*
  2. coworker learning *[Jarosch et al., 2021; HLMP, 2023]*
  3. increasing returns to labor quality *[Kremer, 1993]*

Implications for AI impact on labor markets (1): overview

Main

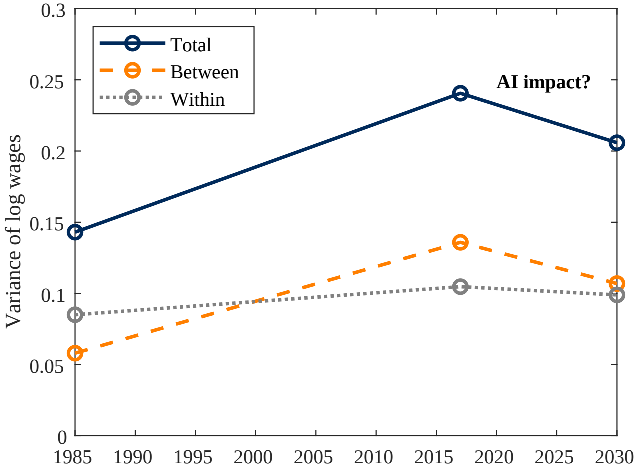
- Literature moving toward richer way of thinking about tech. change. Here: shifts in across-worker productivity differences *and* in specialization/interdependence

- Example: **AI** – early/conjectural evidence on impact of LLMs etc
  - *absolute* adv. less important ("leveller") *[e.g., Brynjolfsson et al., 2023]*
  - comparative adv. less important (e.g., can interpret medical imaging w/o radiologist)

- **Illustrative model counterfactual:** relative to the 2010s
  1. everyone's productivity ↑ by equivalent to 20% of lowest type's productivity – so in proportional terms, lower-x benefit more
  2. coworker complementarity ↓ by 20%

Implications for AI impact on labor markets (2): illustrative exercise

- Model prediction: AI could lead to reversal of historical trends – $\sigma^2_w \downarrow$, $\sigma^2_{\bar{w}} \downarrow$, $\rho_{xx} \downarrow$ – along side a productivity boom
  - driven by $\downarrow$ firm-level wage inequality



Implications for AI impact on labor markets (3): conjectures

Main

- Literature moving toward richer way of thinking about tech. change. Here: shifts in across-worker productivity differences *and* in specialization/interdependence
- Example: AI – early/conjectural evidence on impact of LLMs etc
  - *absolute* adv. less important? ("leveller") *[e.g., Brynjolfsson et al., 2023]*
  - comparative adv. less important? (e.g., can interpret medical imaging w/o radiologist)
- Illustrative model counterfactual: **AI could lead to reversal of historical trends...**
- **...and perhaps also...**
  1. ...**flatter organizations**, with managerial span of control ↑
  2. ...↓ **barriers to entry** for self-employment/start-ups
  3. ...**easier job transitions** & shorter training durations

Within-industry calibration: overview

▸ Main ▸ Data moments

- Baseline: calibration and evaluation based on economy-wide moments
- But the model does not incorporate between-industry differences in, e.g., production technology
- Alternative considered here: target $\hat{\beta}_c$ and σ_w^2 computed as within-industry average, evaluate against within-industry trends
  - keep other targets (e.g., job separation) as before

Within-industry calibration: model fit & counterfactual

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