Superstar Teams: The Role of Firms in Shaping Wage Inequality

Lukas B. Freund

University of Cambridge

October 25, 2023

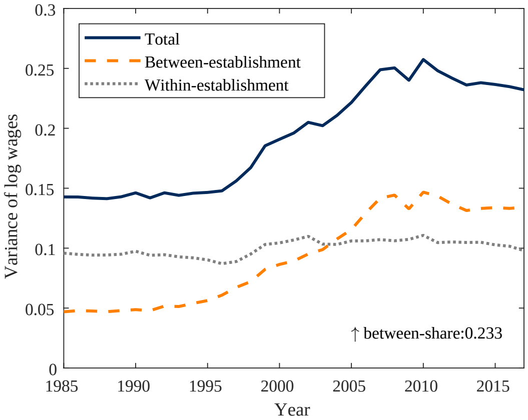
Cambridge Society for Economic Pluralism: Paper Zero Series

Motivation: important role of firms in the evolution of labor market inequality

- **Evidence:** high & ↑ firm-level wage dispersion

▸ Details ▸ USA

⇒ **Wage inequality increasingly a *between-firm* phenomenon**

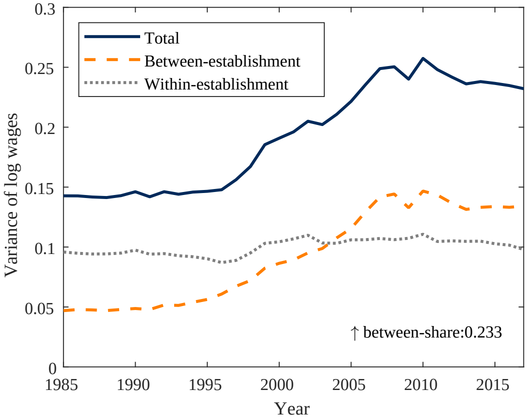


Motivation: important role of firms in the evolution of labor market inequality

- **Evidence:** high & ↑ firm-level wage dispersion

→ canonical **explanations?**

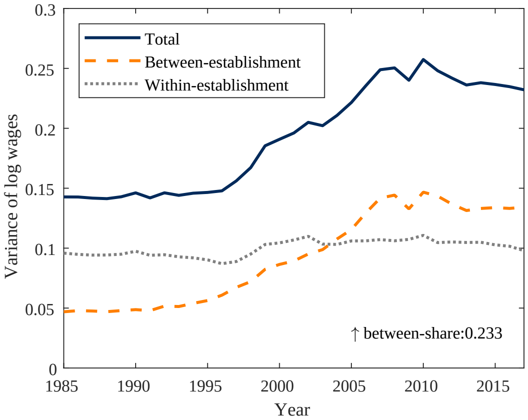
→ firm-oriented **policies?**



Motivation: important role of firms in the evolution of labor market inequality

- **Evidence:** high & ↑ firm-level wage dispersion

Why?



The main idea of the paper in one slide

- **Novel theory:**
  1. individuals differ in talent & their productivity varies across tasks (**specialization**)
  2. production involves **teamwork**
  3. employers and employees **search** for each other
- **Mechanism: specialization $\Rightarrow$ coworker complementarity $\Rightarrow$ coworker sorting**

  $\leadsto$ talented workers gain more from talented colleagues
- **Application: specialization $\uparrow$ $\Rightarrow$ "firming up" of wage inequality**

Roadmap

1. High-level **explanation**: how does the paper formalize & test this idea?
    1. (a tiny bit of) theory
    2. data
    3. quantitative analysis using theory + data
2. The **bigger picture**: economics, inequality, theory & data, pluralism

A Tiny bit of Theory

Model ingredients

- **Environment:**
  1. **Specialization:** individuals vary in their talent, $x_i \in [0, 1]$, and each is *specialized:* the greater is parameter $\chi \in [0, \infty]$, the more their productivity varies across tasks
  2. **Team production:** in each workplace, n individuals collaborate
  3. **Search:** it takes time to find an employer/employee
- The firm acts as a **"team assembly technology":**
  1. facilitates division of labor
  2. assigns tasks to maximize total production

Result: specialization $\Rightarrow$ coworker complementarity

Proposition: Aggregation result

Team output Y can be written as a function $\mathbb{R}_+^n \to \mathbb{R}_+$

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

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

1. **Efficiency gains** $\nearrow \chi$

Result: specialization $\Rightarrow$ coworker complementarity

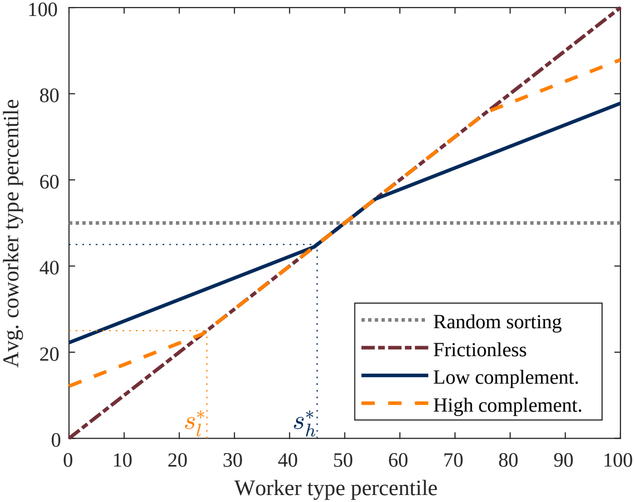
Proposition: Aggregation result

Team output Y can be written as a function $\mathbb{R}_+^n \to \mathbb{R}_+$

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

1. Standard efficiency gains $\nearrow \chi$
2. **Coworker complementarity** $\nearrow \chi$: talent bottlenecks

Implications: who works with whom?

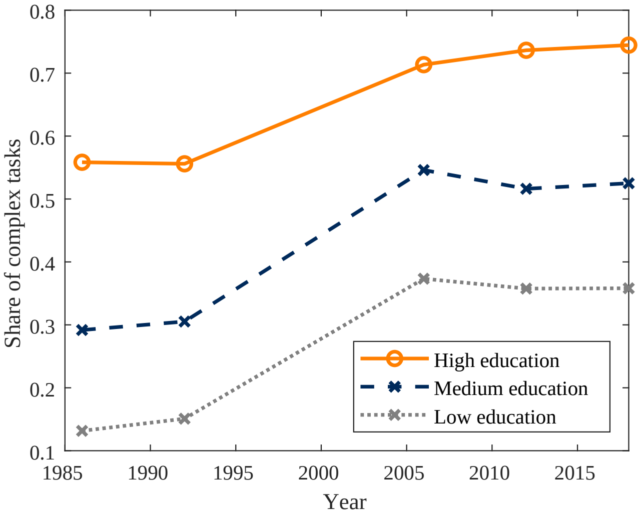


Data

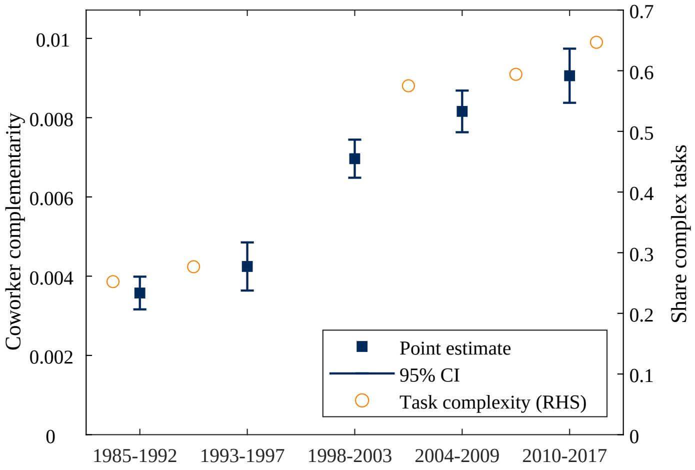
Micro data (DE): matched employer-employee data *and* surveys

- **Matched employer-employee data:** who works for and with whom?
- **Long-running surveys:** what tasks do people perform at work?

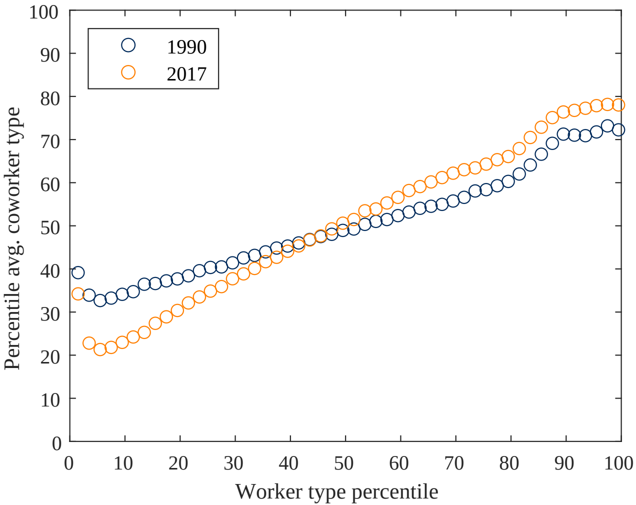
Evidence: specialization (χ) has $\uparrow$ since the mid-1980s



Coworker complementarity has intensified concurrently

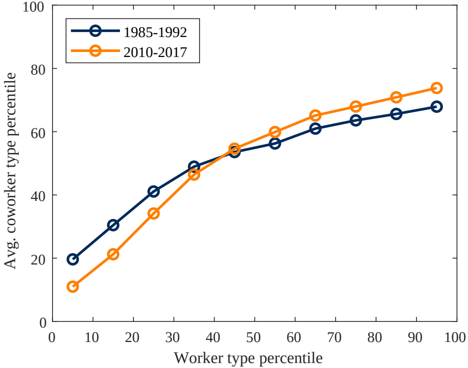


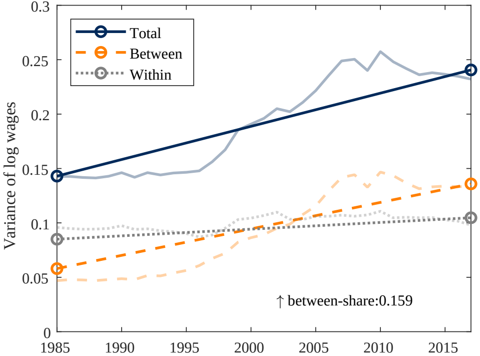
Talented workers increasingly collaborate



Quantitative analysis

Model matches empirical trends





Models allow us to study counterfactual worlds

- **Q:** How much of $\uparrow$ between-firm share of wage variance is due to $\uparrow$ complementarities (CC)?
- **Counterfactual:** between-firm share in 2010s absent $\uparrow$ CC
- **A:** **$\uparrow$ CC accounts for 59% of model-predicted $\Delta \leftrightarrow \approx$40% of empirical Δ**

	Δ model	Implied % Δ model due to Δ parameter
Model baseline	0.16	-
Cf.: fix period-1 complementarity	0.065	59

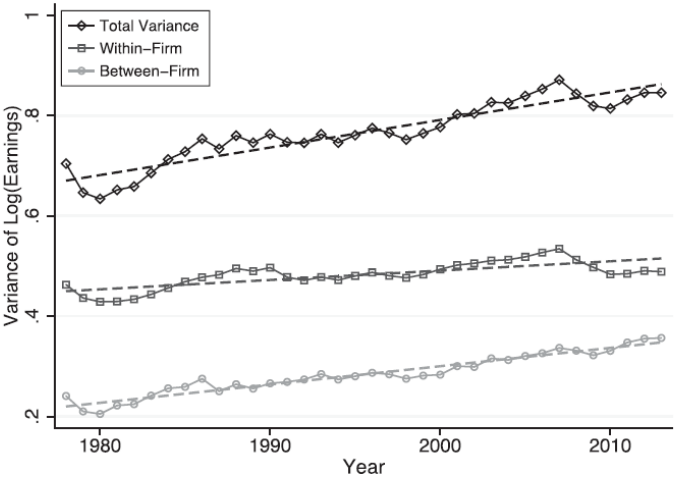
Concluding thoughts

A few discussion points

- **So what?** Individual inequality, firms, social interactions
- **Small steps:** Science is collaborative and, at most times, incremental
- **Why do firms exist?** Form & coordinate teams $\leftarrow$ matters $\uparrow$ as complexity rises
- **Technological change** is multi-dimensional; AI?
- **Theory & data:** fruitful interplay!
- **Mainstream & pluralism:** fruitful interplay?

Extra Slides

Evidence for USA: Bloom et al. (2019) Back



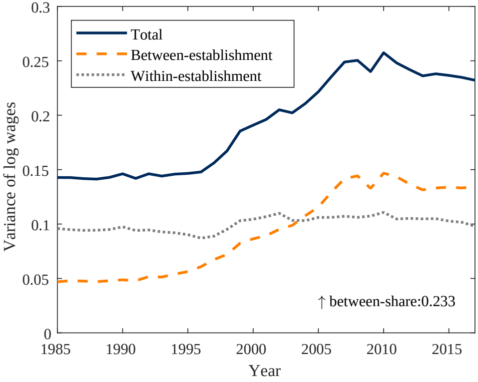
Motivating fact 1: ↑ between-firm share of wage inequality

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.

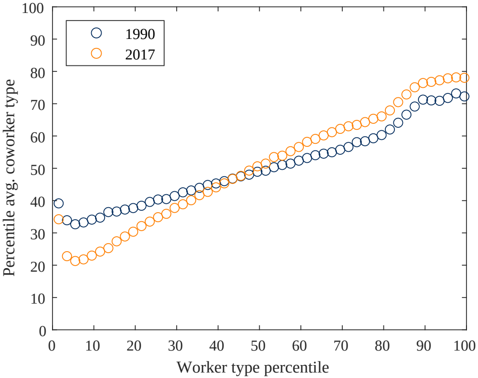
Motivating fact #2: talented workers increasingly collaborate

Intro Var. decomp.
Fact #3

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

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

Table Within-occ. nonlinear
Hakanson et al. (2021)



Motivating fact #3: increased education premium due to workplace effects

▸ Main

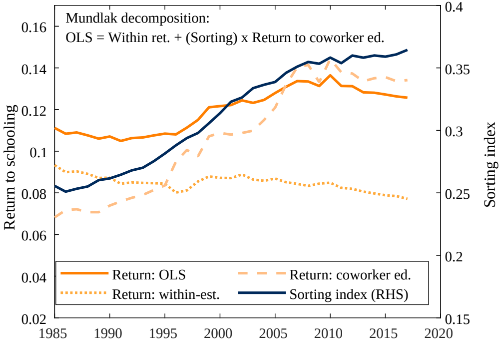
- **Fact 3: increase in return to schooling is primarily due to workplace effects**

- 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.

Workers increasingly tend to perform similar tasks across different jobs

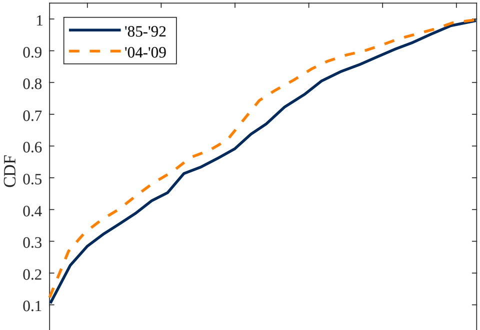
Back

- ✓Workers move to jobs with similar task requirements 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)
  - unconditional average: 0.253 → 0.227: 10% decline

- Robustness: regress distance on sample-period dummies, controlling for potential experience, gender, unemployment rate, origin occupation



Brief summary of methodology: occupations and task space

Details

1. **Measure task similarity between occupations**
   - $\bar{\mathbf{l}}_o = (\bar{l}_{o1}, \ldots, \bar{l}_{o|\hat{\mathcal{T}}|})$: vector of task content of occupation *o*, with $\bar{l}_{o\tau}$ denoting the fraction of workers in occupation *o* performing task $\tau \in \hat{\mathcal{T}}$, where $|\hat{\mathcal{T}}| \in \mathbb{Z}_{++}$ (i.e., discretized)
   - *distance in task space* between any two occupations *o* and *o′*,

$$\varphi(\bar{\mathbf{l}}_o, \bar{\mathbf{l}}_{o'}) = \frac{1}{\pi}\cos^{-1}\left(\frac{\bar{\mathbf{l}}_o'\bar{\mathbf{l}}_{o'}}{||\bar{\mathbf{l}}_o|| \cdot ||\bar{\mathbf{l}}_{o'}||}\right) \in [0, 1]$$

   - **implemntation:** BIBB longitudinal microdata ($|\hat{\mathcal{T}}| = 15$)
2. **Occupational movers:** an individual *i* who in period *t* is employed at *j* in occupation *o* counts as an *occupational mover* if in $t+1$, *i* is employed at $j' \neq j$ and $o' \neq o$
   - implementation: matched employer-employee micro data (entire biography!)
3. Merge $\{\varphi(\bar{\mathbf{l}}_o, \bar{\mathbf{l}}_{o'})\}_{oo'}$ into mover sample $\rightarrow$ **project moves onto task distance space**
   - harmonized occupational classification KldB1988-2d

Methodology (1): occupations and task space

▸ Main

- **Occupations as points in task space – definitions:**
  - $\bar{\mathbf{l}}_o = (\bar{l}_{o1}, \ldots, \bar{l}_{o|\hat{\mathcal{T}}|})$: vector of task content of occupation o, with $\bar{l}_{o\tau}$ denoting the fraction of workers in occupation o performing task $\tau \in \hat{\mathcal{T}}$, where $|\hat{\mathcal{T}}| \in \mathbb{Z}_{++}$ (i.e., discretized)
  - *distance in task space* between any two occupations o and o',

$$\varphi(\bar{\mathbf{l}}_o, \bar{\mathbf{l}}_{o'}) = \frac{1}{\pi}\cos^{-1}\left(\frac{\bar{\mathbf{l}}_o'\bar{\mathbf{l}}_{o'}}{||\bar{\mathbf{l}}_o|| \cdot ||\bar{\mathbf{l}}_{o'}||}\right) \in [0, 1]$$

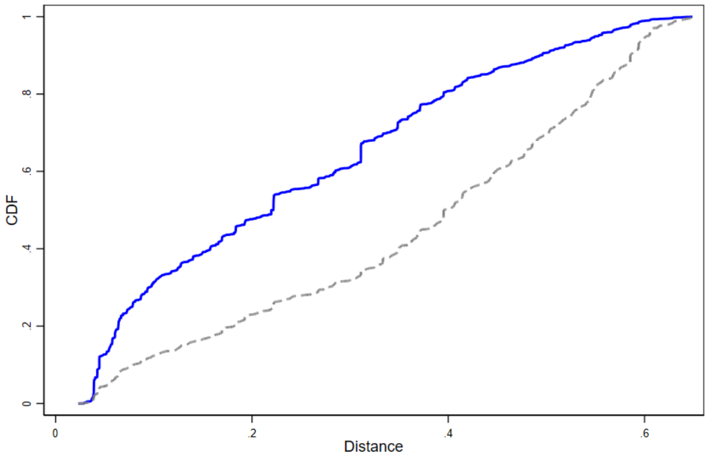
- **Implementation:** BIBB longitudinal microdata
  - 15 harmonized tasks across survey waves
  - measure $\bar{l}_{o\tau}$: share of individuals belonging to o performing τ
  - construct matrix of bilateral distances for each wave, then take *average* of $\varphi(\bar{\mathbf{l}}o, \bar{\mathbf{l}}o')$ across waves for each (o, o')
  - *note:* I thus hold the distance in task space between any two occupations fixed over time – letting it vary across time: conceptually possible, but scale difficult to compare across

Methodology (2): occupational movers

- Consider individual i who in period t is employed at j in occupation o. I consider i an *occupational mover* if in $t+1$, i is employed at $j' \neq j$ and $o' \neq o$.
    - only counted as occ. mover if employed in different job: cf. Kambourov and Manovskii (2008)
    - considering switch in adjacent periods: more likely that it is a *voluntary* move
- Function $\varphi(\bar{\mathbf{l}}_o, \bar{\mathbf{l}}_{o'})$ maps each $o \to o'$ move onto $[0, 1]$
- Implementation: merge $\{\varphi(\bar{\mathbf{l}}_o, \bar{\mathbf{l}}_{o'})\}_{oo'}$ into SIEED
    - harmonized occupational classification KldB1988-2d
    - restrict myself to 1985-2009, because subsequent stark change in occupational classification limits comparability (missing notifications, etc.)

Comparison of realized movements in task space vs. implied by random mobility

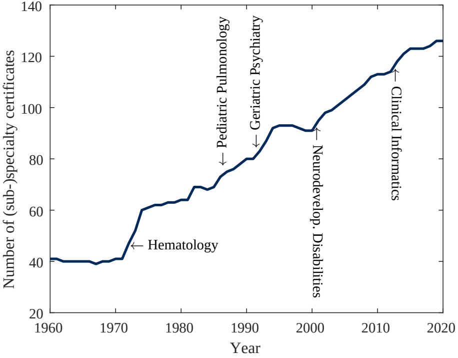
Back



Examples: rising specialization

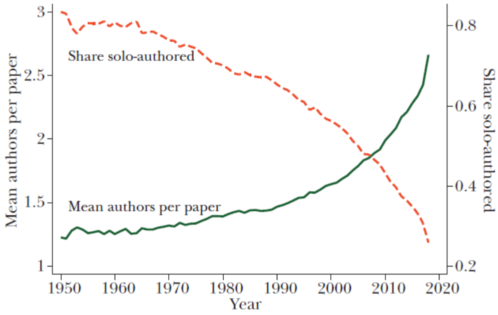
Main

- **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]*



Data sources: short description of main datasets

▸ Main ▸ Imputation procedure

- **Germany**: SIEED linked employer-employee dataset
  - *establishment* and individual data generated in administrative processes
  - built up from a 1.5% sample of all 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
  - top coding (affects >50% of university-educated men in regular full-time employment) $\rightarrow$ adopt standard imputation methods (Dustmann et al., 2009; CHK, 2013)
  - Much larger sampling frame than more familiar LIAB
- **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)

Sample restrictions

Main

- Data cleaning $\Rightarrow$ broadly harmonized samples
- Main restrictions
  - age 20-60
  - full-time employed
  - drop agriculture, public sector, utilities industries
  - firms (and their employees) with at least 10 employees
- DE: West Germany
- Transform SIEED spell-level data into annual panel
- PRT: at least. official minimum wage

Wage Imputation procedure

- 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 $\times$ 99th percentile.

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 coworker complementarities arise precisely when workers are *differentiated* in their task-specific productivities
- **Averaging step:**
  - equally-weighted averaging ignores non-linearity in coworker aggregation implied by the structural model
  - in 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

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 distribution estimated from CPS data for the U.S.
  Recover $\gamma = 0.84$ for 2006 but *very* noisy estimates
- **Ongoing work:** use the extended microfoundation to identify $\chi \Rightarrow$ variation in output – and, hence, wages – conditional on workforce size and talent composition

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 at least qualitatively consistent with my "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
- John's operationalization of worker types and tasks does *not* directly map onto my model
  - 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 what I found from structural estimation using evidence from wage spillovers!**

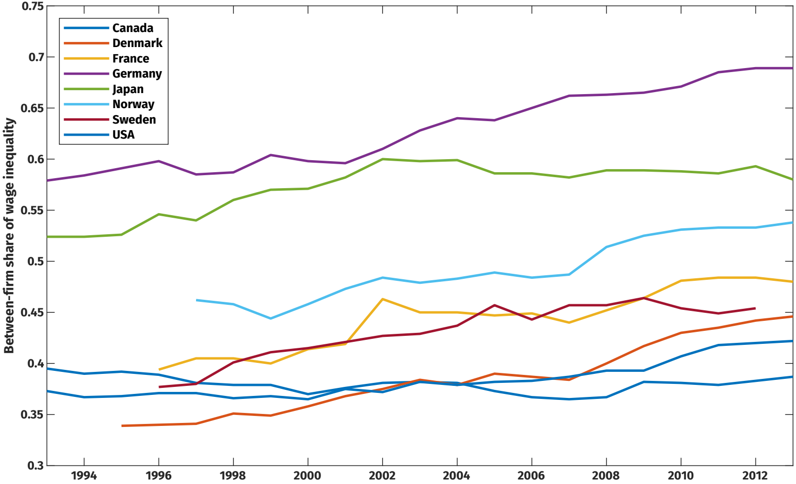
Semi-structural back-of-envelope calculation for γ

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 $i \neq j$, where subscripts denote partial derivatives
- 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!

Firming up inequality: cross-country evidence

Main

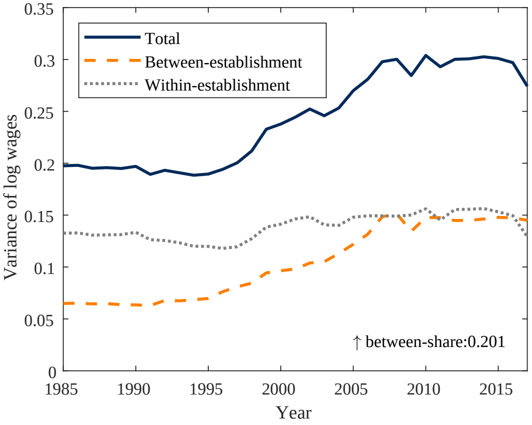


Notes. Data from Tomaskovic-Devey et al. (2020). Measures of earnings differ across countries and, for Germany, between T-D et al. and my study based on the SIEED.

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

Main

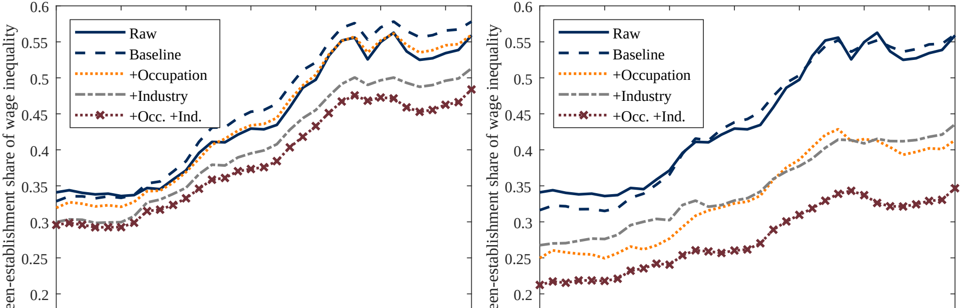
- Instead of considering *all* employers, restrict attention to “panel establishments”



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

▸ Main

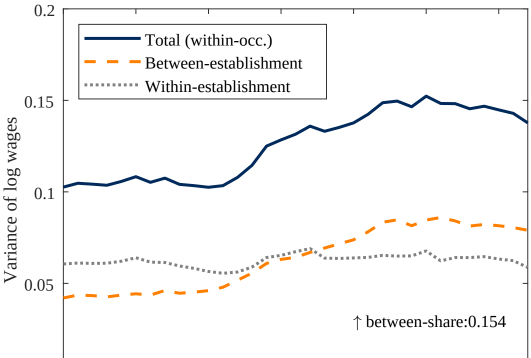
- "With worker FEs": regress $\ln \tilde{w}_{it} = \alpha_i + X_{it}'\hat{\beta} + \epsilon_{it}$, and construct residuals $\ln w_{it} = \ln(\tilde{w}_{it} - X_{it}'\hat{\beta})$.
- "Without worker FEs": regress $\ln \tilde{w}_{it} = \alpha_0 + X_{it}'\hat{\beta} + \epsilon_{it}$, and consider residuals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



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

▸ Main

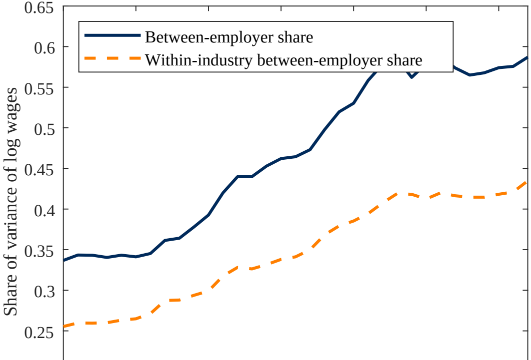
- Decomposition: between-occupation, within-occupation-within-employer, within-occupation-between employer ⇒ within-occupation-between employer/(within-occupation-within-employer + within-occupation-between-employer)



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

Main

- Decomposition: between-industry ($\rightarrow$ between-employer), within-industry-within-employer, within-industry-between-employer $\Rightarrow$ compare "total" between-employer share (baseline) & within-industry between-employer share



Evolution of coworker sorting: correlation coefficient

Main

- Estimate $\hat{x}_i$ and $\hat{x}_{-it}$ separately for 5 periods, compute coworker sorting correlation coefficient
- In addition to the baseline, also consider ranking within-occupation

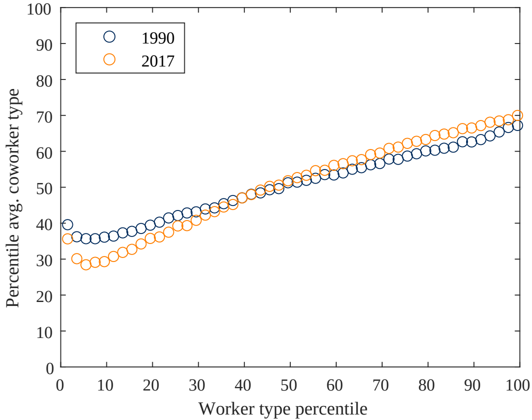
	Sorting	
Period	Spec. 1	Spec. 2
1985-1992	0.427	0.423
1993-1997	0.458	0.443
1998-2003	0.495	0.452
2004-2009	0.547	0.470
2010-2017	0.617	0.519

Notes. The column labelled "Sorting" indicates the correlation between a worker's estimated type and that of their average coworker, separately for five sample periods. Under "Spec. 1" workers are ranked economy wide

Evolution of coworker sorting: within-occupation ranking binscatter

Main

- Reproduce non-linear sorting plot, but now $\hat{x}_i$ is based on *within-occupation* ranking



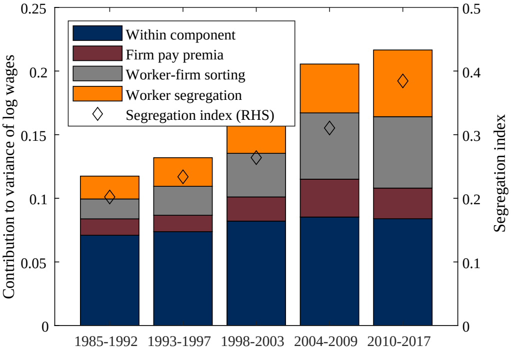
AKM-based wage variance decomposition

Main

- AKM-based var. decomp. *[Song et al., 2019]*

$$\mathrm{Var}(w_{it}) = \underbrace{\mathrm{Var}(\alpha_i - \bar{\alpha}_{j(i,t)}) + \mathrm{Var}(\epsilon_{i,j})}_{\text{within-component}}$$
$$+ \underbrace{\mathrm{Var}(\psi_{j(it)}) + 2\mathrm{Cov}(\bar{\alpha}_{j(it)}, \psi_{j(it)}) + \mathrm{Var}(\bar{\alpha}_{j(it)})}_{\text{between-component}}$$

  - $\mathrm{Var}(\psi_j)$: **firm-specific pay premia**
  - $\mathrm{Cov}(\bar{\alpha}_j, \psi_j)$: **(worker-firm) sorting**
  - $\mathrm{Var}(\bar{\alpha}_j)$: **(worker-worker) segregation**

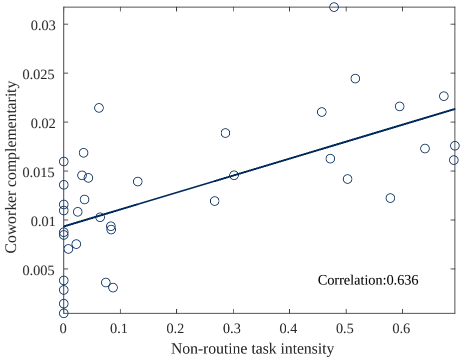


- Segregation index *[Kremer-Maskin, 1996]*:

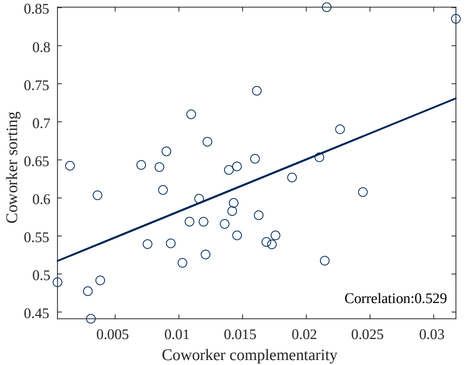
Occupations: task complexity ⇒ complementarity ⇒ sorting

▸ PRT data + industry evidence

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



- ↑ Coworker wage complementarity ⇒ ↑ coworker sorting

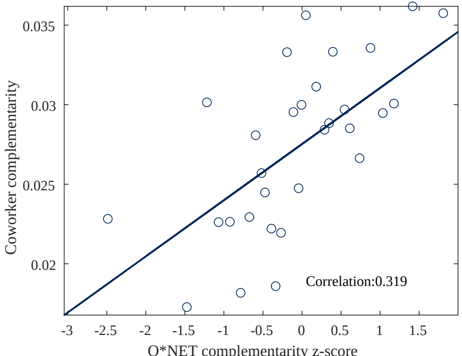


Notes. Quadros de Pessoal microdata. Horizontal axis indicates occupation's reliance

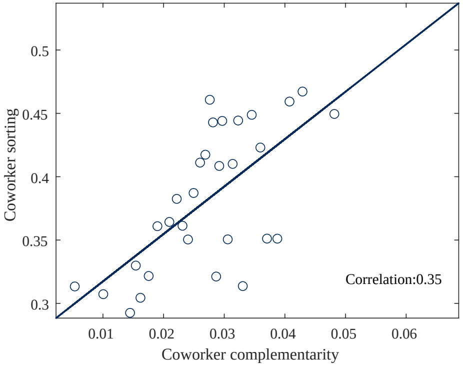
Industries: coworker importance ⇒ complementarity ⇒ sorting

▸ Main

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



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

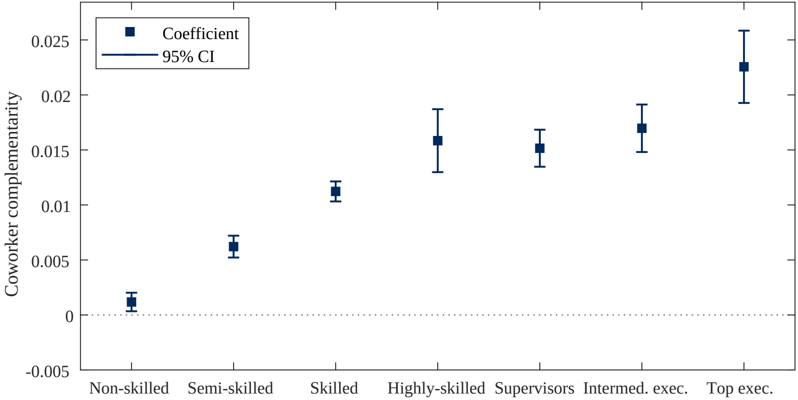


Notes: NACE 4 digit industries

Hierarchies: complexity ⇒ complementarities

▸ Main ▸ Occupations

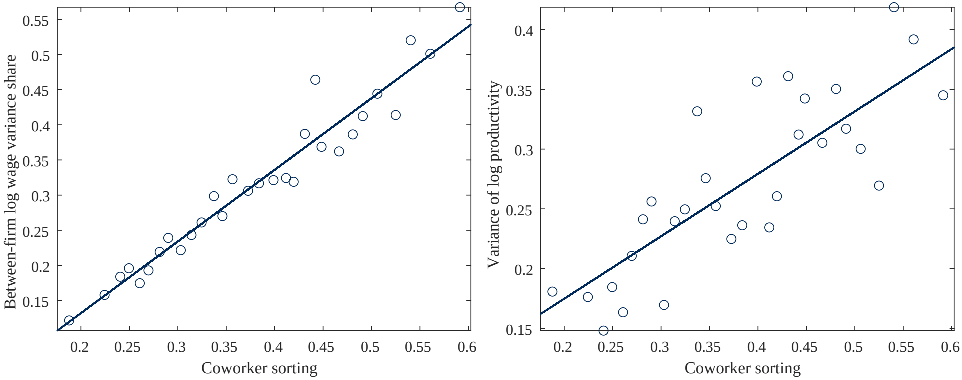
⇒ Coworker wage complementarities are (weakly) monotonically ↑ in the layer of a firm's hierarchy



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.



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

Cluster-based methodology: implementation

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

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

  - $k(1), \ldots, k(J)$: 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
  - $H_1, \ldots., H_K$: generic cdf's

- Implementation here:
  - baseline value of $K = 10$, as in BLM, but experiment with $K = 20$ and $K = 100$
  - use firms' cdf's over entire sample period on a grid of 20 percentiles
- "Half-BLM": take step (1), impute class to each worker-year observation, then estimate two-way fixed effect wage regression using cluster effects instead of firm effects:

Robustness checks: complementarity regression

Main

- Types from non-parametric ranking algorithm instead of AKM-based Jump
- Non-parametric, FD approximation Jump
- Schooling as a non-wage measure of types Jump
- Lagged types Jump

Complementarity estimates using years of schooling

Main

	Dependent variable: wage				
	'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
Fixed effects	Yes	Yes	Yes	Yes	Yes
Obs. (100,000s)	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 are given 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 100,000s.

Sorting & complementarity based on non-parametric ranking algorithm

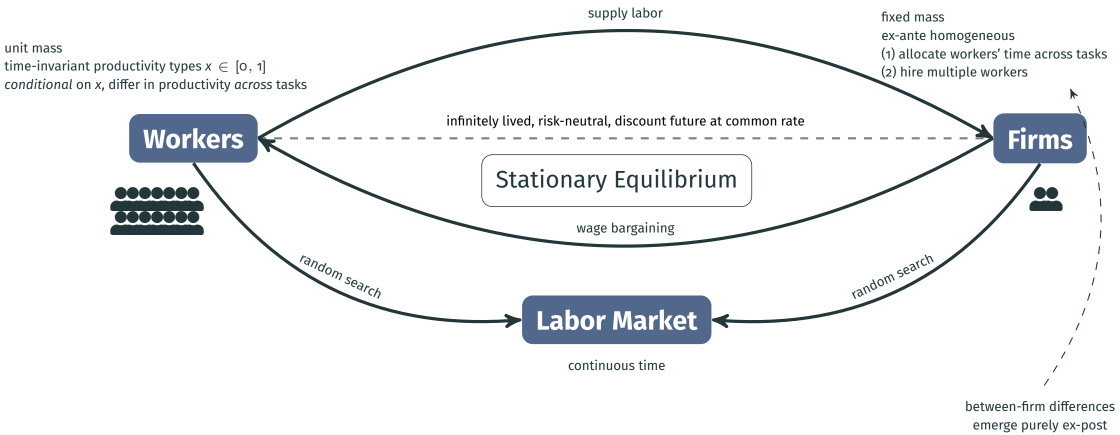
▸ Main

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

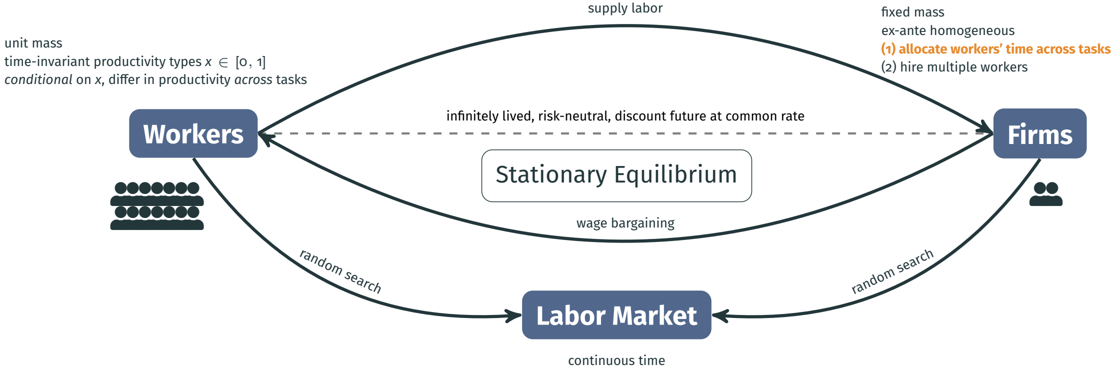
	Sorting		Complementarities	
Period	Spec. 1	Spec. 2	Spec. 1	Spec. 2
1985-1992	0.47	0.38	0.001	0.000
1993-1997	0.56	0.46	0.002	0.001
1998-2003	0.60	0.48	0.004	0.002
2004-2009	0.65	0.50	0.005	0.002
2010-2017	0.68	0.51	0.005	0.004

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,

Overview of model environment: firm organization meets labor search



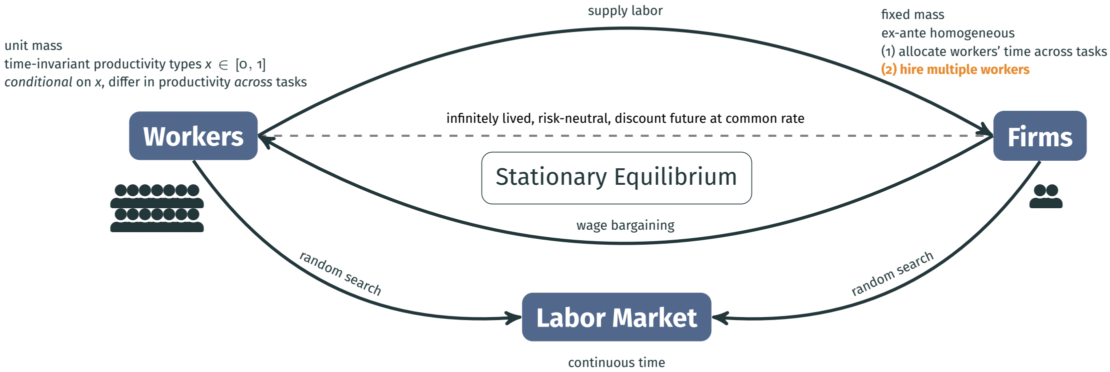
Overview of model environment: firm organization meets labor search



- **Analyze in 2 steps**

  1. **task assignment: derive production function for *one* firm, treating workforce as exogenous**

Overview of model environment: firm organization meets labor search



- **Analyze in 2 steps**
  1. task assignment: derive production function for *one* firm, treating workforce as exogenous

The firm as team assembly technology: organizational optimization problem

Main

- **Role of firms**: facilitate division of labor & **coordinate**
- **Choose** $\{q(\tau)\}_{\tau\in\mathcal{T}}$ & $\{\{y_i(\tau)\}_{\tau\in\mathcal{T}}\}_{i=1}^n$ & $\{\{l_i(\tau)\}_{\tau\in\mathcal{T}}\}_{i=1}^n$ **to max.** Y

$$\mathcal{L}(\cdot) = Y + \lambda\underbrace{\left[\left(\int_{\mathcal{T}} \ln q(\tau)d\tau\right) - Y\right]}_{\text{task aggregation}}$$

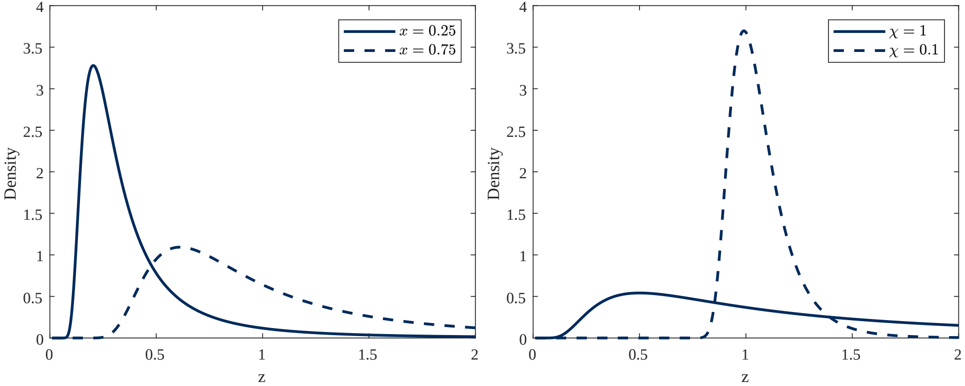
$$+\sum_{i=1}^{n}\left\{\int_{\mathcal{T}}\tilde{\lambda}(\tau)\underbrace{\left(\sum_{i=1}^{n} y_i(\tau) - q(\tau)\right)}_{\text{division of labor}}d\tau + \int_{\mathcal{T}}\lambda_i(\tau)\underbrace{\left(z_i(\tau)l_i(\tau) - y_i(\tau)\right)}_{\text{task production}}d\tau + \lambda_i^L\underbrace{\left(1 - \int_{\mathcal{T}} l_i(\tau)d\tau\right)}_{\text{time constraint}}\right\}$$

$+$ non-negativity constraints

- **Solve** mini-planner problem using insight from trade literature

Illustration: Fréchet distribution

Main



(a) Scale parameter

(b) (Inverse) shape parameter

Figure 1: Illustration of properties of Fréchet distribution

Taylor approximation to CES

▸ Back to team production ▸ Back to stylized model

- 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

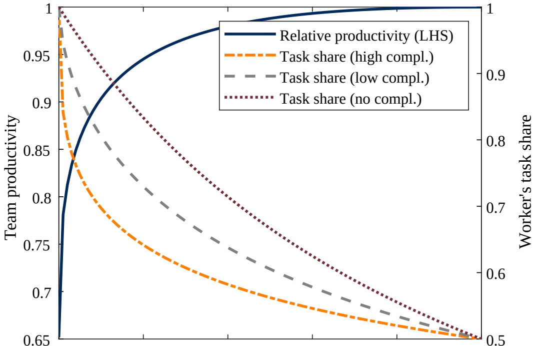
$$\bar{x} - \frac{1}{2}\underbrace{(1-\gamma)}_{\approx\xi}\frac{\sigma_x^2}{\bar{x}},$$

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

Quality mismatch & task mismatch

Main

- When the high-quality worker is paired with a low-quality worker, she ends up "wasting time" on tasks that she is relatively less efficient in and wouldn't have to do if teamed up optimally.



Extension: team production with communication costs

Main

- Assumption till now: division of labor incurs 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: rich microdata from Fortune-100 company to describe communication behavior

Extension: specialization vs. comparative advantage

Main

- Suppose the joint distribution of task-specific productivities across coworkers satisfies

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

  where $\tilde{\chi}$ is the common shape parameter of the the marginal Fréchet distributions, while $\xi \in (0, 1]$ controls correlation in draws across workers

- Can derive

$$Y = f(\mathbf{x}; \tilde{\chi}, \xi) = n^{1+\tilde{\chi}\xi}\left(\frac{1}{n}\sum_{i=1}^{n}(a_1 x_i)^{\frac{1}{\chi\xi+1}}\right)^{\tilde{\chi}\xi+1} \tag{2}$$

- All results go through with $\chi = \tilde{\chi}\xi$

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 – ex-ante homogeneous
  - 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' denote *coworker*
  - team production: $f(x, x') \leftarrow$ see microfoundation

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 λ_u
  - probability for a firm to be contacted by an(y) unemployed: $\lambda_{v.u} = \lambda_u \times u$
  - baseline: no on-the-job search
- **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

Key hiring decision

Main

- **Notation:**
  - Value functions for unemployed worker, $V_u(x)$; worker x employed with coworker x', $V_e(x|x')$
  - Value functions for vacant firm, $V_{f.0}$; firm producing with x and x': $V_f(x, x')$
  - $d_u(x)$: 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 with worker x, $\Omega(x)$

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

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

Environment: firm & worker states

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

$$d_w(x) = d_u(x) + d_{m.1}(x) + \int d_{m.2}(x, \tilde{x}')d\tilde{x}'$$

  - $d_u(x)$: 'density' of unemployed of type x
  - $d_m(x)$, shorthand for $d_m(x, \emptyset)$: 'density' of matches w/ x as only worker
  - $d_m(x, x')$: 'density' of "team matches" b/w x and x'

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

$$d_f = d_{f.0} + \int d_{m.1}(x)dx + \frac{1}{2}\int\int d_{m.2}(x, x')dxdx'.$$

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

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

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 continuously renegotiated

- **One-worker firm:**

$$(1-\omega)\big(V_{e.1}(x) - V_u(x)\big) = \omega\big(V_{f.1}(x) - V_{f.0}\big) \tag{3}$$

Two-worker firm

$$(1-\omega)\big(V_{e.2}(x|x') - V_u(x)\big) = \omega\big(V_{e.2}(x'|x) + V_{f.2}(x,x') - V_{e.1}(x') - V_{f.1}(x')\big) \tag{4}$$

HJB: unmatched

Main

- Unmatched firm:

$$\rho V_{f.o} = (1-\omega)\lambda_{v.u} \int \frac{d_u(x)}{u} S(x)^+ dx \tag{5}$$

- Unmatched worker:

$$\rho V_u(x) = b(x) + \lambda_u \omega \left[\int \frac{d_{f.o}}{v} S(x)^+ + \int \frac{d_{m.1}(\tilde{x}')}{v} S(x|\tilde{x}')^+ d\tilde{x}' \right] \tag{6}$$

HJB: surpluses

- Surplus of coalition of firm with worker x

$$(\rho+\delta)S(x)=f_1(x)-\rho\big(V_u(x)+V_{f.0}\big)+\lambda_{v.u}(1-\omega)\int\frac{d_u(\tilde{x}')}{u}S(\tilde{x}'|x)^+d\tilde{x}' \tag{7}$$

- Surplus from adding x to x'

$$S(x|x')\big(\rho+2\delta\big)=f_2(x,x')-\rho\big(V_u(x)+V_u(x')+V_{f.0}\big)+\delta S(x)-(\rho+\delta)S(x') \tag{8}$$

KFE: unemployed

$$\delta\left(d_{m.1}(x) + \int d_{m.2}(x, \tilde{x}')d\tilde{x}'\right) = d_u(x)\lambda_u\left(\int \frac{d_{f.0}}{v} h(x, \tilde{y}) + \int \frac{d_{m.2}(\tilde{x}')}{v} h(x|\tilde{x}')d\tilde{x}'\right). \tag{9}$$

KFE: one-worker matches

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

KFE: two-worker matches

$$2\delta d_{m.2}(x, x') = d_u(x)\lambda_u \frac{d_{m.1}(x')}{v} h(x|x') + d_u(x')\lambda_u \frac{d_{m.1}(x)}{v} h(x'|x). \tag{11}$$

Lemma: monotonicity of unemployment value and wage in x

Main

Lemma

Assume that $\frac{f_1(x)}{\partial x} > 0$, $\frac{\partial f_2(x,x')}{\partial x} > 0$, and $\omega > 0$. Then: (i) The value of unemployment $V_u(x)$ is monotonically increasing in x, and (ii) so is the wage function $w(x|x')$.

Proof.

See paper appendix. Key: surplus representation.

Wage equation

▸ Main

- NB: here allow for ex-ante firm heterogeneity $\Rightarrow$ measurement result extends

$$\begin{aligned} w(x|y,x') = {} & \rho V_u(x) + \omega\Big[f(x,y,x') - \rho\left(V_u(x) + V_u(x') + V_{f.o}(y)\right) \\ & + \delta S(x|y) - (\rho+\delta)S(x'|y)\Big] - \delta\omega S(x|y) \\ = {} & \omega\big(f(x,y,x') - f(x',y)\big) + (1-\omega)\rho V_u(x) \\ & - \omega(1-\omega)\lambda_{v.u}\int \frac{d_u(\tilde{x}'')}{u} S(\tilde{x}''|y,x')^{+} d\tilde{x}''. \end{aligned}$$

Characterization using a stylized model: setup

Main

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]

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 $x \in \mathcal{X} \rightarrow$ 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: 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*

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

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

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

$$\mu(x) = x$$

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

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

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

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

Parameterization, including estimation results

Main

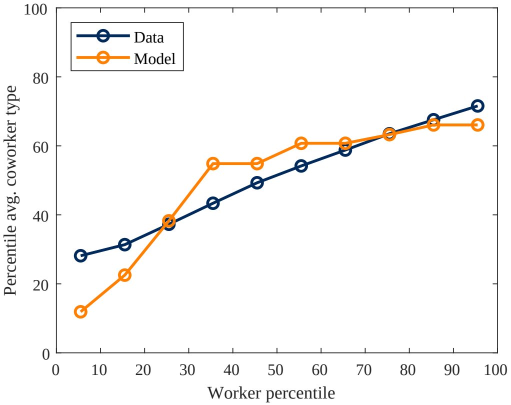
Parameter	Description	Targeted moment	Value	m	$\hat{m}$
γ	Elasticity of complementarity	$\hat{\beta}_c$	**0.837**	0.0091	0.0091
a_0	Production, constant	Avg. wage (norm.)	0.239	1	1
a_1	Production, scale	Var. log wage	1.557	0.241	0.241
b_1	Replacement rate, scale	Replacement rate	0.664	0.63	0.63
δ	Separation hazard	Job loss rate	0.008	0.008	0.008
λ_u	Meeting hazard	Job finding rate	0.230	0.162	0.162
ρ	Discount rate	External	0.008		
ω	Worker bargaining weight	External	0.50		
a_2	Production, team advantage	External	1.10		

Notes. This table lists for each of the estimated parameters, the targeted moment, the estimated value, and

Validation: model reproduces empirical coworker sorting patterns

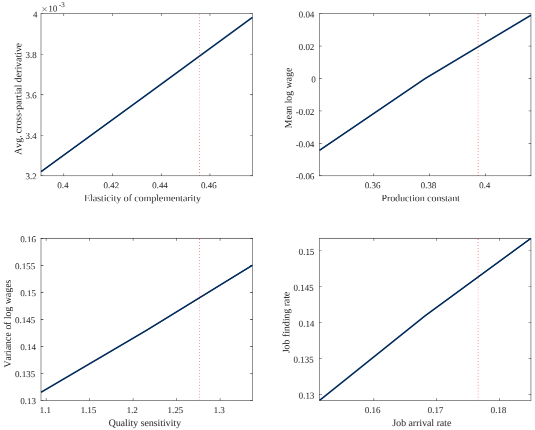
Main

- Key untargeted moment (1): **coworker sorting patterns**
- Coworker correlation matches data well, $\rho_{xx} = 0.53$ (vs. 0.62 in the data)
- Model slightly underestimates the quality of coworkers at both bottom and top
    - OJS will help



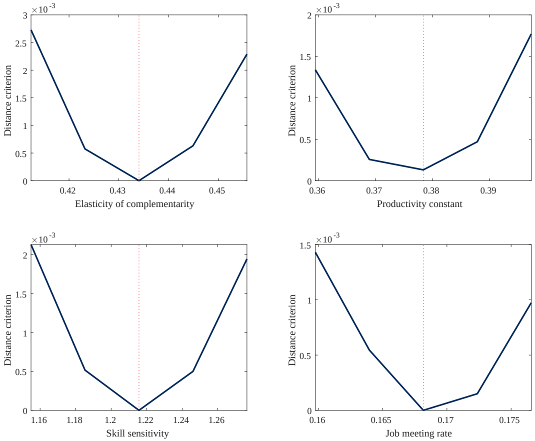
Identification validation exercise 1

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 $\mathcal{G}$.

Between-share adjustment procedure (1)

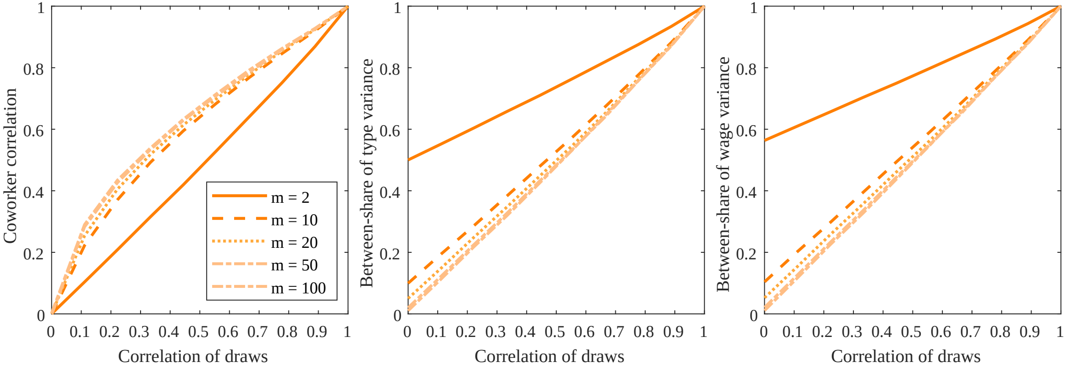
Main

- **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.

Overview of extensions, robustness checks, and other implications Main

- **Extensions and robustness checks**
  - Declining search frictions ▸ Jump
  - Outsourcing & within-occupation analysis ▸ Jump
  - OJS: theoretical & empirical EE patterns ▸ Jump
  - Larger firm size ▸ Jump
- **Implications for overall inequality:** needn't rise due to coworker complementarity ↑ ▸ Jump
- **Implications for aggregate productivity:** reallocation limited rise in mismatch costs ▸ Jump

The effect of declining search frictions

Main

- **Model-based concern:** reduction in search frictions could also explain ↑ coworker sorting
- Yes: job arrival & separation rates estimated to ↑ from p1 to p2
- **Counterfactual analysis:** explains 6% of empirically observed ↑ 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 complementarity	0.076	61.47

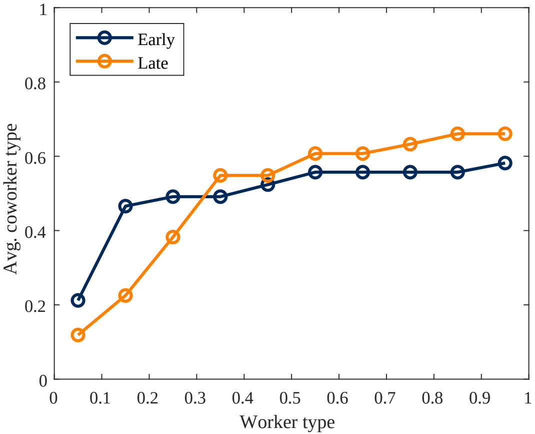
Extension: model with OJS

Main

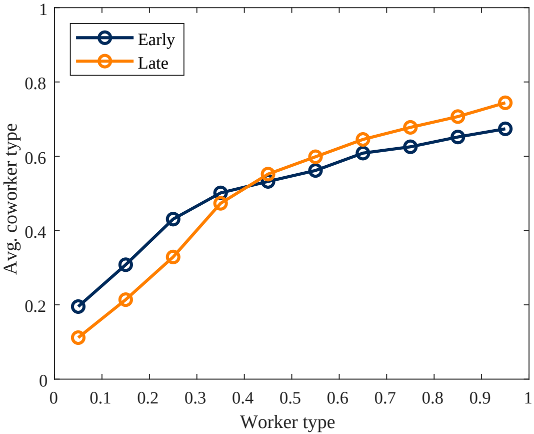
- Baseline model abstracted from OJS
  - transparent trade-off, connection to closed-form matching results
- **Consider extension to OJS:** employed worker meet vacancies at Poisson rate λ_e
  - wages both off and on the job are continuously renegotiated under Nash bargaining, with unemployment serving 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

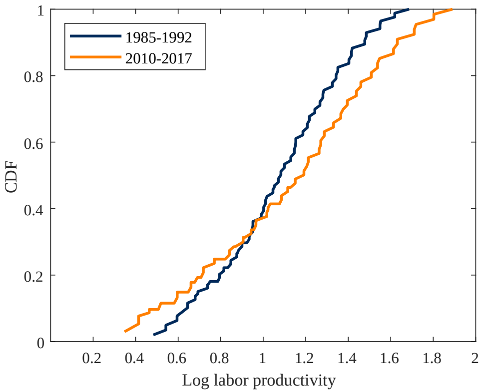
- **Theoretical prediction:** EE transitions move workers in surplus-maximizing direction $\Rightarrow \Delta\hat{x}_{-it} = \hat{x}_{-i,t} - \hat{x}_{-i,t-1}$ should be *positively* correlated with $\hat{x}_i$
  - $h_{2.1}(x, x''|x') = 1$ – worker x in a two-worker firm with coworker x'' would move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

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

Productivity dispersion

Main

- Firm dynamics literature: increased firm-level productivity dispersion *[Autor et al., 2020; de Ridder, 2023]*



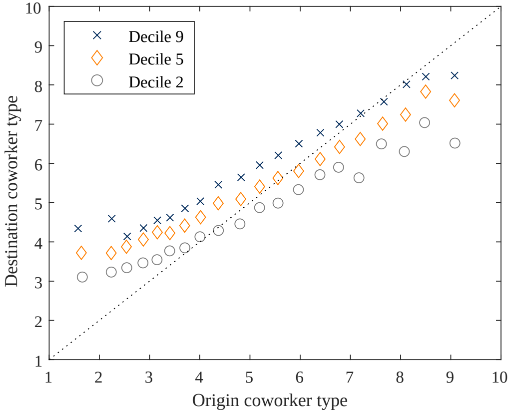
Empirical coworker sorting changes due to EE moves Main

- Non-parametric perspective:
  origin-destination plot by worker type

  $\Rightarrow$ **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* empirically EE transitions "move up" low types more than predicted by theory ***[Add figure]***

  $\Rightarrow$ "**Coworker job ladder**" with both absolute and type-specific dimension? Learning?



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

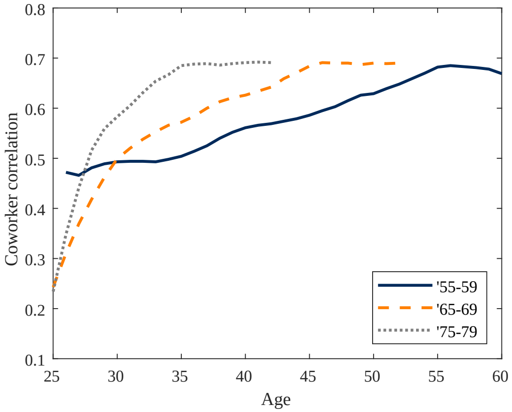
	Data		Model	
Change in coworker type	'85-'92	'10-'17	Period-1	Period-2
Own type	**0.0883*** **	**0.118*** **	**0.214**	**0.270**
	(0.000799)	(0.000918)		
Controls	Year FEs, Origin	Year FEs, Origin	Origin	Origin
N	196,098	282,718	∞	∞
adj. R^2	0.284	0.204		

Table 1: Change in coworker type due to EE moves is 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.

Related: life-cycle patterns of sorting consistent with cohort effects

- Coworker sorting increases over life-cycle
- Coworker sorting higher in recent cohorts
    - consistent with each cohort becoming more specialized
    - but also with improvement in search frictions over time, which primarily affects sorting of young people
- Currently don't have life-cycle effects in my model, but could be incorporated quite straightforwardly



Notes. Figure displays the correlation of own type with average coworker type, by age, for three separate cohorts.

The implications of "horizontal" complementarity

Main

- **Baseline:** only "vertical" complementarity *due to* horizontal differentiation
- **Concern:** but couldn't it better to match a high-type (A) with a low-type (L) coworker than a high-type coworker (H) if A & L are specialized in *different* tasks while A & H are specialized in the *same* tasks?
  - through lens of model: correlation between $\mathbf{z}_A$ and $\mathbf{z}_L$ or $\mathbf{z}_H$ vs. baseline: assume *iid* draws
  - ofc correlation could be *endogenous:* if task-specific skills "persistent," they may be observable upon matching ("inspection") or afterwards ("experience") $\rightarrow$ shape formation or separation probabilities
- **Important & requires careful consideration**
  - conceptual/theoretical: 3 cases (no specialization; baseline with iid draws; case with correlated draws) $\Rightarrow$ conditional on correlation, $\chi \uparrow$ leads to increase 'vertical' complementarity

Match-specific shocks

- Suppose that $f(x, x', \zeta) = \zeta f(x, x')$, where $\ln \zeta \sim G$
  - in progress: micro-foundation for ζ through multi-variate Fréchet: interpretation in terms of negative correlation of $\mathbf{z}$ vectors
- Sketch of solution approach to extended model
  - define

$$h(x|x') = 1 - G\left(\bar{\zeta}(x|x')\right)$$

$$p(k) = 1 - G(k)$$

$$\zeta^*(k) = \frac{\int_k^\infty \zeta dG(\zeta)}{1 - G(k)} \text{ if } G(k) < 1 \text{ and } \zeta^*(k) = k \text{ otherwise}$$

where $\bar{\zeta}(x|x')$ solves $S(x|x', \zeta) = 0$

  - then, e.g., value of unemployment for type x satisfies

$$\rho V_u(x) = b(x) + \lambda_u \omega \left[\int \frac{d_{f.0}}{v} S(x)^+ + \int \frac{d_{m.1}(\tilde{x}')}{v} p\left(\bar{z}\left(x|\tilde{x}'\right)\right) \times S\left(x|\tilde{x}', \zeta^*\left(\bar{\zeta}\left(x|\tilde{x}'\right)\right)\right) d\tilde{x}' \right]$$

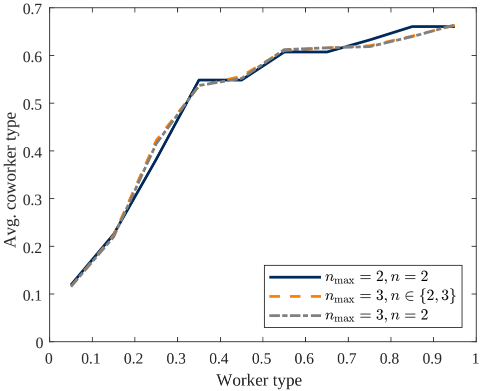
Match-specific shocks

- **Quantitative** evaluation: *in progress*
- **Initial findings:**
    - much better fit to coworker sorting patterns
    - much better fit to wage functions
    - precise calibration of σ_ζ not easy: within-(x, x') wage dispersion captures also e.g. measurement error
    - main prediction of $\uparrow$ between-share as $\gamma \uparrow$: *robust*

Extension to $n_{max=3}$

Main

- Baseline model imposes $n_{\max} = 2$ for reasons of (i) tractability and (ii) transparency
- Can extend to $n_{\max} = 3$ (or $n_{\max} = 4$) & find that results are robust



Productivity costs of complementarity & labor market functioning

▸ Main

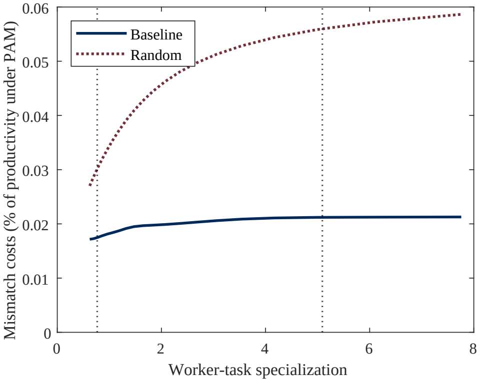
"The benefits of the division of labor are limited by the functioning of the labor market"

- 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, the 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])



Implications for overall inequality?

▸ Main

- **Coworker complementarities do not necessarily $\uparrow$ variance of person-level wages**
  - (un-)surprising? AKM-variance decomposition perspective vs. common intuition/question
  - counterfactual: variance of log wages 0.2166 in 2010 under 1990-complementarities vs. 0.210
  - mechanisms: (i) reallocation effect, (ii) valuation effect, (iii) outside option effect
- Importance of within-firm wage compression: if high, then $\uparrow$ coworker sorting also pushes up *overall* inequality (currently conjecture only!)
- Other caveats: coworker learning *[Jarosch-Oberfield-RossiHansberg, 2021]*, increasing returns to labor quality *[Kremer, 1993]*, sharing in monopoly rents, ... $\Rightarrow$ future research
- Not straightforward to analyze formally: alternative bargaining setup with coalition

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, protect-	

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 $\bar{T}^{\text{abstract}}$ in pooled 2012 & 2018 waves.

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