Submitted to *Econometrica*

SUPERSTAR TEAMS

LUKAS B. FREUND
Columbia University

Production increasingly requires specialized expertise. To study the macroeconomic implications, I develop a tractable theory in which firms assemble teams of workers with heterogeneous task-specific skills. Deriving the firm's production function from optimal task assignment shows that output is maximized when coworkers excel at different tasks yet possess similar overall talent. Crucially, greater skill specificity, while raising potential productivity, endogenously amplifies talent complementarities, i.e., the productivity loss from talent mismatch. This promotes talent concentration into select firms with "superstar teams," though search frictions prevent perfect sorting. Using German panel micro data, I document industry patterns consistent with this mechanism and calibrate the model. The quantified model shows that, first, growing skill specificity since the mid-1980s has amplified sorting, explaining a significant share of the widely documented "firming up of inequality". Second, "Smithian" productivity gains from specialization are muted when labor market frictions impede the matching of coworkers with complementary expertise.

KEYWORDS: firms, inequality, productivity, specialization, teams.

Lukas B. Freund: lukas.beat.freund@gmail.com

I am grateful to Vasco Carvalho for his invaluable guidance and support. In addition, I thank Zsófia Báráy (discussant), Maarten De Ridder, Wouter Den Haan, Luis Garicano, Kyle Herkenhoff, Tomer Ifergane, Guido Menzio, Simon Mongey, Christian Moser, Ezra Oberfield (discussant), Tommaso Porzio, Pascual Restrepo, Richard Rogerson, Esteban Rossi-Hansberg, Bastian Schulz (discussant), and Gianluca Violante for helpful conversations. I also thank participants at numerous seminars and at the Annual Meeting of the CEPR Macroeconomics and Growth Programme 2023, BSE Summer Forum 2024, Bristol Macroeconomics Workshop 2024, EES Workshop on 'New Developments in the Macroeconomics of Labor Markets' 2024, Minneapolis Fed Junior Scholar Conference 2023, Philadelphia Workshop on Macroeconomics 2025 (Fed/Penn/Wharton), NBER SI 2024 (Macro Perspectives), NBER Organizational Economics Workshop Fall 2024, and SED 2024 for comments. This study uses the Sample of Integrated Employer-Employee Data (SIEED 7518) and the Linked Employer Employee Data longitudinal model (LIAB LM7519) from the German Institute for Employment Research (IAB). Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the IAB and subsequently remote data access.

1. INTRODUCTION

Production is often too complex for any individual to master all required tasks. Indeed, specialization appears to have deepened over recent decades. For example, the number of distinct specialty certificates issued by the American Board of Medical Specialties nearly doubled between 1980 and 2020. This pattern extends to the wider economy: routine tasks have increasingly given way to those requiring specialized expertise. Economic development itself is widely understood to involve deepening complexity and specialization.[1]

A classic literature argues that specialization shapes firm productivity – and thus wages – through the division of labor, with firms playing a central role in organizing production with heterogeneously skilled, specialized workers: "Team production" allows output to exceed the separable sum of individual contributions.[2] Yet, how this aggregates to macroeconomic outcomes – how specialization and its rise affect overall productivity and inequality – remains poorly understood. Classic theories are typically qualitative or microeconomic, and modern macro models usually abstract from specialization, assuming workers differ only in efficiency units or positing production functions additively separable in labor.[3]

To bridge this gap, I develop a model of firms assembling teams of workers with heterogeneous skills across tasks that is sufficiently tractable for empirical and macroeconomic analysis. The core innovation is a microfounded firm-level production function whose shape depends endogenously on skill specificity, i.e., the degree of dispersion in task-specific skills. It reveals that greater specificity both raises potential productivity and makes workers less substitutable, and thus output more vulnerable to deviations from optimal team composition, where coworkers excel at different tasks yet possess similar overall talent. These talent complementarities promote concentration of talent into select firms with "superstar teams," though mismatch can persist in equilibrium due to search costs.

I use the model to quantitatively analyze the macro consequences of rising specialization. First, regarding labor market inequality, the model calibrated to German micro data

[1]See, e.g., Acemoglu and Restrepo (2018), Atalay et al. (2020), Bandiera et al. (2022), Autor et al. (2024).

[2]The description of team production follows Alchian and Demsetz (1972). Other seminal references include Arrow (1974), Rosen (1978), Becker and Murphy (1992), Kremer (1993) and Garicano (2000).

[3]Recent advances focus on either task-based production (Acemoglu and Restrepo, 2018), coworker learning (Jarosch et al., 2021, Herkenhoff et al., 2024), or multi-dimensional skills (Lindenlaub, 2017).

shows thta growing skill specificity since the mid-1980s has amplified complementarities and thereby promoted talent sorting, explaining a significant share of the much-discussed "firming up" of wage inequality (Song et al., 2019), i.e., growing between-firm wage dispersion. Second, regarding productivity, because vulnerability to mismatch rises endogenously with specialization, labor market frictions – especially prevalent in developing economies (Donovan et al., 2023) – can substantially curtail "Smithian" growth.

Theory. The theory has four core features. First, production requires a continuum of imperfectly substitutable tasks. Second, workers have limited time and vary in their task-specific skills. Third, production occurs in teams. Ex-ante identical firms hire multiple workers – as in Herkenhoff et al. (2024), a team comprises two workers – and assign them tasks. Fourth, production is embedded in a labor market with information frictions modeled as random search: Firms sequentially encounter unmatched individuals, observe their previously unknown skills, and decide whether to hire or continue searching.

The key analytical step is to microfound a tractable, firm-level production function by deriving it from the optimal task assignment problem. Tractability comes from modeling workers' task-specific skills as realizations of correlated Fréchet distributions, in the spirit of trade models à la Eaton and Kortum (2002): Individual *talent* governs expected skill; an economy-wide *skill specificity* parameter controls within-worker dispersion of skill across tasks; and the *horizontal distance* between any two workers captures dissimilarity in task specialization. Under optimal assignment by comparative advantage, aggregation across tasks yields a reduced-form production function where talents and horizontal distance are sufficient statistics for output – facilitating integration into a macro model.

This microfoundation reveals that skill specificity amplifies both potential productivity and the cost of deviations from optimal team composition ("coworker mismatch"). Optimal composition requires coworkers who are horizontally distant but close in talent. While the benefits of horizontal distance are straightforward, what is the intuition for talent complementarities?[4] A worker must perform a larger share of tasks when their coworker is less talented. When skills are one-dimensional, realized productivity is unaffected; but skill specificity magnifies the opportunity cost of handling tasks a worker is relatively slow at.

[4]Technically, "coworker talent complementarity" (or super-modularity) refers to a positive cross-partial derivative of the production function with respect to two distinct coworkers' talent levels.

These losses are larger in absolute magnitude for more talented workers, as their potential output is higher. The theory formalizes this intuition. Output is the product of two terms: the first captures efficiency gains from specialization under team production; the second is a constant-elasticity-of-substitution (CES) function of talents, where the strength of complementarities rises endogenously with skill specificity.

At the macro level, skill specificity promotes equilibrium talent sorting – workers of similar talent cluster in the same workplaces – which generates firm-level dispersion in productivity and average pay. This occurs because endogenous talent complementarities allow firms with talented employees to outbid competitors for other talented candidates. Search frictions, however, create a trade-off between match quality and waiting costs, resulting in imperfect sorting that intensifies with skill specificity. Teams spanning wide talent gaps emerge only when firms encounter workers who are horizontally very distant.

Measurement. To test and quantify these predictions, I take the theory to administrative matched employer-employee panel data for Germany, treating each establishment as one team. As skill specificity is not directly observable, I infer it from wage changes induced by variation in coworker talent. This allows calibrating the model without targeting sorting or firm-level wages. Using 2010-2017 data, the quantified model successfully matches these untargeted sorting patterns. Moreover, consistent with the theory, industries like architecture or chemical manufacturing that involve more complex tasks – which I treat as a proxy for skill specificity – exhibit stronger talent complementarities and sorting than industries where routine tasks dominate, such as accommodation services or furniture manufacturing.

Applications. Theory and data jointly provide a quantitative answer to how skill specificity shapes inequality and aggregate productivity.

First, growing skill specificity helps explain a prominent feature of labor market inequality since the mid-1980s: the increased share of wage dispersion attributable to between-firm gaps in average pay, which I show was accompanied by rising talent sorting. Re-estimating model parameters indicates skill specificity has increased concurrently, consistent with the suggestive evidence. Under my preferred specification, the model predicts a 17 p.p. increase in the between-establishment share of the log wage variance from 1985-1992 to 2010-2017 (close to the 24 p.p. in the data). Counterfactual exercises show greater specificity accounts for 77% of this “firming up” of inequality.

Second, because productivity gains from team production and complementarities are endogenously linked, aggregate gains from specialization are limited by the quality of labor markets. The microfoundation revealed that while greater specificity raises potential productivity – in line with classic Smithian and Ricardian intuition – this comes with a catch: output simultaneously becomes more vulnerable to mismatch. Quantitatively, I find that in the German economy, around two-thirds of potential gains are realized. This share shrinks to a quarter when information frictions are more severe, suggesting they can act as a bottleneck for specialization-driven growth.

These analyses jointly highlight that increased sorting and, hence, between-firm gaps in productivity and pay need not reflect worsening frictions, such as labor or product market power (Deb et al., 2024); in this model, they are, instead, endogenous responses to rising skill specificity that keep the economy close to the productivity frontier.

Discussion. The paper's primary contribution is a framework to study how specialization inside firms shapes macroeconomic outcomes. As a first step, it involves some strong assumptions – notably the absence of within-firm coordination frictions and two-worker teams – so the quantitative analyses are best viewed as explorations rather than definitive assessments. The framework opens avenues to study the macro effects of specialization and team quality in domains such as innovation, management, and firm dynamics.

Related literature. An emerging literature studies how worker interdependencies shape macroeconomic outcomes. Notably, Jarosch et al. (2021) and Herkenhoff et al. (2024) examine coworker learning as a source of human capital accumulation, identifying large aggregate consequences of structural shifts in coworker talent sorting.[5] My paper is complementary, providing a richer account of production and thus the determinants of sorting while abstracting from learning. Relative to Herkenhoff et al. (2024) specifically, on whose labor market matching setup I build, I make two contributions. First, theoretically, I allow skills to be multi-dimensional instead of one-dimensional – a precondition for analyzing specialization that introduces significant complexity. The microfounded production function I develop is crucial to capture the implications while preserving tractability. Second, endogenizing productivity and complementarities – treated as exogenous in Herkenhoff

[5] Also see Akcigit et al. (2018), Chade and Eeckhout (2020), Boerma et al. (2021), Pearce (2022).

et al. (2024) – reveals that growing skill specificity over time explains trends in labor market inequality, and that productivity gains are constrained by labor market frictions.

The paper builds on seminal theories of firm production, notably Kremer (1993) and Garicano (2000).[6] While sharing Kremer's (1993) emphasis on coworker complementarities, my theory differs in two ways. First, it microfounds a production function wherein complementarities endogenously vary with skill specificity, instead of increasing mechanically with team size. Second, many influential results in Kremer (1993) hinge on increasing returns to team quality, whereas this paper isolates complementarity effects. Modeling the organizational problem underlying a production function is inspired by Garicano (2000) and Garicano and Rossi-Hansberg (2006). The focus differs substantially: Where knowledge hierarchy theory considers vertical relationships – better workers enable larger spans of control – this paper examines the origins and implications of horizontal complementarities. Compared to both papers, I relax the assumption of frictionless markets; in Kremer (1993), for example, sorting is perfect for *any* positive degree of complementarity. Introducing search, and thus mismatch, allows me to address two key questions: how sorting shifts due to variation in complementarities, and how the productivity effects of rising specialization are shaped by endogenously increasing vulnerability to mismatch.

Important building blocks come from the literature on task-assignment – notably Acemoglu and Restrepo (2018) and Ocampo (2022) – and trade. Technically, the contribution is to show how Eaton and Kortum's (2002) stochastic formulation of technological heterogeneity in terms of extreme-value distributions, generalized by Lind and Ramondo (2023), can be leveraged to aggregate across a continuum of tasks given multi-dimensional skill heterogeneity.[7] The practical payoff is a microfounded, aggregative production function with intuitive properties that is sufficiently parsimonious to study macro-level outcomes.

This paper also contributes to a theoretical literature on worker-firm dynamics. While a production function additively separable in labor is commonly assumed, there are frequent

---

[6]More broadly, the paper relates to studies of firm organization, including Lucas (1978), Sattinger (1993), Kremer and Maskin (1996), Saint-Paul (2001) and Gabaix and Landier (2008); and, more recently, Porzio (2017), Adhvaryu et al. (2020), Caliendo et al. (2020), Adenbaum (2022), Bloesch et al. (2022), Kohlhepp (2022), Minni (2022), Kuhn et al. (2023) and Bassi et al. (2023).

[7]Deming's (2017) study of social skills inspired the analogy to trade, but without offering a tractable team production function or analyzing coworker complementarities and equilibrium matching.

calls for richer, "empirically implementable theories of the firm" (Hagedorn et al. (2017, p.33); Lentz and Mortensen (2010, pp. 593, 598)). This paper answers these calls by developing a tractable model of firm production emphasizing the interdependence of coworkers. In modeling heterogeneous skills across tasks, this paper echoes recent theories of multi-dimensional skills (Lindenlaub (2017), Lindenlaub and Postel-Vinay (2023), Baley et al. (2022)), integrating elements of circular models à la Marimon and Zilibotti (1999), Gautier et al. (2010), and Martellini and Menzio (2020). Whereas these theories abstract from team production, this paper distinctively demonstrates that worker heterogeneity across tasks is precisely what generates coworker interdependencies.

Finally, the application to wage inequality connects two lines of research: the long-standing analysis of how changes in the nature of work affect inequality,[8] and more recent, statistical work highlighting the large and growing role of firms.[9] However, explanations for this latter shift are mostly confined to qualitative hypotheses – notably, Song et al. (2019, V.B.) conjecture that coworker complementarities and sorting may play a role – while structural accounts are scarce. This paper contributes such an account: by formally microfounding a richer model of firm production it ties the abstract notion of complementarities to skill specificity and thus concrete shifts in the nature of work; and quantifies how this mechanism explains rising between-firm inequality.

Outline. Section 2 develops the theory, Section 3 confronts it with data, Section 4 presents applications, and Section 5 discusses limitations and directions for future work. Further material is contained in (i) a *Supplemental Online Appendix* ("Appendix O") , and (ii) *Additional Materials* ("Appendix A") available from the author's website.

2. THEORETICAL MODEL

2.1. *Environment*

Time is continuous, runs forever, and we consider the economy in steady state.

Demographics. The economy is populated by workers and firms. All agents are infinitely-lived and have linear preferences over the numeraire final good discounted at

[8] See, notably, Autor et al. (2003), Acemoglu and Autor (2011), and Autor et al. (2024).

[9] Prominent examples of this burgeoning literature include Card et al. (2013), Barth et al. (2016), Helpman et al. (2017) and Song et al. (2019).

rate $\rho \in (0,1)$. There is a unit measure of workers indexed by $i \in \mathcal{I} = [0,1]$. Workers are endowed with heterogeneous, time-invariant task-specific skills, $\{z_i(\tau)\}_{\tau \in \mathcal{T}}$ for any $i \in \mathcal{I}$, where $\mathcal{T} = [0,1]$ denotes the set of tasks. Both tasks and skill distribution are described below. Workers are either employed or unemployed. When employed, a worker inelastically supplies a finite amount of time in exchange for a wage. When unemployed, they receive a flow utility b_i. There is also a fixed unit measure of ex-ante homogeneous firms (or "entrepreneurs") that are either idle or actively producing by employing workers.

Task-based production technology. Production of the final good requires a unit continuum of imperfectly substitutable tasks. I have a granular notion of tasks in mind - producing economics research, for instance, might involve coming up with an idea, writing down a model, solving it, collecting and analyzing data, writing a paper, and so on. An active firm j employs a discrete number $n \in \{1,2\}$ of workers – so a "team" consists of two workers – who produce tasks that are combined into the final good according to

$$Y_j = \left(\int_{\mathcal{T}} q_j(\tau)^{\frac{\eta-1}{\eta}} d\tau \right)^{\frac{\eta}{\eta-1}}, \tag{1}$$

where Y_j is the quantity of the final good produced by j, $\eta \in (0,\infty)$ is the task elasticity of substitution, and $q_j(\tau)$ denotes the amount of task τ used, given by

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

Equation (2) captures the division of labor is possible, in that $q_j(\tau)$ can be positive even as $y_i(\tau) = 0$ as long as $y_k(\tau) > 0$ for $k \neq i$. The amount of task τ produced by worker i, $y_i(\tau)$, depends linearly on i's skill at task τ and the time dedicated by i to τ:

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

where $l_i(\tau)$ has to satisfy the time constraint

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

Labor market matching and information structure. Firms and workers meet through random search. Specifically, unemployed workers contact firms with a vacant position (i.e.,

$n < 2$) at a Poisson rate λ_u. After meeting, but not before, the firm and the worker can observe and contract on the task-specific skills. The parties decide whether to form a match or, instead, to keep on searching. If a match is formed, the wage is determined through generalized Nash bargaining, the worker's bargaining power being $\omega \in [0, 1]$. Further details are relegated to Section 2.3.1; for now it suffices to say that this assumption ensures that the wage will be monotonically increasing in firm-level output and all matching decisions are privately efficient. In an existing match, a worker i is separated from their employer at an exogenous Poisson rate δ_i.

Task-specific skills. To model skill heterogeneity over the infinite-dimensional task space parsimoniously, I adopt a stochastic formulation in the spirit of Eaton and Kortum (2002). Consider first the marginal distribution of skills for individual $i \in \mathcal{I}$. The skill in task τ is the realization of a random variable Z_i, drawn independently for each τ, which follows an individual-specific Fréchet distribution, $\Pr[Z_i \leq z_i] = \exp\left(-\left(\frac{z_i}{\iota x_i}\right)^{-\frac{1}{\chi}}\right)$. Here, $x_i \in \mathbb{R}_+$ is an individual-specific scale parameter and $\chi \in \mathbb{R}_+$ is the common inverse shape parameter. In the population, $x \in \mathbb{R}_+$ is distributed according to a distribution function $\tilde{\Phi}(x)$; sometimes it will be useful to work with the rank $\hat{x} = \tilde{\Phi}(x)$, i.e. $\hat{x} \in [0, 1]$.[10]

The scale term x_i captures a worker's *talent*, or absolute advantage: A larger x_i means that a high skill draw for any task τ is more likely. The economy-wide, structural parameter of central interest in the paper is χ, which governs the degree of *skill specificity*: Higher values imply greater within-worker dispersion in task-specific skills – a worker is good at some tasks but less good at others, so productivity drops more steeply as we move away from a worker's best task. Appendix O Figure D.1 provides a graphic illustration.

To model skill correlation across workers, I introduce additional geometric structure: Each worker is located on a cylinder, which is latent in the sense that their position shapes the distribution of task-specific skills. We already discussed height, which pins down talent x_i. Conditional on talent, individuals are uniformly distributed on the circle, which has circumference two. Letting $\mu_i \in [0, 2)$ be i's position, $d_{il} = d(\mu_i, \mu_l) = \min\{|\mu_i - \mu_l|, 2 -$

[10] A couple of technical remarks: First, throughout the paper it is assumed that $1 + \chi(1 - \eta) > 0$. Beyond requiring that tasks not be too substitutable, under the maintained assumptions the exact value of η will not influence task assignment. Second, $\iota := \Gamma\left(1 + \chi(1 - \eta)\right)^{1/(1-\eta)}$, with Γ denoting the Gamma function, is a common scaling term which ensures that varying χ or η does not mechanically change productivity.

$|\mu_i - \mu_l|\}$ measures the arc-length distance between any two points (i,l) on the circle, so $d_{il} \in [0,1)$. Then for any $(i,l), i \neq l$, the bivariate dependence structure of skills is governed by a pair-specific Gumbel copula:[11]

$$C_{il}(u,v) = \exp\left\{-\left[(-\log u)^{1/\xi_{il}} + (\log v)^{1/\xi_{il}}\right]^{\xi_{il}}\right\} \tag{5}$$

where the tail dependence parameter ξ_{il} is *distance-dependent* and governed by $\xi_{il} = g(d_{il})$, with $g : [0,1) \to [0,1)$ increasing. Intuitively, greater distance along the circle and hence greater ξ_{il} – which I also, somewhat loosely, refer to as "horizontal distance" – implies greater dissimilarity in what tasks any pair of workers excel at.

Importantly, the dependence structure depends only on the circular distance, not the specific locations (μ_i, μ_l). Furthermore, for any $i \in \mathcal{I}$, the pairwise distances to all other workers are uniformly distributed, so $\Pr(d_{il} \leq d|i) = d$, for $d \in [0,1)$. Of course, the distribution of talents and ξ_{il} in *teams* will be endogenously determined.

Discussion. The theory has four central features: (i) tasks are imperfectly substitutable; (ii) workers have heterogeneous skills for these tasks and limited time; (iii) production can involve multiple, but not infinitely many, workers; and (iv) hiring is subject to search. This latter feature is empirically salient – information about workers' granular skills is scarce – and proves important for both substantive insights and measurement. In principle, though, the production model can be embedded in different labor market environments.

A few important but implicit simplifications merit highlighting. First, firms are ex-ante identical – all productive knowledge is embodied in workers, with firms simply organizing them. This assumption transparently reveals how firm heterogeneity in observables can emerge endogenously, but could be relaxed. Second, I do not explicitly model a span-of-control problem and abstract from multi-team firms or agency problems. Third, tasks and skills are symmetric – none inherently more valuable, difficult or scarce – while workers are equally specialized overall. The latter two assumptions isolate the core interaction between specialization and complementarities while maintaining tractability.

[11] By focusing on pairwise relationships, the model captures the most relevant dependencies while preserving tractability and, through the latent circular structure, heuristically ensuring consistency of the bivariate margins. Appendix O.A.3 discusses a more general approach.

Two trade-offs arise from the theory's features. First, (i)-(iii) imply there is scope for gains from the division of labor, yet workers must perform a variety of tasks. Second, search (iv) creates a trade-off between productivity gains from well-matched teams and the opportunity cost of the time required to form them. The next two sections analyze these trade-offs. I first consider a single firm with an arbitrary workforce and analytically derive the production function from microfoundations. This tractable production function then allows analyzing equilibrium team formation.

2.2. *Team production*

Consider a single firm – hence the j subscript is omitted – that employs n workers. Treating the skill composition as exogenous, how much output can be produced?

2.2.1. *Organizational problem*

The firm solves a "mini-planner problem" by choosing task usage $\{q(\tau)\}_{\tau\in\mathcal{T}}$, individual task production $\{\{y_i(\tau)\}_{\tau\in\mathcal{T}}\}_{i=1}^n$ and individual time allocation $\{\{l_i(\tau)\}_{\tau\in\mathcal{T}}\}_{i=1}^n$ to maximize production Y subject to equations (1)-(4).

REMARK 1—Reduced-form production function: The reduced-form production function f is defined as the solution to the problem of optimally assigning tasks,[12]

$$f(\cdot) = \max Y \quad (6)$$
$$\text{s.t. } (1)-(4).$$

The Lagrangean associated to the firm's problem is

$$\mathcal{L}(\cdot) = Y + \lambda\left[\left(\int_{\mathcal{T}} q(\tau)^{\frac{\eta-1}{\eta}} d\tau\right)^{\frac{\eta}{\eta-1}} - Y\right] + \int_{\mathcal{T}} \lambda(\tau)\left(\sum_{i=1}^{n} y_i(\tau) - q(\tau)\right) d\tau$$
$$+ \sum_{i=1}^{n}\left\{\lambda_i^L\left(1 - \int_{\mathcal{T}} l_i(\tau)d\tau\right) + \int_{\mathcal{T}} \lambda_i(\tau)\left(z_i(\tau)l_i(\tau) - y_i(\tau)\right)d\tau + \int_{\mathcal{T}} \bar{\lambda}_i(\tau)y_i(\tau)d\tau.\right\},$$

Here, λ, λ_i^L, $\lambda_i(\tau)$ and $\lambda(\tau)$ denote the shadow values of, respectively, total production, i's time, a unit of task τ produced by i, and a unit of task τ used in final good production,

[12] I am deliberately loose in specifying the domain of f for now, as Proposition 1 will clarify this point.

while $\bar{\lambda}_i(\tau)$ attaches to a non-negativity constraint in task production. Importantly, each worker's time is scarce, with multiplier λ_i^L capturing the opportunity cost of i's time.

Considering the demand for tasks first, and treating the shadow prices as known, the first-order condition (FOC) for $q(\tau)$ and standard algebra yields the usual isoelastic functions of (relative) task-specific shadow costs, $\{\lambda(\tau)\}_{\tau\in\mathcal{T}}$ (see Appendix O.A.1.2). These costs in turn depend on task assignment. The FOC with respect to $y_i(\tau)$ implies that $\lambda(\tau) = \lambda_i(\tau)$ if $y_i(\tau) > 0$. Since some worker will provide a given task τ, and with task production featuring constant returns to scale, cost-minimization requires that

$$\lambda(\tau) = \min_{i=1,\ldots,n} \left\{ \frac{\lambda_i^L}{z_i(\tau)} \right\}, \tag{7}$$

where I used the FOC for $l_i(\tau)$ to substitute for $\lambda_i(\tau)$. Thus, any task τ will optimally be performed by the worker with the lowest ratio of the task-invariant shadow cost of time to task-specific skill; and the firm's shadow cost for that task, $\lambda(\tau)$, equals that minimum. The optimal assignment thus features complete division of labor – no strictly positive mass of tasks is performed by more than one worker – with every worker performing that interval of tasks where she has a comparative advantage.

A second look at equation (7) clarifies that taking a step further to derive f is non-trivial. Doing so requires knowledge of the entire distribution of the minimum of the worker-task specific shadow costs. This object itself depends on $\{\lambda_i^L\}_{i\in\mathcal{I}}$, which in turn is an endogenous function of skills.

The max-stability property of the Fréchet distribution allows overcoming this challenge, as shown by Eaton and Kortum (2002) in a trade context. Exploiting this property, Lemma A.3 in Appendix O pins down optimal task shares and λ.

2.2.2. *The microfounded production function*

We are now in a position to analytically characterize the production function f. It is useful to keep the following, easy-to-derive benchmark in mind:

REMARK 2—No division-of-labor benchmark: When $\chi = 0$, or for any value of $\chi > 0$ but no division of labor, firm-level output is the separable sum of talents:

$$Y^{\text{no DoL}} = \sum_{i=1}^{n} x_i \tag{8}$$

By contrast, aggregating over the task continuum under the *optimal* assignment yields the following reduced-form, firm-level production function:

PROPOSITION 1—Reduced-form production function: Output Y in a team of $n = 2$ workers can be written as a function of employee talents $\{x_i\}$ and horizontal distance ξ_{il}:

$$
\begin{aligned}
Y &= f(x_i, x_l, \xi_{il}; \chi) \\
&= \left((x_i)^{\frac{1}{1+\chi\xi_{il}}} + (x_l)^{\frac{1}{1+\chi\xi_{il}}} \right)^{1+\chi\xi_{il}} && (9) \\
&= \underbrace{2^{1+\chi\xi_{il}}}_{\text{efficiency gains}} \times \underbrace{\left(\frac{1}{2}(x_i)^{\frac{1}{1+\chi\xi_{il}}} + \frac{1}{2}(x_l)^{\frac{1}{1+\chi\xi_{il}}} \right)^{1+\chi\xi_{il}}}_{\text{talent complementarity}} && (10)
\end{aligned}
$$

PROOF: See Appendix O.A.1.3. The proof uses Lemma A.3, normalizing $\lambda = 1$. *Q.E.D.*

Thus, despite skills being high-dimensional, $f(\cdot)$ is highly tractable. Aggregating over the task continuum ensures output is deterministic, and optimal assignment with Fréchet skills implies talents and horizontal distance are sufficient to determine output – without requiring knowledge of either task-specific skills or the micro-level task assignment.[13].

Turning to the shape of f, the next result immediately follows from equation (9).

COROLLARY 1: For any $\chi\xi_{il} > 0$, the production function $f(\cdot)$ is

(i) *super-additive*, as $f(\cdot) > Y^{\text{no DoL}}$;

(ii) *super-modular* in talent, as $\frac{\partial^2 f}{\partial x_i \partial x_l} > 0$;

(iii) monotonically increasing in ξ_{il}.

When skills are task-specific, joint output exceeds production absent division of labor, and coworkers' realized productivities are inherently interdependent. Following Alchian and Demsetz (1972, Section II), I call this "team production". The theory clarifies what it

[13]This aggregation result relates to the seminal paper of Houthakker (1955) and, more recently, Jones (2005), Lagos (2006) and Acemoglu and Restrepo (2018). In independent work, Dvorkin and Monge-Naranjo (2019) similarly derive a CES function, though their analysis is less general by assuming $\xi_{il} = 1$.

means for workers to be "complementary" or have "complementary skills": Two workers are complementary insofar as they excel at *different* tasks but are *similar* in talent.[14]

Equation (10) expresses output as the product of two terms, showing how the *degree* of skill specificity, χ. The first term captures potential efficiency gains from team production versus no division of labor. These gains increase with χ because, intuitively, division of labor matters only when workers' realized productivity varies substantially across tasks. The presence of ξ_{il} reveals the first of two ways in which output depends on team composition: Potential gains are realized only if coworkers are skilled at different tasks (high χ_{il}).

The second term captures the other dependence: As skill specificity χ increases, output becomes more sensitive to *talent* composition. This term has a CES structure, with the strength of talent complementarities endogenously increasing with χ:[15]

COROLLARY 2: Let the inverse elasticity of substitution ("elasticity of complementarity") across talents be defined as $\gamma_{il} := \frac{\partial \ln(f_l/f_i)}{\partial \ln(x_i/x_l)}$, for any homothetic production function $f(x_1, ..., x_n; \cdot)$, where $f_i = \partial f / \partial x_i$ denotes a partial derivative. For the production function (9), the elasticity of complementarity is equal to $\gamma = \frac{\chi \xi_{il}}{1+\chi \xi_{il}}$.

Intuitively, the more specialized skills are, the more beneficial talented coworkers become, with the marginal contribution of greater worker talent higher when matched with other high-talent workers: $\frac{\partial^3 f(\cdot)}{\partial x_i \partial x_l \partial \chi} > 0$.[16] Correspondingly, by the power mean inequality, greater χ increases the relative weight on the least-capable team member's talent in determining Y. So, other things equal, output is greater when coworkers have similar talent.

---

[14]Note that allowing for heterogeneity in worker time endowments, denoted l_i, shows l_i to be isomorphic to x_i in $f(\cdot)$; see Appendix O equation (A.20). Thus, the microfoundation for complementarity in talent also explains complementarity in overall hours worked.

[15]The elasticity of substitution across tasks, η, does not show up in equation (9). Technically, as noted by Eaton and Kortum (2002, Footnote 18), this holds as long as $1 + \chi - \eta\chi > 0$, in which case η only appears in a constant term that cancels with the scaling term ι. The irrelevance of η in that sense is a tight implication of the Fréchet, and it highlights that coworker complementarities do not hinge on the assumption that tasks combine in a Leontief fashion (e.g., Kremer, 1993).

[16]The cross-partial connects directly to super-modularity, measuring how one worker's marginal talent productivity changes with coworker talent. It also links to the elasticity of complementarity, as $\gamma_{il} = (f f_{il})/(f_i f_l)$ for any $i \neq l$, where subscripts indicate partial derivatives. This observation goes back to Hicks (1932, 241-246).

COROLLARY 3: Consider a team with talents $\boldsymbol{x} = (x_i, x_l)$, where $x_i > x_l$. Let $\boldsymbol{x}'$ be the vector $\boldsymbol{x}$ such that x_i is replaced with $x_i + \epsilon$ and x_l with $x_l - \epsilon$, where $\epsilon > 0$. Then for $\chi\xi_{il} > 0$, $f(\boldsymbol{x}, \xi_{il}) > f(\boldsymbol{x}', \xi_{il})$.

What underpins these properties is the assignment of tasks, and how it varies with χ.

REMARK 3—Task shares: The share of tasks produced by worker i is equal to

$$\pi_i = \left(x_i^{\frac{1}{1+\chi\xi_{il}}}\right)\left(\sum_{k=1}^{n}(x_k)^{\frac{1}{1+\chi\xi_{il}}}\right)^{-1}. \tag{11}$$

Hence: (i) $\frac{\partial \pi_i}{\partial x_i} > 0$ and $\frac{\partial \pi_i}{x_k} < 0$ for $k \neq i$; and (ii) supposing that $x_i > x_l$, it holds that $\pi_i > \pi_l$ for $\chi\xi_{il} < \infty$, $\pi_i/\pi_l \to x_i/x_l$ as $\chi\xi_{il} \to 0$, and $\pi_i/\pi_l \to 1$ as $\chi\xi_{il} \to \infty$.

PROOF: Equation (11) follows directly from Lemma A.3 given $\lambda = 1$. *Q.E.D.*

Consider a team where $x_i > x_l$. Per equation (11), i optimally performs more tasks than l. Otherwise, tasks where i excels would be inefficiently overproduced. If skills are not task-specific, i's average usage-weighted skill – i.e., $\int_{\mathcal{T}} l_i(\tau)z_i(\tau)d\tau$ – is unaffected. But when $\chi > 0$, as a worker's marginal task is their worst, $z_i(\tau)$ diminishes as i's task share expands. Thus, when χ is greater, *relatively* more tasks are performed by the less talented worker l than when χ is low – this intuitively underpins the power mean intuition above. Further, i's realized average skill is greater when her coworker is more talented, allowing her to focus on the tasks she is best at, with the absolute magnitude of this effect increasing with x_i. Simply put, low-talent workers' task allocation matters little in absolute terms; high-potential workers' task allocation matters greatly.

Why microfound the production function? What concrete insights do we gain by deriving the production function from deeper microfoundations, instead of assuming a reduced-form CES function? The answer is twofold. First, rather than taking coworker complementarities as given, the theory reveals they increase endogenously with skill specificity. Thus, when specificity varies across sectors or over time, this generates economic effects through amplified complementarities that a reduced-form approach would miss. Second, equation (9) is *not* simply a CES function with endogenous elasticities. The theory links potential productivity gains from specialization to increased complementarities and, thus, vulnerability to coworker skill mismatch: Gains from specialization require workers being matched up who posses complementary expertise. This identifies labor market frictions as a potential constraint on productivity growth. Looking ahead, Sections 3.3 and 4.2 explore the first insight; Section 4.3 examines the second.

2.3. *Team formation*

Equipped with production function (9), which summarizes output for *any* coworker combination, we can now study which teams are, in fact, formed in equilibrium.

2.3.1. *Further assumptions and notation*

To ease notation, I suppress subscripts for talents, denoting that of the incumbent by x and that of the coworker by x', with pairwise horizontal distance ξ.

Tractability. In a dynamic context, hiring decisions potentially depend on all task-specific skills of both candidate and incumbent. These skills are observable to agents upon meeting, raising the prospect of an infinite-dimensional state variable. Proposition 1 resolves this challenge, showing that (x, x', ξ) are "sufficient statistics" for joint output. While the econometrician observes only workers' talent types, not task-specific skills (see Section 3), unobserved match quality is isomorphic to the pairwise scalar term ξ. To prevent the full skill vector from becoming a state through other objects, I assume separation hazards δ_i and unemployment flow utility b_i depend only on talent: $\delta(x)$ and $b(x)$.

Distributions. Let $d_{m.1}(x)$ denote the measure of producing matches consisting of a firm and one worker of talent x, while $d_{m.2}(x, x')$ is the corresponding measure of matches with an additional coworker with talent x' for any value of ξ.[17] The population measure of workers with talent x, $d_w(x) = \tilde{\Phi}'(x)$, is the sum of those who are unemployed, in one-worker matches or in two-worker matches: $d_w(x) = d_u(x) + d_{m.1}(x) + \int d_{m.2}(x, \tilde{x}')d\tilde{x}'$. The aggregate unemployment rate is obtained by integrating over x, i.e., $u = \int d_u(x)dx$. Similarly for firms, the total measure of firms with fewer than two employees is $v = d_{f.0} + \int d_{m.1}(x)dx$, where $d_{f.0}$ is the mass of idle firms, and adding-up requires $d_f = d_{f.0} + \int d_{m.1}(x)dx + \frac{1}{2}\int\int d_{m.2}(x, x')dxdx'$. Dividing by 2 in the last term avoids double-counting, $d_{m.2}(x, x')$ being symmetric.

[17] For notational simplicity, and with some abuse of language, I use "measure" to refer to these and analogous functions. In the context of the continuous-type model, these should be interpreted as densities, while in the numerical implementation, which is based on a discretized type space, they represent point measures or masses.

Joint value and surplus. The joint value of a match between a firm and a worker of talent x is $\Omega_1(x) = V_{f.1}(x) + V_{e.1}(x)$, where $V_{e.1}(x)$ is x's value of being employed alone, whose value is $V_{f.1}(x)$. The surplus generated by such a match is $S(x) = \Omega_1(x) - V_{f.0} - V_u(x)$, where $V_u(x)$ is the value of unemployment and $V_{f.0}$ is the value of an idle firm.

The joint value of a firm with a team characterized by (x, x', ξ) is $\Omega_2(x, x', \xi) = V_{f.2}(x, x', \xi) + V_{e.2}(x|x', \xi) + V_{e.2}(x'|x, \xi)$, with $V_{e.2}(x|x', \xi)$ denoting the value of x being employed together with a coworker of talent x' at horizontal distance ξ. Hence, the surplus generated when the latter is added to the former is $S(x'|x, \xi) = \Omega_2(x, x', \xi) - \Omega_1(x) - V_u(x')$. This surplus is increasing in the joint value generated by the team and declines with both sides' outside values. Note $S(x|x', \xi)$ is not symmetric in x and x' even if $\Omega_2(x, x', \xi)$ is, because it matters for outside options who the incumbent employee is.

Surplus sharing. Wage bargaining takes the generalized Nash form and concerns the entire surplus. The firm treats each employee as marginal (so their outside option is unemployment), and renegotiation is continuous. This protocol ensures that all matching decisions are privately efficient and characterized by parties' joint surplus (cf. Bilal et al., 2022); and the wage is invariant to the order with which workers join a team.

Thus, the wage $w(x|x', \xi)$ of a worker of talent x employed with a coworker of talent x' who is horizontal distance ξ apart satisfies

$$(1-\omega)\big(V_{e.2}(x|x', \xi) - V_u(x)\big) = \omega\big(V_{e.2}(x'|x, \xi) + V_{f.2}(x, x', \xi) - V_{e.1}(x') - V_{f.1}(x')\big). \tag{12}$$

2.3.2. *Equilibrium conditions*

We are looking for a stationary equilibrium in which matching decisions are optimal and consistent with the distribution of agents to which they give rise being stationary. See Appendix O.A.2.3 for a definition. Decisions and associated values are characterized by a set of Hamilton-Jacobi-Bellman (HJB) equations, while a system of Kolmogorov-Forward equations (KFEs) describes flows across states. Their interaction means the equilibrium needs to be computed numerically.

Policy functions. The matching decisions are related to surplus values as follows:

$$h(x) = \mathbf{1}\{S(x) > 0\}, \quad h(x|x', \xi) = \mathbf{1}\{S(x|x', \xi) > 0\} \tag{13}$$

$$h(x|x') = \mathbf{P}\{S(x|x', \xi) > 0\} \tag{14}$$

The functions in equation (13) describe, respectively, whether a match between an unmatched firm and a type-x worker will be consummated; and whether a firm that already employs worker x' is willing to hire a worker x given shock ξ. Equation (14) defines conditional matching probabilities.

Value functions. The asset value of an idle firm, $V_{f.0}$, and that of an unemployed worker with talent x are straightforward and given by

$$\rho V_{f.0} = (1-\omega)\lambda_{v.u} \int \frac{d_u(x)}{u} S(x)^+ dx, \tag{15}$$

$$\rho V_u(x) = b(x) + \lambda_u \omega \left[\frac{d_{f.0}}{v} S(x)^+ + \int\int \frac{d_{m.1}(\tilde{x}')}{v} S(x|\tilde{x}', \tilde{\xi})^+ dH(\tilde{\xi}) d\tilde{x}' \right] \tag{16}$$

where, to ease notation, for any r, $r^+ = \max\{r, 0\}$ denotes the optimal decision.

Turning to matched agents, the joint value of a firm employing a team $(x, x', \xi)$and that of a firm talent-x employee respectively satisfy

$$\rho\Omega_2(x, x', \xi) = f(x, x', \xi) - \delta(x) S(x|x', \xi) - \delta(x') S(x'|x, \xi), \tag{17}$$

$$\begin{aligned} \rho\Omega_1(x) &= f(x) + \delta(x)\big[-\Omega_1(x) + V_u(x) + V_{f.0} \big] \\ &\quad + \lambda_{v.u} \int\int \frac{d_u(\tilde{x}')}{u} \big(\underbrace{-\Omega_1(x) + V_{e.2}(x|\tilde{x}', \tilde{\xi}) + V_{f.2}(x, \tilde{x}', \tilde{\xi})}_{(1-\omega)S(\tilde{x}'|x,\tilde{\xi})} \big)^+ dH(\tilde{\xi}) d\tilde{x}'. \end{aligned} \tag{18}$$

This equation embodies the key matching decision: Beyond production flow value and match separation risk, at rate $\lambda_{v.u}$ the firm is contacted by an unmatched worker with talent $\tilde{x}'$ at horizontal distance ξ. This worker is hired only if joint surplus is positive. Appendix O.A.2.1 derives the recursions for these surplus values.

Population dynamics. The KFEs summarize inflows and outflows for each state, with the matching decisions (13) modulating the flow intensities implied by exogenous meeting rates and the distributions themselves. As these equations are straightforward but lengthy, they are collected in Appendix O.A.2.2.

2.3.3. *Key trade-off and equilibrium properties*

I next discuss the trade-off shaping equilibrium team composition and how skill specificity χ affects sorting.

Trade-off. When a firm with an employee of talent x meets an unmatched worker of talent x' at horizontal distance ξ, the policy function $h(x'|x,\xi)$ balances match quality against search costs: hiring yields flow output $f(x,x',\xi)$; rejecting means continued search with flow payoffs $f(x)$ and $b(x')$, respectively. The properties of $f(x,x',\xi)$ imply team formation is more likely between workers similar in talent and horizontally distant. A low-x candidate will never be paired with a high-x incumbent, even if ξ is high, since continuing search offers the prospect of finding a coworker at similar horizontal distance but closer in talent. Conversely, teams with similar talent form readily. For intermediate talent gaps, matching requires sufficient horizontal distance to offset the talent mismatch penalty. Appendix O Figure D.2a illustrates equilibrium matching probabilities.

Coworker sorting. A useful statistic to quantify talent sorting is the correlation between coworkers' talent ranks (Lopes de Melo, 2018). Appendix O.A.2.5 defines this statistic, $\rho_{xx'}$, in terms of equilibrium objects. This is as good a place as any to note that, since firms are ex-ante homogeneous, which firms get the most productive teams is random, varying with the first worker drawn. "Sorting" thus exclusively refers to *coworker* sorting (or "segregation"), not sorting between ex-ante heterogeneous firms and workers.

Comparative statics for χ. Greater skill specificity, by amplifying coworker talent complementarities, fosters more positively assortative matching. Appendix O Figure D.2b demonstrates this positive relationship between χ and $\rho_{xx'}$. Succinctly put, under strong skill specificity, equilibrium features firms with "superstar teams" composed of the most talented workers and others with "laggard teams."

3. MODEL MEETS DATA

This section evaluates the model's empirical support and quantitative relevance. Section 3.1 introduces the data, Section 3.2 maps model objects to data, Section 3.3 tests core mechanisms using industry-level variation, and Section 3.4 calibrates parameters and explores quantitative properties.

3.1. *Data*

I use the *Sample of Integrated Employer-Employee Data* (SIEED), an administrative matched employer-employee dataset for Germany provided by the Institute for Employment Research (IAB). It covers the entire workforce of a 1.5% sample of all establishments and these workers' complete employment biographies, including spells outside sampled establishments. The production unit is the establishment (defined by ownership, industry, and location); with that proviso, I use "establishment" and "firm" interchangeably. The dataset enables analysis of workers' employment trajectories, including their coworkers and earnings, over a long period.

For the analysis, I construct an annual panel of workers aged 20–60 employed full-time at West German private-sector establishments. The earnings variable is the residual daily wage, controlling for year dummies, education-specific age profiles, and job tenure. This section uses the 2010-2017 sub-sample (3,911,751 person-year observations); Section 4 extends the analysis back to 1985 (15,359,711 person-year observations). Appendix O.B.1.1 provides details on data construction and descriptive statistics.

3.2. *Measurement*

This section describes how three theoretical objects are mapped to the data: worker talent types; teams and coworkers; and skill specificity.

3.2.1. *Worker talent types*

In the theory, a worker's type $\hat{x}$ indexes their time-invariant talent. Empirically, I measure $\hat{x}$ by exploiting the panel structure of data and model. As the wage increases monotonically in talent – a standard propery of matching models with individual-level heterogeneity in absolute advantage – more talented workers have higher lifetime earnings, even if in a single cross-section unemployment or poor match quality temporarily reduces wages relative to less talented workers. I therefore estimate individual fixed effects (FE) from panel wage regressions, controlling for observable characteristics and time-invariant employer effects. Details are relegated to Appendix A.F.1. Ranking workers by their FEs and binning them into deciles – matching the ten-point talent grid that will be used in the numerical solution – yields $\hat{x}_i$ for each worker i.

As an alternative to ranking workers economy-wide, I also rank them within occupations. Under this interpretation, the coworker correlation $\rho_{xx'}$, for example, captures whether top performers in their respective occupations cluster together.

3.2.2. *Teams and coworker(s)*

Real-world production units typically comprise more than two workers. In the spirit of Jarosch et al. (2021) and Herkenhoff et al. (2024), I construct a "representative coworker" type for each worker-year observation to map the two-worker model to the data. I define worker i's coworkers in year t as employees in the same establishment-year cell, denoted $S_{-it} = \{k : j(kt) = j(it), k \neq i\}$, where $j(it)$ is the identifier of i's employer in year t. Importantly, this step aligns the theoretical notion of teams with establishments. Then i's representative coworker in year t is the leave-out mean, $\hat{x}_{-it} = \frac{1}{|S_{-it}|} \sum_{k \in S_{-it}} \hat{x}_k$.[18]

This empirical aggregation step is immensely useful yet amounts to a significant simplification, so it is worth considering what could go wrong. First, should only same-occupation employees count as coworkers, rather than everyone in the same establishment? The theory suggests otherwise – coworker complementarities arise from horizontal skill differentiation, which is more likely across different occupations contributing to joint output. Second, the theory strictly speaking implies that lower-talent coworkers should receive higher weights in aggregation. Taking an unweighted average is in line with existing studies but ignores these complementarities across an individual's coworkers. Appendix O.B.2.5 explains why the resulting bias is minor. Third, ignoring team size variation may seem problematic. Yet, averaging naturally captures that a single coworker change matters less in larger teams.

3.2.3. *Skill specificity*

Measuring skill specificity (χ) presents the greatest challenge. Ideally, we would measure χ directly using detailed individual-level data on task-specific skills. Such data do not exist.[19] Instead, I exploit the model's structure to infer χ indirectly, and show that these estimates align with a task-based, ordinal proxy.

Proposition 1 and Corollary 2 showed that $\partial^2 f(\cdot)/\partial x \partial x'$ systematically increases with χ. How can we quantify this cross-partial derivative with our data? The central insight:

---

[18]The arithmetic average implicitly assume that intervals between (ordinal) *types* are similar. This is consistent with the discretization of talent in Section 3.4. Alternatively, one could work with cardinal worker types throughout.

[19]Some datasets such as the NLSY and administrative data for Sweden and Denmark contain skill-specific test scores, but these are typically limited to broad categories like verbal or technical ability.

REMARK 4—Cross-partial proportionality: The cross-partial derivative of the production function is proportional to the corresponding cross-partial of the wage function:

$$\frac{\partial^2 f(x, x', \xi)}{\partial x \partial x'} \propto \frac{\partial^2 w(x|x', \xi)}{\partial x \partial x'}. \tag{19}$$

This result follows from differentiating the wage function (contained in Appendix O.A.2.4) with respect to x and x'. The worker receives an ω-share of the increase in production from being matched into a team, but the wage level also reflects outside options. While this complicates full identification of production functions using wage data (Eeckhout and Kircher, 2011, Hagedorn et al., 2017), Corollary 4 clarifies that the cross-partial derivative is unaffected, as the outside options are separable. Intuitively, the new hire's option value is foregone regardless of coworker quality, while the firm and incumbent's options are foregone regardless of who is hired.

Operationalization. Operationalizing Corollary 4 requires three more steps. In step one, as ξ is unobserved, I construct the average wage of a type-x worker employed together with a coworker of talent x', integrating over ξ, denoted $\bar{w}(x|x')$.[20] In step two, I approximate $\frac{\partial \bar{w}(x|x')}{\partial x \partial x'}$ by estimating the following auxiliary regression of individual-year level wages using OLS:

$$\begin{aligned}\frac{w_{it}}{\overline{w}_t} = \beta_0 + \sum_{d=2}^{10} \beta_{1d} \mathbf{1}\{\hat{x}_i = d\} + \sum_{d=2}^{10} \beta_{2d} \mathbf{1}\{\hat{x}_{-it} = d\} + \beta_c(\hat{x}_i \times \hat{x}_{-it}) \qquad (20)\\ + \psi_{j(i,t)} + \nu_{0(i,t)t} + \zeta_{S(i,t)t} + \epsilon_{it}\end{aligned}$$

where $\mathbf{1}\{\hat{x}_i = d\}$ and $\mathbf{1}\{\hat{x}_{-it} = d\}$ are dummies for worker and coworker types, and $\psi_{j(i,t)}$, $\nu_{o(i,t)t}$, and $\zeta_{s(i,t)t}$ are employer fixed effects (FE), occupation-year FEs, and industry-year FEs, respectively. Each observation is weighted by the inverse frequency of the ($\hat{x}_i$, $\hat{x}_{-it}$) match to ensure equal coverage of the state space. The coefficient of interest, β_c, indicates how much more the wage of an individual rises with a one-decile increase in coworker talent compared to someone whose rank is one decile lower.

[20]This step is not trivial, as, in the theory, the distribution of ξ among matches reflects selection (Borovickova and Shimer, 2024): for $\chi > 0$, the horizontal distance ξ tends to be larger for pairs with larger talent gap. Importantly, the calibration treats data and theory liked for like, i.e., the same auxiliary regression (20) will be run on model-generated data.

Identification comes from variation in coworker quality over time, which in the theory occurs naturally due to search frictions. This variation includes both movers (changes in coworker quality for individuals who switch employer) and stayers (changes in coworker quality when other employees join or leave). Using pre-determined type measures – rather than estimating types and coworker effects jointly – avoids the reflection problem (Manski, 1993), while non-parametric controls for individual types address coworker sorting.[21] Additionally, fixed effects absorb unobserved time-invariant employer heterogeneity and shocks at the occupation-year or industry-year level.

The baseline estimate of β_c is 0.0061 and statistically significant at the 1% level. To provide a sense of magnitude, this implies that a one-decile improvement in average coworker talent increases wages roughly 3% more for a top-decile worker compared to a fifth-decile worker (0.0061·(10-5)·1=0.0305). Appendix O.B.2 describes a battery of robustness analyses, including alternative controls, years of schooling as a non-wage wage type, and a log wage specification.

Finally, in step 3, the auxiliary regression coefficient β_c serves as the target disciplining skill specificity χ in a simulated-method-of-moments (SMM) procedure detailed in the next section – a larger β_c implies greater χ.

Discussion. Alternatively, one could calibrate χ by targeting the correlation coefficient $\rho_{xx'}$. The strategy adopted here has two principal advantages, however. First, the model is parsimonious and omits other mechanisms generating positive coworker sorting, such as preference-based homophily (workers preferring colleagues with similar talent) or sorting due to ex-ante firm heterogeneity. Targeting $\rho_{xx'}$ would therefore risk over-attributing sorting to χ. Under the current approach, by contrast, the model is given the chance to generate little sorting, letting the data determine whether skill specificity accounts for observed sorting patterns. Second, measuring skill specificity and sorting separately allows testing the model-implied link between them, as undertaken in Section 3.3.

[21] Thus, unlike typical peer effects studies (e.g., Cornelissen et al., 2017, Nix, 2020), I leverage the model's structure to separate type measurement from spillover estimation. Moreover, I focus on the interaction term rather than an average treatment effect.

Task complexity proxy. Beyond structurally estimating χ, I use a task-based proxy: the share of abstract and non-routine tasks in production, termed "task complexity."[22] Intuitively, workers likely perform similarly across different routine tasks, which almost by definition require few specific skills (Martellini and Menzio, 2021, p. 340). Complex tasks like investigating, teaching, or coordinating, by contrast, demand cognitive skills that typically require specific training, creating greater scope for productivity variation across tasks.[23] Following Deming (2017), I therefore interpret task complexity as proxy for skill specificity. I construct this proxy using multiple waves from the BIBB Employment Surveys, building on Spitz-Oener (2006). Appendix O.B.1.2 details the survey methodology, task classification, and linkage to the SIEED.

3.3. *Validation of model mechanisms: industry-level variation*

Before turning to a quantitative exploration of the model's properties, this section reports industry-level correlations consistent with the core mechanisms.[24]

To anticipate the time-trend analysis in Section 4.2, I expand the sample to span 1985-2017, dividing these years into five periods indexed by p: 1985-1992, 1993-1997, 1998-2003, 2004-2009, and 2010-2017. I measure worker and coworker types, construct the sorting coefficient separately for each (2-digit industry s, sample period p) cell, and obtain a point estimate for the interaction coefficient β_c by estimating regression (20) within each (s, p) cell. I also construct the task-complexity measure separately for each (s, p), cell, capturing variation from differences in occupational employment shares.

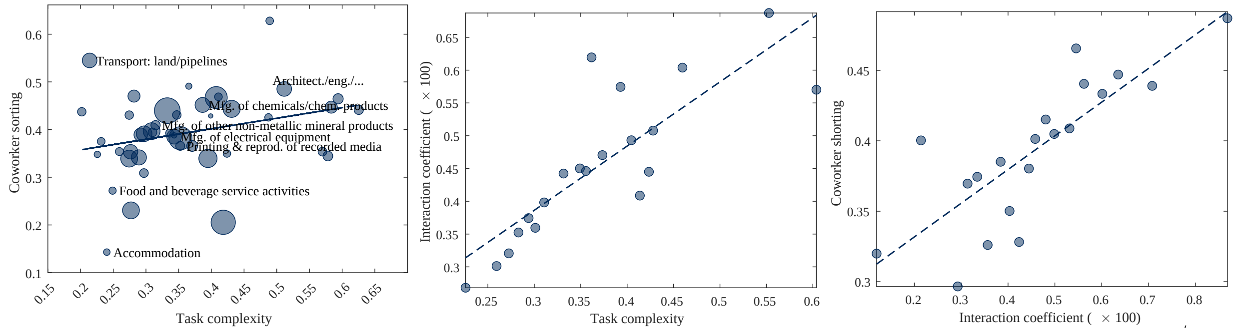
Figure 1a demonstrates that industries involving more complex tasks, such as architecture and engineering, exhibit more pronounced talent sorting than industries where routine tasks dominate, like accommodation services. Panels 1b and 1c unpack this positive association through the lens of the theory. Greater task complexity is associated with higher β_c, which in turn predicts stronger coworker talent sorting – consistent with the theory.

---

[22] This proxy does not map directly to the model. Analytical aggregation fails with two types of tasks, as a mixture of two Fréchets with different shape parameters is not Fréchet.

[23] Consistent with this idea, Caplin et al. (2022) find that the time needed to reach maximal productivity is highest for knowledge-intensive occupations.

[24] An earlier version of this paper (Freund, 2023) also reported cross-sectional evidence using Portuguese micro data. Also see the findings in a companion note (Criscuolo et al., 2024).



(a) Task complexity and sorting (b) Task complexity vs. β_c (c) β_c vs. ρ_{xx}

FIGURE 1.—Industry-level correlations are consistent with the model mechanisms

Notes. Panel (a) is based on data pooled across periods. The binscatters in panels (b) and (c) include period FEs.

While these panels rely on cross-industry variation, Appendix A.F.3 shows similar patterns using within-industry variation over time. Moreover, the structurally estimated degree of skill specificity correlates positively with task complexity at the industry level.

3.4. *Calibration*

Turning to the quantitative analysis of the model, I next describe how structural parameters are disciplined, then discuss calibration outcomes.

3.4.1. *Methodology*

I proceed in four steps. First, I impose some functional form assumptions to reduce the dimensionality of the estimation problem. Second, I preset three parameters based on literature and data. I directly infer parameters mapping one-to-one to empirical moments. Fourth, the remaining parameters are jointly inferred using the SMM.

Functional form assumptions. The unemployment flow value is proportional to solo output with factor b_1. Moreover, the separation rate varies linearly with talent, indexed by δ_0 and δ_1.

The production function builds directly on equation (10), while allowing greater flexibility for the quantitative exploration. First, since we measured ordinal worker types $\hat{x}_i$, while equation (10) took cardinal talent x_i as its input, I introduce two parameters mapping between the spaces, a_0 and a_1. Thus, solo output becomes $\hat{f}(\hat{x}) = a_0 + a_1\hat{x}$. Second, I introduce a size adjustment, since real-world teams often exceed two members: The second hire in the model appears as the $\bar{n}^{\text{th}}$ hire in the production function. Larger $\bar{n}$ mutes the degree to which skill specificity χ boosts team productivity relative to solo production. Third, according to Proposition 1, a greater value of horizontal distance ξ_{il} raises both productivity *and* the strength of coworker talent complementarities; the latter effect lowers productivity. I abstract from the link between ξ_{il} and talent complementarities to isolate how *horizontal* distance *raises* output whereas *talent* distance *lowers* it. Thus, the team production function in the quantitative analysis is:

$$\hat{f}(\hat{x}_i, \hat{x}_l, \xi_{il}) = 2 \times \left(\frac{\bar{n}}{\bar{n}-1}\right)^{\chi \xi_{il}} \times \left(a_0 + a_1 \left(\frac{1}{2}(\hat{x}_i)^{\frac{1}{\chi+1}} + \frac{1}{2}(\hat{x}_l)^{\frac{1}{\chi+1}}\right)^{\chi+1}\right). \quad (21)$$

Bilateral distances on the circle, d_{il}, maps to pair-specific tail dependence ξ_{il} via $g(d) = H^{-1}(d)$, where H is uniform over [0,0.5]. For any individual, potential ξ_{il} draws are thus uniformly distributed over [0,0.5]. The upper bound reflects that in larger teams, the effect of ξ might partly "wash out." Geometrically, the maximum average distance between one individual and many others distributed around the circle equals one quarter of the circle's circumference, i.e., 0.5.

Preset parameters. Regarding preferences and bargaining, I follow Herkenhoff et al. (2024). The discount rate ρ is set to 0.008, consistent with an annual interest rate of 10%. Such a high rate is common in this type of model, where a high discount factor effectively proxies for concavity in the utility function, which tractability requires us to abstract from. The bargaining parameter $\omega = 0.50$ implies equal sharing of surplus. Lastly, I set $\bar{n}$ equal to 20, which lies between the median and mean establishment size in the data.

Directly inferred parameters. Parameters indexing separation rates, δ_0 and δ_1, are estimated offline by matching type-specific monthly job losing rates, computed from an auxiliary dataset, the LIAB (see Appendix A.F.2).

Internally estimated parameters. The remaining five parameters, $\psi = \{\chi, a_0, a_1, b_1, \lambda_u\}$, are estimated using indirect inference methods by matching moments. The estimated values minimize the objective function

$$\mathcal{G}(\psi) = \sum_{j=1}^{5} \left(\frac{\hat{m}_j - m_j(\psi)}{\frac{1}{2}|\hat{m}_j| + \frac{1}{2}|m_j(\psi)|}\right)^2,$$

where $\hat{m}_j$ refers to the empirical moment and $m_j(\psi)$ denotes its model counterpart.

To construct the model moments I solve the model in continuous time over ten grid points for talent and simulate the model at monthly frequency for a large number of workers. All wage moments are constructed from the subsample of workers in teams. I average moments over five runs to mitigate noise. The minimum of $\mathcal{G}(\psi)$ is found using a hybrid particleswarm-patternsearch algorithm.

While the elements of ψ are jointly estimated, each is closely informed by one of these moments (see Appendix A Figure G.2). Most importantly, $\hat{\beta}_c$ directly informs χ, as already discussed in Section 3.2.3 and validated in a Monte Carlo study described in Appendix A. G.2. Concretely, I estimate an auxiliary regression analogous to equation (20) on model-generated data. Note that in theory and data the moment used is β_c scaled by the standard deviation of coworker types, $\sigma_{x'}$, to account for the possibility that constructing the representative coworker variable lowers the variability of coworker talent in the data relative to the model.

Parameters a_0 and a_1 target the average wage (normalized to unity) and the total log wage variance (0.23), respectively. While both parameters increase the average wage, a_0 reduces wage dispersion whereas a_1 amplifies it. Parameter b_1 reflects the rate at which unemployment flow benefits replace (type-specific) average wages. Drawing on official administrative replacement rates (post-Hartz labor market reforms) and Koenig et al.'s (2021) estimates of non-monetary opportunity costs of employment, I target a ratio of 0.66. Finally, λ_u targets a monthly job finding rate of 16.2% from the LIAB.

3.4.2. *Calibration outcomes and validation*

Table I summarizes the model parameters and evaluates the model fit. The model is capable of matching the targeted moments perfectly.

Talent complementarities. To quantify the strength of coworker talent complementarities under the estimated χ value of 2.47, I examine productivity gains from reallocating workers from two talent-mixed teams into two talent-homogeneous teams. Absent skill specificity this would be zero. Under the estimated parameters, the mean productivity gain is 7%; -matching workers within the bottom and top talent deciles rather than across them

Parameter	Description	Value	Source	m	$\hat{m}$
ρ	Discount rate	0.008	External		
ω	Worker barg. weight	0.50	External		
$\bar{n}$	Team size	20.00	External		
δ_0	Sep. rate, constant	0.015	Offline est.		
δ_1	Sep. rate, scale	-0.84	Offline est.		
χ	Skill specificity	2.47	$\beta_c \times \sigma_{\hat{x}'}$	0.012	0.012
a_0	Production, constant	0.27	Avg. wage	1.00	1.00
a_1	Production, scale	1.56	Var. log wages	0.23	0.23
b_1	Unemp. flow utility, scale	0.65	Replacement rate	0.66	0.66
λ_u	Meeting rate	0.23	Job finding rate	0.16	0.16

TABLE I

MODEL PARAMETERS (2010-2017 PERIOD)

Notes. This table lists for each of the parameters the value. For internally estimated parameters, it also indicates the targeted moment, and moment values in data ($\hat{m}$) and model (m).

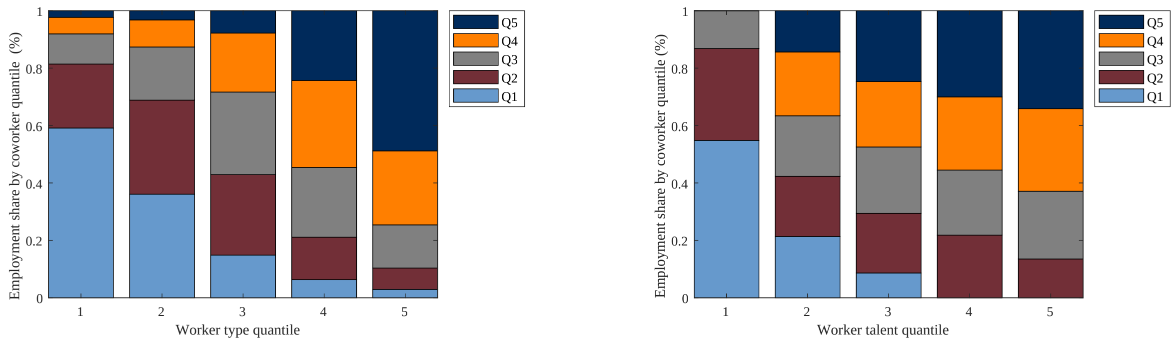
yields an 31% increase in output.[25] Appendix O Figure D.3 visualizes this statistic for all possible talent combinations.

Horizontal distance. Turning to the influence of horizontal distance ξ, and averaging over all possible talent combinations, output is 6% lower for $\xi = 0$ than when $\xi = 0.5$. For external validation, the model is qualitatively consistent with Jäger and Heining's (2022) findings about the heterogeneous wage effects from unexpected coworker deaths: While the average wage effect for remaining workers is positive, the death of a worker with highly specialized human capital causes the wage of incumbents in other occupations to *decrease* by 0.42% in the short-run.[26] As a model counterpart, I pool all workers who experienced a coworker separation shock, denoting by t^* the event period and the wage change by $\Delta w_{i,t^*} = \frac{w_{i,t^*} - w_{i,t^*-1}}{w_{i,t^*-1}}$. Estimating the regression

$$\Delta w_{i,t^*} = \beta_0 + \beta_1 \mathbf{1}\{\xi_{i,t^*-1} \in \text{top } 20\%\} + \sum_{d=2}^{10} \beta_{2d} \mathbf{1}\{\hat{x}_i = d\} + \sum_{d=2}^{10} \beta_{3d} \mathbf{1}\{\hat{x}_{-i,t^*-1} = d\} + \epsilon_{i,t^*}$$

[25] The estimated complementarities in talent are qualitatively consistent with the evidence documented in Shao et al. (2023) that coworker hours are gross complements.

[26] Many thanks to the authors for sharing statistics in percentage terms with me.



(a) Data

(b) Theory

FIGURE 2.—Model generates empirically realistic talent sorting patterns

Notes. For each quantile of the talent distribution, these plots – data on the left, model on the right – show the employment share of coworker talent types, where those are binned into quantiles, too.

yields estimates $\hat{\beta}_0$ = 0.016 and $\hat{\beta}_1$ = -0.023. The model thus offers a structural interpretation of the quasi-experimental evidence: when skills are specialized, losing a coworker has negative effects if that coworker was sufficiently horizontally distant.

Aggregate employment outcomes. Turning to macro-level outcomes, OA Figure D.4 shows that more talented workers face lower unemployment risk. Consistent with data (e.g. Cairó and Cajner, 2018), this is due to lower separation rates, not higher finding rates. In the model, talented workers only selectively accept offers.

Labor market sorting and firm-level dispersion. The model reproduces empirically observed talent sorting patterns – an important piece of validation since no evidence on sorting was used in the calibration. Specifically, the model-implied coworker correlation 0.51 nearly matches the empirical value (0.64). Figure 2 provides a more disaggregated picture, comparing predicted and actual conditional distributions of coworker talent by quantiles. The model's predictions (panel 2b) fit the data fairly well (panel 2a). The model also reproduces the observed decomposition of wage dispersion into between- and within-firm components, as detailed in Section 4.2. Thus, despite assuming ex-ante identical firms, the model generates ex-post differences in observables – some firms are more productive and pay higher wages simply through superior team quality.

4. THE MACRO IMPLICATIONS OF RISING SKILL SPECIFICITY

In this section, I use the calibrated model to show how rising skill specificity, mediated by firm organization, affects macro outcomes. After briefly reviewing suggestive evidence for rising specificity (Section 4.1), I first demonstrate how, through the lens of the theory, this explains the "firming up of inequality" in Germany (Section 4.2); I then consider how the endogeneity of coworker complementarities, interacting with labor market frictions, moderates "Smithian" productivity gains from specialization (Section 4.3).

4.1. *Growing skill specificity*

The idea that skills have become more specialized over recent decades has been articulated across multiple research strands.[27] First, extensive literature documents a within-country transformation of work involving jobs becoming less routine over time while demanding more specialized expertise (see Section 1 for references). Consistent with the argument in Section 3.2.3 and the cross-industry evidence in Section 3.3, Deming (2017) interprets this trend as implying increased skill specificity. The BIBB survey used in Section 3.2.3 shows this transformation is also visible in Germany. Figure 3a depicts an upward trend in the share of complex tasks reported, increasing from 0.25 in 1986 to 0.65 in 2018. The rise is especially pronounced between the 1990s and early 2000s and manifests across all education groups.

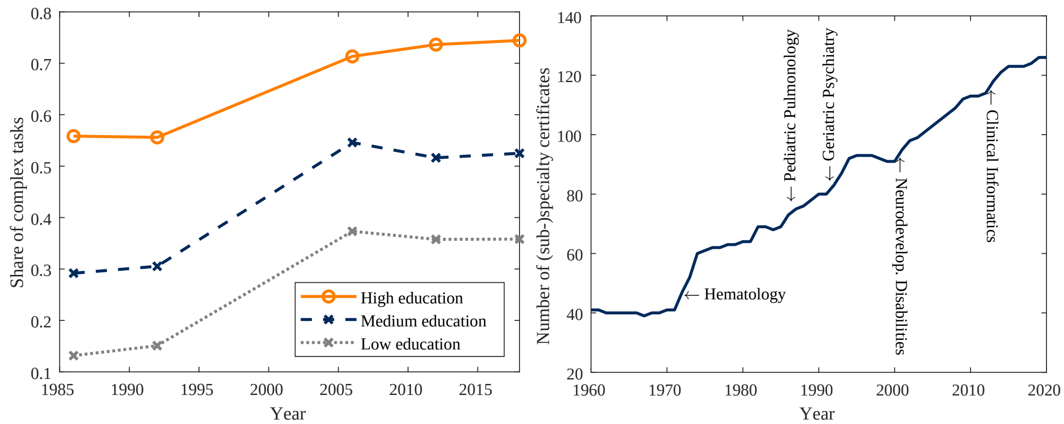
Jones (2009) offers a complementary argument, contending that the growing "burden of knowledge" – the cost of reaching the frontier – necessitates increasingly narrow individual expertise in science. This logic extends to knowledge work more broadly (Neffke, 2019). For example, Figure 3b shows that the number of distinct specialty certificates issued in the U.S. nearly doubled over the previous four decades.

There is also a long-standing idea, supported by cross-country evidence (Bandiera et al., 2022), that economic development broadly involves deepening skill specialization. Indeed, Caunedo et al. (2023) show that production in developed countries uses complex tasks more intensively than developing countries.

Discussion. It is worth stepping back to make two points about trends in skill specificity. First, regarding measurement, owing to the difficulty of measuring task-specific

---

[27] Staying closer to the mathematical structure of my model, Appendix O.A.4 shows that if education augments task-specific skills randomly, the sharp acceleration in secondary education since the 1980s could have caused a rise in the Fréchet skill specificity parameter χ. In addition, Grigsby (2023) uses a model-based approach to find that workers have become more specialized in the U.S.



(a) Rising task complexity (b) Deepening medical specialization

FIGURE 3.—Rising economy-wide task complexity & deepening medical specialization

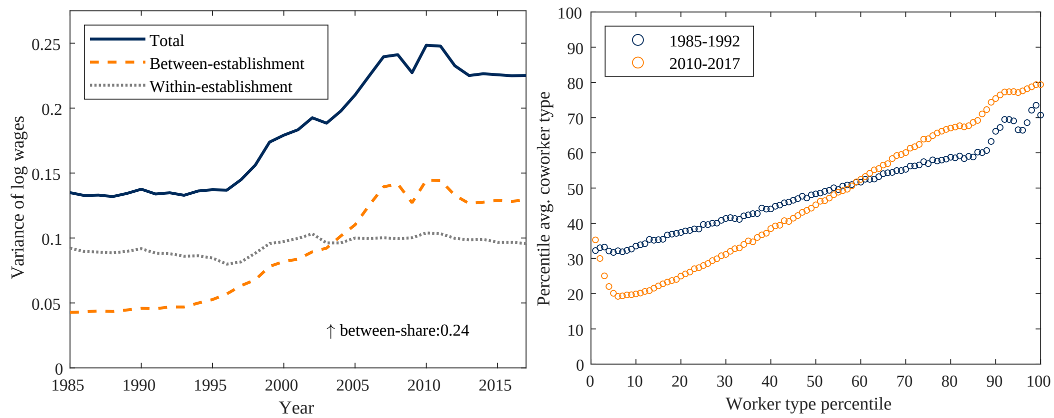
Notes. Panel 3a is based on the BIBB data and depicts the average share of complex, or abstract non-routine, tasks in individual workers' set of activities for five points in time and distinguishing between three education groups. Panel 3b is sourced from the American Board of Medical Specialities. For 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.

skills, direct, incontrovertible evidence for rising skill specificity does not yet exist.[28] Instead, I rely on several pieces of suggestive evidence and a structural approach to estimating χ, which together paint a consistent picture. Second, regarding interpretation, I do not attempt to disentangle technological shifts in task composition and human capital acquisition decisions as drivers of rising χ, as they are intricately intertwined. Automation exemplifies this interconnection. Automation displaces humans from routine tasks, leaving us to handle more complex ones (Acemoglu and Restrepo, 2018). But learning these complex tasks likely involves greater fixed learning costs, too, so any individual will acquire the requisite skills for only a subset of complex tasks (Rosen, 1983).

4.2. *A model-based explanation for the "firming up" of inequality*

An extensive empirical literature documents that, across many advanced economies, a prominent feature of wage dispersion and its rise since the 1980s is the large and increased share attributable to between-firm differences in pay (see Section 1). My theory offers a

[28] Freund and Mann (2025) develop a method to structurally estimate the distribution of task-specific skills, which when applied to historical data could enable more direct measurement of trends in specificity.



(a) Wage inequality (b) Coworker sorting

FIGURE 4.—Firming up of wage inequality & rising coworker sorting in Germany

Notes. The left panel shows the variance of log (residual) wages, decomposed into between-employer (the person-weighted variance of firm-level average log wages) and within-employer components. The indicated change in the between-share compares the average over 1985-1992 and 2010-2017. The right panel plots, for any percentile of the worker type distribution the percentile rank of the average coworker type. For visual clarity, types are grouped into 50 cells, then coworker quality is computed for each cell.

parsimonious explanation: Growing skill specificity has amplified talent complementarities, increasing sorting and, thus, firm-level wage inequality.

Empirical evolution of labor market inequality in Germany. I start by briefly revisiting the evolution of wage inequality in Germany to provide a quantitative reference point. Figure 4a presents the yearly total variance of log wages (solid line) and decomposes it into between-establishment (dashed) and within-establishment (dotted) components. The total variance rose from an average of 0.13 in 1985-1992 to 0.23 in 2010-2017. Rising between-firm gaps account for 90% of this increase, so the between-employer share of total variance rose from 33% to 57%.

Concurrently, highly productive workers have increasingly clustered in the same workplaces, segregated from less productive workers. Figure 4b visualizes this shift, based on worker and representative coworker types constructed separately for five sample periods. The headline finding is that coworker correlation has risen from an average of 0.38 during 1985-1992 to 0.64 during 2010-2017, a 68% increase. When workers are ranked within occupation, the increase is slightly smaller (+0.16 vs. +0.26) but still amounts to a 50% increase – top-performing workers within each occupation have become increasingly likely to work together. Appendix O Table D.1 provides details.

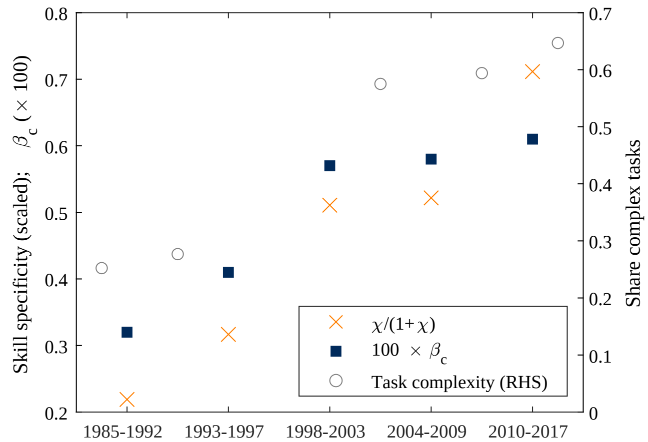
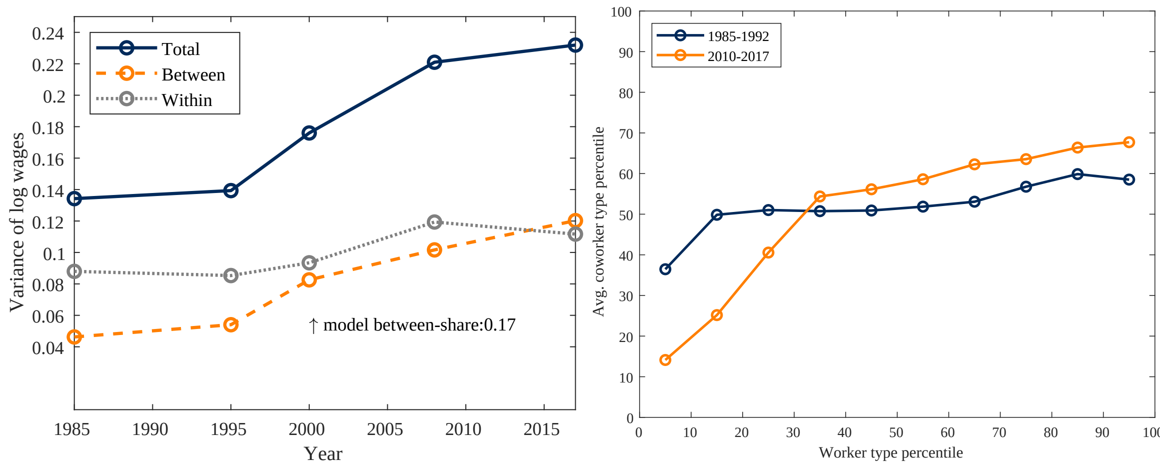


FIGURE 5.—Skill specificity has increased alongside task complexity

Notes. This figure reports the estimated value of skill specificity χ (shown as $\frac{\chi}{\chi+1}$) alongside the point estimate for the coefficient β_c ($\times$ 100) estimated from regression (20), separately for five sample periods. The circles reproduce the average share of complex tasks from Figure 3a. The years of the survey waves and the sample split in the matched employer-employee data do not align perfectly, so the task measures are placed approximately at the mid-points of the closest sample period.

Model calibration. To evaluate whether shifts in skill specificity χ can account for these trends, I re-calibrate the model for each of the four sample periods prior to 2010-2017. Specifically, I re-estimate the parameter vector ψ, targeting the same set of moments as before, but measured for earlier years. To discipline χ I estimate equation (20) period-by-period to recover the interaction coefficient β_c. Regarding the replacement rate, I summarize the Hartz reforms as roughly a 10% reduction in the monetary replacement rate, similar to Jung et al. (2023).

The headline estimation result is that skill specificity has intensified. Figure 5 displays the evolution of point estimates for β_c, as a reduced-form moment, and the indirectly inferred values of χ. The estimated χ rises from 0.28 to 2.47 – co-moving with task complexity. Correspondingly, the mean productivity gain associated with reallocating workers from two talent-mixed teams into two talent-homogeneous teams rises from 2% to 7%. The full set of parameter estimates appears in Appendix O Table D.2. Notably, a_0 has decreased while a_1 has increased over time, consistent with technological change that has made output more elastic with respect to talent. Additionally, increases in both job arrival and job separation rates are consistent with declining search frictions (cf. Martellini and Menzio, 2020), plausibly due to the emergence of online job portals.



(a) Variance decomposition (b) Coworker sorting

FIGURE 6.—Model-implied evolution of wage inequality and sorting tracks data

Notes. In the left panel, the solid lines indicate the model-predicted variance of log wages, decomposed into between- and within-employer components. The model-generated between-within decomposition is corrected for a mechanical bias, as described in Appendix O.C.1. The right panel plots the average coworker talent type (vertical axis) for each worker talent type (horizontal axis), for two separate sample periods.

Model-implied trends in firm-level dispersion. The calibrated model replicates the empirically recorded evolution of between- and within-firm inequality. Figure 6a depicts the model-implied decomposition of the variance of log wages for each sample period. Though only the total variance of log wages was targeted, the model captures a large fraction of the empirically recorded increase in the between-firm share (+0.17 vs. +0.24). The model also explains widening labor productivity gaps between firms (Appendix O Figure D.5) – a salient trend in the firm dynamics literature (e.g., Autor et al., 2020, Sorkin and Wallskog, 2021). These shifts mirror rising talent sorting, with Figure 6b providing the model counterpart to Figure 4b. These figures also help clarify where the model fails to capture the data: it understates the level of sorting in earlier years and hence overstates the total rise in sorting (+0.33 vs. +0.26).

Counterfactual-based decomposition. What portion of the model-implied rise in the (statistical) role of firms in explaining wage inequality can be attributed to greater skill specificity? The answer is not obvious since all estimated parameters changed. For instance, in the theory, a higher job arrival rate amplifies sorting, as workers can accept offers more selectively. Moreover, as long as the original equilibrium exhibits positive sorting,

Δ *Between-firm share*	Partial contribution				Total Δ
	χ	(a_0, a_1)	b	$(\delta_0, \delta_1, \lambda_u)$	
Baseline	0.13	0.05	-0.02	0.01	0.17
	77%	26%	-10%	8%	100%
Within-occ. ranking	0.18	0.04	-0.02	0.01	0.22
	84%	20%	-8%	5%	100%
Avg. within-industry	0.19	0.01	-0.01	-0.002	0.20
	99%	7%	-6%	-1%	100%

TABLE II

RISING SKILL SPECIFICITY IS AN IMPORTANT DRIVER OF THE FIRMING UP OF INEQUALITY

Notes. This table shows the results of a Shapley-Owen-Shurrocks decomposition of the model-implied change in the between-firm share of the log wage variance from 1985-1992 to 2010-2017. See Audoly et al. (2024, Appendix E) for a summary of the methodology. The first row corresponds to the baseline; the two others are described in the main text. The parameter groups (a_0, a_1) and $(\lambda_u, \delta_0, \delta_1)$ were each treated as a single factor. Contributions may not exactly sum to total, respectively 100%, due to rounding.

technological change that amplifies the return to talent can mechanically lead to greater between-firm wage inequality.[29]

The structural model facilitates an answer through counterfactual exercises. I perform a Shapley-Owen-Shurrocks decomposition (Shorrocks, 2013) of the *change* in the between-firm share of the variance of log wages between 1985-1992 and 2010-2017. This decomposition involves evaluating all possible combinations of parameters and weighting them appropriately to quantify each parameter's marginal effect. This approach is attractive because it provides an exact additive decomposition and is permutation-invariant. Table II summarizes the results.

The key takeaway is that rising skill specificity χ is an important driver of the firming up of inequality. Under the baseline calibration, it contributed 0.13 of the 0.17 point increase in the between share, or 77%. This corresponds to just more than half (55%) of the empirically observed increase. For comparison, shifts in parameters a_0 and a_1 – capturing an amplified return to talent – jointly contributed 26% to the model-implied rise. The decline

[29] Within the AKM-framework, Song et al. (2019) attribute 9% of the increase in worker-firm sorting and 35% of rising worker segregation in the U.S. to mechanical effects due to increased return to talent.

in search frictions was also a positive contributor, accounting for 8% of the increase.[30] The deteriorated unemployment outside option, b_1, operated in the opposite direction by pushing workers to accept lower-quality matches.

Sensitivity analysis & Caveats. The headline result is robust when addressing two important concerns relating to occupational and industry dynamics.

First, the measurement underlying the baseline figures may be confounded by outsourcing dynamics or related shifts in the boundary of the firm. To address this possibility, I re-estimate the model parameters targeting β_c estimates obtained when ranking workers within-occupation rather than economy-wide, ensuring that results are not affected by changes in the sorting of high- and low-wage occupations into different establishments. OA Table D.3 displays the parameter estimates. The model-predicted increase in the between-firm share is equal to 22 p.p., of which – as shown in Table II – the great majority is attributed to a rise in χ.

Second, the analysis so far has abstracted from industry-level differences in the production function (e.g., Haltiwanger and Spletzer, 2020). I therefore repeat the calibration, this time targeting the within-industry average of the variance of log wages and the mean industry-level β_c estimate (see OA Table D.4 for parameter estimates). As shown in Table II, counterfactual exercises under this alternative specification imply the rise in skill specificity explains almost the entirety of the empirically observed rise in the within-industry between-firm share of wage dispersion.[31]

Person-level inequality and comparison to Kremer (1993). The theory also reveals that rising coworker complementarities do *not*, by themselves, exacerbate person-level inequality. By itself, the induced rise in talent sorting increases wage dispersion between firms but reduces it within firms.[32] Rising specificity and talent-biased technological change *jointly*

---

[30] The relatively limited contribution made by declining search frictions is consistent with Kantenga and Law's (2016) study. A caveat regarding this exercise is that the transition rates are treated as exogenous in the structural model, despite being plausibly endogenous to the cost of mismatch.

[31] As mentioned above, the model understates the degree of sorting in earlier years, when χ is estimated to be quite low (Table D.2). This may lead the model to overstate the contribution of χ to rising between-firm inequality.

[32] Extensions of the model could introduce channels through which greater sorting amplifies person-level inequality. For instance, sorting may dynamically increase lifetime inequality through coworker learning effects (Jarosch et al., 2021, Herkenhoff et al., 2024) – a possibility worth exploring in future research.

lead to rising person-level inequality through heightened firm-level wage dispersion. This result differs notably from Kremer (1993), which is sometimes understood as saying that complementarities amplify inequality. However, this prediction hinges on increasing returns to team quality (cf. Kremer, 1993, III) – an assumption I do not make, thus isolating the role of complementarities. Indeed, changes in sorting play no role at all in Kremer (1993), wherein matching is perfectly assortative for *any* degree of complementarity.

4.3. *Gains from specialization with endogenous complementarities*

Rising specialization over the course of development is widely understood to yield aggregate productivity gains. The microfounded approach reveals a crucial nuance: "Smithian" growth is limited by endogenously increasing vulnerability to mismatch when labor market frictions prevent workers from finding colleagues with complementary expertise.

Intuition. In the microfounded production function, higher skill specificity χ boosts potential productivity but simultaneously increases vulnerability to coworker mismatch. In perfect labor markets, this endogenous link is without consequences – mismatch never exists in equilibrium – but it becomes operative under information frictions. Ignoring the endogenous interaction between specificity, productivity, complementarities, and frictions will, thus, overstate Smithian productivity gains.[33]

Quantitative exploration. Figure 7 illustrates how labor productivity varies with χ under different worker allocations. I hold all other parameters constant at their baseline values and normalize labor productivity to unity at the lowest χ value. The solid line represents the estimated German economy. The dashed line shows the productivity frontier, featuring perfect talent sorting for any value of χ, holding constant the marginal distribution of workers in teams $\phi(x)$ – so only the matching changes – and setting $\xi_{il} = \frac{1}{2}$ for every team (i, l). The dotted line depicts a highly frictional economy, fixing talent matching to be random, again holding fixed $\phi(x)$, and $\xi_{il} = \frac{1}{4}$.

[33] "Rising specialization" could also be interpreted as deepening division of labor with fixed skill specialization. Numerical exercises using an extension of the production model featuring communication frictions suggest that, at least qualitatively, the same endogenous link between productivity potential and complementarities would operate in such a setting.

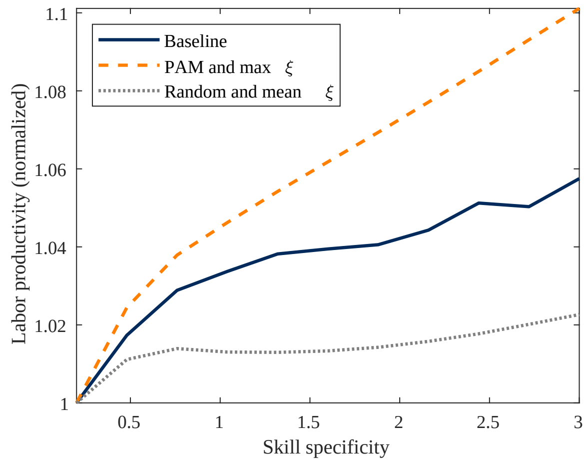


FIGURE 7.—With endogenous complementarities, frictions limit gains from specialization

Notes. The solid line shows (normalized) labor productivity in the equilibrium of the economy for different χ values. Letting $\phi(x)$ denote the unconditional density of talent in teams, the dashed line captures $\frac{1}{2}\int f(x,x,\frac{1}{2})\phi(x)dx$; the dotted line corresponds to $\frac{1}{2}\int f(x,x',\frac{1}{4})\phi(x)\phi(x')dx'dx$.

The figure reveals that realized productivity gains from greater specialization are substantially constrained by the interaction of endogenous complementarities and labor market frictions. While in the baseline economy – where equilibrium talent sorting and horizontal distances in teams endogenously increase in response to rising χ – roughly two-thirds of potential gains are realized, this fraction falls to approximately one-quarter in the "highly frictional" economy. Additionally, Appendix A.G.1 compares outcomes with fixed versus endogenous complementarities.

Discussion. Simply put, specialization and labor market quality are complements in a broad macroeconomic sense. In the context of recent literature, this insight seems particularly relevant for understanding growth obstacles in developing countries, where labor markets typically feature severe information frictions (Donovan et al., 2023). For instance, the theory rationalizes cross-country results results by Bandiera et al. (2024) indicating that improved matching is crucial for unlocking gains from frontier technology adoption. Going beyond the model, the theory may also explain why specialization remains limited in poorer countries (Bassi et al., 2023, Atencio-De-Leon et al., 2024): anticipating that labor market frictions would constrain specialization benefits, workers may rationally choose broader skill portfolios. Extending the model to allow for endogenous specialization is challenging, however, and so is left for future research.

5. CONCLUSION: INSIGHTS, SHORTCOMINGS & FUTURE WORK

Takeaways. The paper's primary contribution is a parsimonious theory of firms centered on a simple idea: Production typically involves division of labor among workers with heterogeneous, specialized skills. The implications of this view are summarized in a microfounded firm-level production function whose shape is governed by the degree of skill specificity and which remains tractable enough to analyze economy-wide outcomes. The paper can thus be read broadly as integrating a classic view of the firm with macroeconomic analysis.

More narrowly, the paper delivers five key insights. First, when production involves division of labor among heterogeneously skilled workers, their productivity is non-separable. Optimally, teams comprise workers similar in talent but specialized in different tasks. Second, *both* potential productivity and talent complementarities endogenously increase with the degree of skill specificity. Third, firm-level dispersion in productivity and average pay naturally arises from sorting induced by these endogenous talent complementarities. Indeed, fourth, growing skill specificity since the mid-1980s has amplified complementarities, contributing to the "firming up" of inequality. Fifth, "Smithian" productivity gains from specialization are limited by the quality of labor markets.

Limitations & future research. I close by briefly revisiting strong modeling assumptions, discussing the resulting shortcomings, and highlighting avenues for future research.

First, workers were assumed not to search on the job (OJS). Relaxing this assumption would presumably enable the model to better match empirical sorting patterns, which, as already mentioned, it currently fits imperfectly in earlier years. An extension is feasible in principle, albeit at the cost of introducing ξ_{il} as an additional state variable.

Second, team size was limited to two workers. While sufficient to deliver the key insights, this assumption limits how much persistence in firm-level productivity the model can generate and precludes analyzing firm growth patterns. Appendix O.A.3 discusses ways to generalize the microfounded production function. While beyond the scope of this paper, a "large team" version would allow quantifying the role of team quality in explaining large growth rate differences among young firms (Sterk et al., 2021) – an avenue that reduced-form evidence suggests is promising (Criscuolo et al., 2024). Furthermore, introducing *both* large firms and OJS would also facilitate assessing quantitatively to what extent AKM-based (Abowd et al., 1999) estimates of "firm pay premia" reflect coworker effects.

Third, inside the firm, tasks were always optimally assigned. This assumption facilitated closed-form expressions for output, but relaxing it would open the door to studying additional implications of specialization. For instance, recent evidence suggests some managers are better at coordinating their workforce than others (e.g., Minni, 2022). An extension of my framework that introduces an explicit managerial role could capture this important feature of the data. Intuitively, deepening skill specialization should make management more important, constituting a source of "manager-skill biased technological change" that helps explain rising managerial pay shares.

The theory has many other applications. To give just one example, in science and innovation, a growing "burden of knowledge" and consequent specialization mean that progress increasingly hinges on teamwork (Jones, 2009). The theory highlights a challenge: Since production complementarities and therefore skill bottlenecks grow with specialization, innovation increasingly depends on being able to form teams of knowledge workers who have different specialties but similar talent levels. This carries implications for the optimal allocation of researchers and the pace of scientific progress.

REFERENCES

ABOWD, JOHN M., FRANCIS KRAMARZ, AND DAVID N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67 (2), 251–333. [40]

ACEMOGLU, DARON AND DAVID AUTOR (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, Elsevier, vol. 4, 1043–1171. [7]

ACEMOGLU, DARON AND PASCUAL RESTREPO (2018): "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 108 (6), 1488–1542. [2, 6, 13, 31]

ADENBAUM, JACOB (2022): "Endogenous Firm Structure and Worker Specialization," *mimeo*. [6]

ADHVARYU, ACHYUTA, VITTORIO BASSI, ANANT NYSHADHAM, AND JORGE A. TAMAYO (2020): "No Line Left Behind: Assortative Matching Inside the Firm," . [6]

AKCIGIT, UFUK, SANTIAGO CAICEDO, ERNEST MIGUELEZ, STEFANIE STANTCHEVA, AND VALERIO STERZI (2018): "Dancing with the Stars: Innovation Through Interactions," . [5]

ALCHIAN, ARMEN A. AND HAROLD DEMSETZ (1972): "Production, Information Costs, and Economic Organization," *The American Economic Review*, 62 (5), 777–795. [2, 13]

ARROW, KENNETH JOSEPH (1974): *The Limits of Organization*, New York: Norton, [1st ed.] ed. [2]

ATALAY, ENGHIN, PHAI PHONGTHIENGTHAM, SEBASTIAN SOTELO, AND DANIEL TANNENBAUM (2020): "The Evolution of Work in the United States," *American Economic Journal: Applied Economics*, 12 (2), 1–34. [2]

ATENCIO-DE-LEON, ANDREA, MUNSEOB LEE, AND CLAUDIA MACALUSO (2024): "Does Turnover Inhibit Specialization? Evidence from a Skill Survey in Peru," *American Economic Review: Insights*. [38]

AUDOLY, RICHARD, RORY MCGEE, SERGIO OCAMPO-DIAZ, AND GONZALO PAZ-PARDO (2024): "The Life-Cycle Dynamics of Wealth Mobility," . [35]

AUTOR, DAVID, CAROLINE CHIN, ANNA SALOMONS, AND BRYAN SEEGMILLER (2024): "New Frontiers: The Origins and Content of New Work, 1940–2018*," *The Quarterly Journal of Economics*, 139 (3), 1399–1465. [2, 7]

AUTOR, DAVID, DAVID DORN, LAWRENCE F KATZ, CHRISTINA PATTERSON, AND JOHN VAN REENEN (2020): "The Fall of the Labor Share and the Rise of Superstar Firms," *The Quarterly Journal of Economics*, 135 (2), 645–709. [34]

AUTOR, DAVID H., FRANK LEVY, AND RICHARD J. MURNANE (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 118 (4), 1279–1333. [7]

BALEY, ISAAC, ANA FIGUEIREDO, AND ROBERT ULBRICHT (2022): "Mismatch Cycles," *Journal of Political Economy*, 130 (11), 2943–2984. [7]

BANDIERA, ORIANA, AHMED ELSAYED, ANTON HEIL, AND ANDREA SMURRA (2022): "Economic Development and the Organisation of Labour: Evidence from the Jobs of the World Project," . [2, 30]

BANDIERA, ORIANA, ANANYA KOTIA, ILSE LINDENLAUB, CHRISTIAN MOSER, AND ANDREA PRAT (2024): "Meritocracy across Countries," . [38]

BARRO, ROBERT J. AND JONG WHA LEE (2013): "A New Data Set of Educational Attainment in the World, 1950–2010," *Journal of Development Economics*, 104, 184–198. [Online Appendix - p.9]

BARTH, ERLING, ALEX BRYSON, JAMES C. DAVIS, AND RICHARD FREEMAN (2016): "It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States," *Journal of Labor Economics*, 34 (S2), S67–S97. [7]

BASSI, VITTORIO, JUNG HYUK LEE, ALESANDRA PETER, TOMMASO PORZIO, RITWIKA SEN, AND ESAU TUGUME (2023): "Self-Employment within the Firm," *mimeo*. [6, 38]

BECKER, GARY S. AND KEVIN M. MURPHY (1992): "The Division of Labor, Coordination Costs, and Knowledge," *The Quarterly Journal of Economics*, 107 (4), 1137–1160. [2]

BILAL, ADRIEN, NIKLAS ENGBOM, SIMON MONGEY, AND GIOVANNI L. VIOLANTE (2022): "Firm and Worker Dynamics in a Frictional Labor Market," *Econometrica*, 90 (4), 1425–1462. [17]

BLOESCH, JUSTIN, BIRTHE LARSEN, AND BLEDI TASKA (2022): "Which Workers Earn More at Productive Firms? Position Specific Skills and Individual Worker Hold-up Power," *mimeo*. [6]

BOEHM, JOHANNES AND EZRA OBERFIELD (2022): "Growth and the Fragmentation of Production," *mimeo*. [Online Appendix - p.10]

BOERMA, JOB, ALEH TSYVINSKI, AND ALEXANDER ZIMIN (2021): "Sorting with Team Formation," Tech. Rep. w29290, National Bureau of Economic Research, Cambridge, MA. [5]

BOROVICKOVA, KATARINA AND ROBERT SHIMER (2024): "Assortative Matching and Wages: The Role of Selection," *mimeo*. [22]

CAIRÓ, ISABEL AND TOMAZ CAJNER (2018): "Human Capital and Unemployment Dynamics: Why More Educated Workers Enjoy Greater Employment Stability," *The Economic Journal*, 128 (609), 652–682. [29]

CALIENDO, LORENZO, GIORDANO MION, LUCA DAVID OPROMOLLA, AND ESTEBAN ROSSI-HANSBERG (2020): "Productivity and Organization in Portuguese Firms," *Journal of Political Economy*. [6]

CAPLIN, ANDREW, MINJOON LEE, SØREN LETH-PETERSEN, JOHAN SAEVERUD, AND MATTHEW D. SHAPIRO (2022): "How Worker Productivity and Wages Grow with Tenure and Experience: The Firm Perspective," . [24]

CARD, DAVID, JÖRG HEINING, AND PATRICK KLINE (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *The Quarterly Journal of Economics*, 128 (3), 967–1015. [7, Online Appendix - p.11, Online Appendix - p.12, Online Appendix - p.16]

CAUNEDO, JULIETA, ELISA KELLER, AND YONGSEOK SHIN (2023): "Technology and the Task Content of Jobs across the Development Spectrum," *The World Bank Economic Review*, 37 (3), 479–493. [30]

CHADE, HECTOR AND JAN EECKHOUT (2020): "Competing Teams," *The Review of Economic Studies*, 87 (3), 1134–1173. [5]

CORNELISSEN, THOMAS, CHRISTIAN DUSTMANN, AND UTA SCHÖNBERG (2017): "Peer Effects in the Workplace," *American Economic Review*, 107 (2), 425–456. [23]

CRISCUOLO, CHIARA, LUKAS B. FREUND, AND PETER N. GAL (2024): "Human Capital at Work: Five Facts about the Role of Skills for Firm Productivity, Growth, and Wage Inequality," *OECD Publishing*, No. 36. [24, 39]

DAUTH, WOLFGANG AND JOHANN EPPELSHEIMER (2020): "Preparing the Sample of Integrated Labour Market Biographies (SIAB) for Scientific Analysis: A Guide," *Journal for Labour Market Research*, 54 (1), 1–14. [Online Appendix - p.11]

DE HAAN, L. (1984): "A Spectral Representation for Max-stable Processes," *The Annals of Probability*, 12 (4), 1194–1204. [Online Appendix - p.9]

DEB, SHUBHDEEP, JAN EECKHOUT, ASEEM PATEL, AND LAWRENCE WARREN (2024): "Walras–Bowley Lecture: Market Power and Wage Inequality," *Econometrica*, 92 (3), 603–636. [5]

DEMING, DAVID J. (2017): "The Growing Importance of Social Skills in the Labor Market," *The Quarterly Journal of Economics*, 132 (4), 1593–1640. [6, 24, 30]

DONOVAN, KEVIN, WILL JIANYU LU, AND TODD SCHOELLMAN (2023): "Labor Market Dynamics and Development," *The Quarterly Journal of Economics*, qjad019. [3, 38]

DVORKIN, MAXIMILIANO AND ALEXANDER MONGE-NARANJO (2019): "Occupation Mobility, Human Capital and the Aggregate Consequences of Task-Biased Innovations," *St. Louis Federal Reserve Bank Working Paper 2019-013C*. [13]

EATON, JONATHAN AND SAMUEL KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70 (5), 1741–1779. [3, 6, 9, 12, 14, Online Appendix - p.1]

EECKHOUT, JAN AND PHILIPP KIRCHER (2011): "Identifying Sorting—In Theory," *The Review of Economic Studies*, 78 (3), 872–906. [22]

FREUND, LUKAS B. (2023): "Superstar Teams: The Micro Origins and Macro Implications of Coworker Complementarities," *Available at SSRN 4312245*. [24]

FREUND, LUKAS B. AND LUKAS MANN (2025): "For Whom the Bot Tolls: Specialization and the Earnings Effects of AI," *mimeo*. [31]

GABAIX, XAVIER AND AUGUSTIN LANDIER (2008): "Why Has CEO Pay Increased So Much?" *The Quarterly Journal of Economics*, 123 (1), 49–100. [6]

GARICANO, LUIS (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108 (5), 874–904. [2, 6]

GARICANO, LUIS AND ESTEBAN ROSSI-HANSBERG (2006): "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, 121 (4), 1383–1435. [6]

GAUTIER, PIETER A., COEN N. TEULINGS, AND AICO VAN VUUREN (2010): "On-the-Job Search, Mismatch and Efficiency," *The Review of Economic Studies*, 77 (1), 245–272. [7]

GENEST, CHRISTIAN, IVAN KOJADINOVIC, JOHANNA NEŠLEHOVÁ, AND JUN YAN (2011): "A Goodness-of-Fit Test for Bivariate Extreme-Value Copulas," *Bernoulli*, 17 (1), 253–275. [Online Appendix - p.9]

GRIGSBY, JOHN (2023): "Skill Heterogeneity and Aggregate Labor Market Dynamics," *mimeo*. [30]

HAGEDORN, MARCUS, TZUO HANN LAW, AND IOURII MANOVSKII (2017): "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 85 (1), 29–65. [7, 22, Online Appendix - p.12]

HALTIWANGER, JOHN C AND JAMES R SPLETZER (2020): "Between Firm Changes in Earnings Inequality: The Dominant Role of Industry Effects," Working Paper 26786, National Bureau of Economic Research. [36]

HELPMAN, ELHANAN, OLEG ITSKHOKI, MARC-ANDREAS MUENDLER, AND STEPHEN J. REDDING (2017): "Trade and Inequality: From Theory to Estimation," *The Review of Economic Studies*, 84 (1 (298)), 357–405. [7]

HERKENHOFF, KYLE, JEREMY LISE, GUIDO MENZIO, AND GORDON M. PHILLIPS (2024): "Production and Learning in Teams," *Econometrica*, 92 (2), 467–504. [2, 3, 5, 21, 26, 36]

HICKS, JOHN (1932): *The Theory of Wages*, London: Macmillan. [14]

HOUTHAKKER, H. S. (1955): "The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis," *The Review of Economic Studies*, 23 (1), 27–31. [13]

HOXBY, CAROLINE (2000): "Peer Effects in the Classroom: Learning from Gender and Race Variation," . [Online Appendix - p.17]

JÄGER, SIMON AND JÖRG HEINING (2022): "How Substitutable Are Workers? Evidence from Worker Deaths," *mimeo*. [28]

JAROSCH, GREGOR, EZRA OBERFIELD, AND ESTEBAN ROSSI-HANSBERG (2021): "Learning From Coworkers," *Econometrica*, 89 (2), 647–676. [2, 5, 21, 36]

JONES, BENJAMIN F. (2009): "The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?" *The Review of Economic Studies*, 76 (1), 283–317. [30, 40]

JONES, CHARLES I. (2005): "The Shape of Production Functions and the Direction of Technical Change," *The Quarterly Journal of Economics*, 120 (2), 517–549. [13]

JUNG, PHILIP, PHILIPP KORFMANN, AND EDGAR PREUGSCHAT (2023): "Optimal Regional Labor Market Policies," *European Economic Review*, 152, 104318. [33]

KANTENGA, KORY AND TZUO-HANN LAW (2016): "Sorting and Wage Inequality," *mimeo*. [36]

KOENIG, FELIX, ALAN MANNING, AND BARBARA PETRONGOLO (2021): "Reservation Wages and the Wage Flexibility Puzzle," *SSRN Electronic Journal*. [27]

KOHLHEPP, JACOB (2022): "The Inner Beauty of Firms," *mimeo*. [6]

KREMER, MICHAEL (1993): "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics*, 108 (3), 551–575. [2, 6, 14, 36, 37]

KREMER, MICHAEL AND ERIC MASKIN (1996): "Wage Inequality and Segregation by Skill," Working Paper 5718, National Bureau of Economic Research. [6]

KUHN, MORITZ, JINFENG LUO, IOURII MANOVSKII, AND XINCHENG QIU (2023): "Coordinated Firm-Level Work Processes and Macroeconomic Resilience," *Journal of Monetary Economics*. [6]

LAGOS, RICARDO (2006): "A Model of TFP," *The Review of Economic Studies*, 73 (4), 983–1007. [13]

LENTZ, RASMUS AND DALE T. MORTENSEN (2010): "Labor Market Models of Worker and Firm Heterogeneity," *Annual Review of Economics*, 2 (1), 577–602. [7]

LIND, NELSON AND NATALIA RAMONDO (2023): "Trade with Correlation," *American Economic Review*, 113 (2), 317–353. [6, Online Appendix - p.1]

LINDENLAUB, ILSE (2017): "Sorting Multidimensional Types: Theory and Application," *The Review of Economic Studies*, 84 (2), 718–789. [2, 7]

LINDENLAUB, ILSE AND FABIEN POSTEL-VINAY (2023): "Multidimensional Sorting under Random Search," *Journal of Political Economy*, 131 (12), 3497–3539. [7]

LOPES DE MELO, RAFAEL (2018): "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," *Journal of Political Economy*, 126 (1), 313–346. [19]

LUCAS, ROBERT (1978): "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9 (2), 508–523. [6]

MANSKI, CHARLES F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 60 (3), 531–542. [23]

MARIMON, RAMON AND FABRIZIO ZILIBOTTI (1999): "Unemployment vs. Mismatch of Talents: Reconsidering Unemployment Benefits," *The Economic Journal*, 109 (455), 266–291. [7]

MARTELLINI, PAOLO AND GUIDO MENZIO (2020): "Declining Search Frictions, Unemployment, and Growth," *Journal of Political Economy*, 128 (12), 4387–4437. [7, 33]

——— (2021): "Jacks of All Trades and Masters of One: Declining Search Frictions and Unequal Growth," *American Economic Review: Insights*, 3 (3), 339–352. [24]

MINNI, VIRGINIA (2022): "Making the Invisible Hand Visible: Managers and the Allocation of Workers to Jobs," *Job Market Paper*. [6, 40]

NEFFKE, FRANK M. H. (2019): "The Value of Complementary Co-Workers," *Science Advances*, 5 (12), eaax3370. [30]

NIX, EMILY (2020): "Learning Spillovers in the Firm," Tech. rep., Working Paper. [23, Online Appendix - p.15]

OCAMPO, SERGIO (2022): "A Task-Based Theory of Occupations with Multidimensional Heterogeneity," Tech. rep., mimeo. [6]

PEARCE, JEREMY (2022): "Idea Production and Team Structure," *Working paper*. [5]

PORZIO, TOMMASO (2017): "Cross-Country Differences in the Optimal Allocation of Talent and Technology," *mimeo*. [6]

ROHRBACH-SCHMIDT, DANIELA AND MICHAEL TIEMANN (2013): "Changes in Workplace Tasks in Germany—Evaluating Skill and Task Measures," *Journal for Labour Market Research*, 46 (3), 215–237. [Online Appendix - p.13]

ROSEN, SHERWIN (1978): "Substitution and Division of Labour," *Economica*, 45 (179), 235–250. [2]

——— (1983): "Specialization and Human Capital," *Journal of Labor Economics*, 1 (1), 43–49. [31]

SAINT-PAUL, GILLES (2001): "On the Distribution of Income and Worker Assignment under Intrafirm Spillovers, with an Application to Ideas and Networks," *Journal of Political Economy*, 109 (1), 1–37. [6]

SATTINGER, MICHAEL (1993): "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, 31 (2), 831–880. [6]

SHANBHAG, D. N. AND M. SREEHARI (1977): "On Certain Self-Decomposable Distributions," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 38 (3), 217–222. [Online Appendix - p.10]

SHAO, LIN, FAISAL SOHAIL, AND EMIRCAN YURDAGUL (2023): "Are Working Hours Complements in Production?" *Journal of Economic Dynamics and Control*, 153, 104696. [28]

SHORROCKS, ANTHONY F. (2013): "Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value," *The Journal of Economic Inequality*, 11 (1), 99–126. [35]

SONG, JAE, DAVID J. PRICE, FATIH GUVENEN, NICHOLAS BLOOM, AND TILL VON WACHTER (2019): "Firming Up Inequality," *The Quarterly Journal of Economics*, 134 (1), 1–50. [3, 7, 35]

SORKIN, ISAAC AND MELANIE WALLSKOG (2021): "The Slow Diffusion of Earnings Inequality," *Working Paper*. [34]

SPITZ-OENER, ALEXANDRA (2006): "Technical Change, Job Tasks, and Rising Educational Demands: Looking Outside the Wage Structure," *Journal of Labor Economics*, 24 (2), 235–270. [24, Online Appendix - p.13]

STERK, VINCENT, PETR SEDLÁČEK, AND BENJAMIN PUGSLEY (2021): "The Nature of Firm Growth," *American Economic Review*, 111 (2), 547–579. [39]

VOM BERGE, PHILIPP, LISA SCHMIDTLEIN, STEFAN SETH, TOBIAS GRAF, STEPHAN GRIESSEMER, STEFFEN KAIMER, STEFFEN KOEHLER, CLAUDIA LEHNER, MARTINA OERTEL, AND CHRISTIAN SEYSEN (2020): "The Sample of Integrated Employer-Employee Data (SIEED): SIEED 7518, Version 1," *Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB)*. [Online Appendix - p.10]

Online Supplemental Appendix for "Superstar Teams"

Any references to sections, equations, figures, tables, assumptions, propositions, lemmas or corollaries that are not preceded by a capital letter refer to the main article.

APPENDIX A THEORETICAL APPENDIX

A.1 Production function: proofs and derivations

Following the introduction of auxiliary definitions and lemmas (Section A.1.1), this section states Lemma A.3 and proves Proposition 1. I exploit that the production problem is cast so as to parallel the structure of an Eaton and Kortum (2002) style trade model, allowing for correlation across producers' task-specific skills (Lind and Ramondo, 2023), and formulated as a planner problem.

Throughout I drop the (i,l) subscripts for ξ, with the understanding that an $n=2$ team has been formed. The derivations are also valid for $n>2$ if the pairwise distances are homogeneous within the team (see Appendix O.A.3).

A.1.1 Auxiliary definitions and lemmas

DEFINITION A.1—Fréchet distribution: A random variable Z has a Fréchet distribution with scale parameter $s>0$ and shape parameter $\theta>0$ if $\mathrm{P}[Z\leq z]=e^{-\left(\frac{z}{s}\right)^{-\theta}}$.

LEMMA A.1: *Let the random variable Z be Fréchet-distributed with scale parameter s and shape parameter $\theta>1$. Then $\mathbb{E}[Z]=x\Gamma\left(1-\frac{1}{\theta}\right)$, where $\Gamma(r)=\int_0^\infty t^{r-1}e^{-t}dt$ is the Gamma function.*

PROOF:

$$\mathbb{E}[Z]=\int_0^\infty z\frac{\partial}{\partial z}\mathrm{P}[Z\leq z]\,dz=\int_0^\infty z\left(\theta\frac{1}{s}\left(\frac{z}{s}\right)^{-\theta-1}e^{-\left(\frac{z}{s}\right)^{-\theta}}\right)dz$$
$$=\int_0^\infty ze^{-\left(\frac{z}{s}\right)^{-\theta}}\left(\theta s^{\theta}\left(z\right)^{-\theta-1}\right)dz=s\int_0^\infty t^{-1/\theta}e^{-t}dt=s\Gamma(1-\frac{1}{\theta}).$$

Q.E.D.

DEFINITION A.2—Multivariate Fréchet distribution: A random vector $(Z_1,...,Z_n)$ has a multivariate Fréchet distribution with a vector of scale parameters $(s_1,...,s_n)\in\mathbb{R}_{++}$, shape parameter $\alpha > 0$ and correlation parameter $\xi \in (0, 1]$, if

$$\mathrm{P}\left[Z_1 \leq z_1, \ldots, Z_n \leq z_n\right] = \exp\left[-\left(\sum_{i=1}^{n}\left(\frac{z_i}{s_i}\right)^{-\frac{\alpha}{\xi}}\right)^{\xi}\right].$$

The term "multivariate Fréchet distribution" is used here to refer to the specific, symmetric dependence structure above; in principle a different or more flexible dependence structure could be considered.

LEMMA A.2: *Let the random vector $(Z_1, ..., Z_n)$ be distributed multivariate-Fréchet with scale parameters $(s_1, ..., s_n)$, shape parameter α and correlation parameter ξ. Then for any $B_i > 0$, $i = 1, ..., n$ and $\beta > 0$, the random variable $\max_{i=1,...n} B_i Z_i^{\beta}$ is Fréchet distributed with scale $\left(\sum_{i=1}^{n}\left(s_i^{\alpha} B_i^{\alpha/\beta}\right)^{1/\xi}\right)^{\xi\beta/\alpha}$ and shape α/β.*

PROOF:

$$\begin{aligned}\mathrm{P}\left[\max_{i=1,\ldots,n} B_i Z_i^{\beta} \leq p\right] &= \mathrm{P}\left[Z_1 \leq (p/B_1)^{1/\beta}, ..., Z_n \leq (p/B_n)^{1/\beta}\right] \\ &= \exp\left[-\left(\sum_{i=1}^{n}\left(s_i^{\alpha} B_i^{\alpha/\beta}\right)^{\frac{1}{\xi}}\right)^{\xi} p^{-\alpha/\beta}\right].\end{aligned}$$

Q.E.D.

A.1.2 Lemma A.3: statement and proof

As a preliminary step, note that taking the first-order condition (FOC) for $q(\tau)$, defining $Q(\tau) := \lambda(\tau)q(\tau)$, and standard CES algebra yields expressions for task demand and the shadow cost index λ:

$$Q(\tau) = \left(\frac{\lambda(\tau)}{\lambda}\right)^{1-\eta} \lambda Y, \tag{A.1}$$

$$\lambda = \left(\int_{\mathcal{T}} \lambda(\tau)^{1-\eta} d\tau\right)^{\frac{1}{1-\eta}}. \tag{A.2}$$

LEMMA A.3: *Under the optimal task assignment, assuming $\eta > 1$ and a symmetric horizontal distance across coworker pairs, ξ:*

(i) the shadow cost index is

$$\lambda = \left(\sum_{i=1}^{n} \left(\frac{x_i}{\lambda_i^L} \right)^{\frac{1}{\chi\xi}} \right)^{-\chi\xi} ; \tag{A.3}$$

(ii) and the share of tasks for which worker i is the least-cost provider is

$$\pi_i := \boldsymbol{P}\left[\lambda_i(\tau) = \min_{k=1,..,n} \lambda_k(\tau) \right] = \frac{\left(x_i/\lambda_i^L \right)^{\frac{1}{\chi\xi}}}{\sum_{k=1}^{n} \left(x_k/\lambda_k^L \right)^{\frac{1}{\chi\xi}}}, \tag{A.4}$$

which furthermore corresponds to the fraction of the shadow value of all tasks used in final good production accounted for by i:

$$\pi_i = Q_i/Q \tag{A.5}$$

where $Q_i := \int_{\mathcal{T}} \lambda(\tau) y_i(\tau) d\tau$.

PROOF: For Part (i), start with the expression for the shadow-price index in equation (A.2) and substitute for $\lambda(\tau)$ using equation (7):

$$\lambda = \left[\int_0^\infty (\lambda(\tau))^{1-\eta} d\tau \right]^{\frac{1}{1-\eta}} = \left[\int_0^\infty \left(\min_{i=1,..,n} \frac{\lambda_i^L}{z_i(\tau)} \right)^{1-\eta} d\tau \right]^{\frac{1}{1-\eta}} = \left[\mathbb{E}\left[\max_{i=1,...,n} \left(\frac{Z_i}{\lambda_i^L} \right)^{\eta-1} \right] \right]^{\frac{1}{1-\eta}}$$

$$= \left[\left(\sum_{i=1}^{n} \left(\frac{x_i}{\lambda_i^L} \right)^{\frac{1}{\chi\xi}} \right)^{-\xi\chi(1-\eta)} \iota^{\eta-1} \Gamma(1+\chi(1-\eta)) \right]^{\frac{1}{1-\eta}} = \left(\sum_{i=1}^{n} \left(\frac{x_i}{\lambda_i^L} \right)^{\frac{1}{\chi\xi}} \right)^{-\chi\xi},$$

where $\Gamma(\cdot)$ is the Gamma function, evaluated at the argument $1+\chi(1-\eta)>0$. The third equality uses that λ_i^L and $z_i(\tau)$ are strictly positive and that $\eta > 1$, and the fourth equality follows from Lemma A.2, used to derive the distribution of the random variable $\max_{i=1,...,n} \left(Z_i/\lambda_i^L \right)^{\eta-1}$, followed by application of Lemma A.1.

For Part (ii), we start with the definition of task shares

$$\pi_i := \mathrm{P}\left[\lambda_i(\tau) = \min_{k=1,..,n} \lambda_k(\tau) \right] = \mathrm{P}\left[\left(Z_i/\lambda_i^L \right) = \max_{k=1,...,n} \left(Z_k/\lambda_l^L \right) \right]$$

To ease notation, let $\tilde{x}_i = \iota x_i/\lambda_i^L$ for any i and write the multivariate Fréchet distribution in terms of its tail dependence function $G(r_1, ..., r_n) = \left(\sum_{i=1}^n r_i^{1/\xi}\right)^\xi$, that is, as $\exp\left(-G\left(\left(\frac{z_1}{\iota x_1}\right)^{-1/\chi}, ..., \left(\frac{z_n}{\iota x_n}\right)^{-1/\chi}\right)\right)$. Since $G(\cdot)$ is homogeneous of degree (h.o.d.) one, the partial derivatives $G_i(\cdot) = \partial G(\cdot)/\partial r_i = r_i^{\frac{1}{\xi}-1}\left(\sum_{i=1}^n r_i^{1/\xi}\right)^{\xi-1}$ are h.o.d. 0.

Consider the following expression for any z,

$$
\begin{aligned}
&\mathrm{P}\left[\max_{k=1,...,n}\frac{Z_k}{\lambda_k^L}\le z \text{ and } \frac{Z_i}{\lambda_i^L}=\max_{k=1,...,n}\frac{Z_k}{\lambda_k}\right]=\mathrm{P}\left[\frac{Z_i}{\lambda_i^L}\le z \text{ and } \frac{Z_k}{\lambda_k}\le\frac{Z_i}{\lambda_i^L},\forall k=1,...,n\right]\\
&=\int_0^z\frac{\partial}{\partial z_i}\exp\left(-G\left(\left(\frac{z_1}{\tilde{x}_1}\right)^{-\frac{1}{\chi}},...,\left(\frac{z_n}{\tilde{x}_n}\right)^{-\frac{1}{\chi}}\right)\right)\Big|_{z_1=t,...,z_n=t}dt\\
&=\int_0^z\frac{1}{\chi}z_i^{-\frac{1}{\chi}-1}\tilde{x}_1^{\frac{1}{\chi}}G_i\left(\left(\frac{z_1}{\tilde{x}_1}\right)^{-\frac{1}{\chi}},...,\left(\frac{z_n}{\tilde{x}_n}\right)^{-\frac{1}{\chi}}\right)\times\exp\left(-G\left(\left(\frac{z_1}{\tilde{x}_1}\right)^{-\frac{1}{\chi}},...,\left(\frac{z_n}{\tilde{x}_n}\right)^{-\frac{1}{\chi}}\right)\right)\Big|_{z_1=t,...,z_n=t}dt\\
&=\int_0^z\frac{1}{\chi}t^{-\frac{1}{\chi}-1}\tilde{x}_i^{\frac{1}{\chi}}G_i\left(\left(\frac{t}{\tilde{x}_1}\right)^{-\frac{1}{\chi}},...,\left(\frac{t}{\tilde{x}_n}\right)^{-\frac{1}{\chi}}\right)\times\exp\left(-G\left(\left(\frac{t}{\tilde{x}_1}\right)^{-\frac{1}{\chi}},...,\left(\frac{t}{\tilde{x}_n}\right)^{-\frac{1}{\chi}}\right)\right)dt\\
&=\tilde{x}_i^{\frac{1}{\chi}}G_i(\tilde{x}_1^{\frac{1}{\chi}},...,\tilde{x}_n^{\frac{1}{\chi}})\int_0^z\frac{1}{\chi}t^{-\frac{1}{\chi}-1}\exp\left(-G\left(\tilde{x}_1^{\frac{1}{\chi}},...,\tilde{x}_n^{\frac{1}{\chi}}\right)t^{-\frac{1}{\chi}}\right)dt\\
&=\frac{\tilde{x}_i^{\frac{1}{\chi}}G_i(\tilde{x}_1^{\frac{1}{\chi}},...,\tilde{x}_n^{\frac{1}{\chi}})}{G\left(\tilde{x}_1^{\frac{1}{\chi}},...,\tilde{x}_n^{\frac{1}{\chi}}\right)}\int_0^z\frac{1}{\chi}t^{-\frac{1}{\chi}-1}G\left(\tilde{x}_1^{\frac{1}{\chi}},...,\tilde{x}_n^{\frac{1}{\chi}}\right)\exp\left(-G\left(\tilde{x}_1^{\frac{1}{\chi}},...,\tilde{x}_n^{\frac{1}{\chi}}\right)t^{-\frac{1}{\chi}}\right)dt\\
&=\frac{(x_i/\lambda_i^L)^{\frac{1}{\chi\xi}}}{\sum_{k=1}^n(x_k/\lambda_k^L)^{\frac{1}{\chi\xi}}}\times\left[\exp\left(-G\left(\tilde{x}_1^{\frac{1}{\chi}},...,\tilde{x}_n^{\frac{1}{\chi}}\right)t^{-\frac{1}{\chi}}\right)\right]_0^z\\
&=\frac{(x_i/\lambda_i^L)^{\frac{1}{\chi\xi}}}{\sum_{k=1}^n(x_k/\lambda_k^L)^{\frac{1}{\chi\xi}}}\times\exp\left(-G\left(\tilde{x}_1^{\frac{1}{\chi}},...,\tilde{x}_n^{\frac{1}{\chi}}\right)z^{-\frac{1}{\chi}}\right),
\end{aligned}
$$

where line five exploits $G(\cdot)$ being h.o.d. one and $G_i(\cdot)$ being h.o.d. zero; the last equality follows because $\exp(-\infty) = 0$. Now letting $z \to \infty$ yields the desired result.

Further, the max-stability property of the Fréchet distribution ensures that the conditional distribution of the maximum equals the unconditional distribution of the maximum. This ensures that $Q_i = \pi_i Q$, so that π_i also corresponds to fraction of the shadow value of all tasks used in final good production attributable to i. *Q.E.D.*

A.1.3 Proof of Proposition 1

Letting $\tilde{\chi} = \chi\xi$, part (i) of Lemma A.3 implies, given the normalization $\lambda = 1$, that

$$1 = \sum_{i=1}^{n} \left(\frac{x_i}{\lambda_i^L} \right)^{\frac{1}{\tilde{\chi}}}. \tag{A.6}$$

Substituting this into the expression for π_i in part (ii) of Lemma A.3 yields

$$\pi_i = (x_i/\lambda_i^L)^{\frac{1}{\tilde{\chi}}}, \tag{A.7}$$

which says that the share of tasks performed by i is increasing in her talent and decreasing in the shadow cost of her labor.

Next, to derive an expression for λ_i^L as a function of π_i and Y, the FOC for $l_i(\tau)$ given $y_i(\tau) > 0$ implies $l_i(\tau) = (\lambda_i(\tau)\lambda_i^L)y_i(\tau)$. By integrating over tasks and using the time constraint (4) we find that $\lambda_i^L = \int \lambda_i(\tau)y_i(\tau)d\tau$. Since $\lambda(\tau) = \lambda_i(\tau)$ if $y_i(\tau) > 0$, and given the definition of Q_i, it follows that $\lambda_i^L = Q_i$, which says that the shadow value of worker i's time is equal to the shadow value of all tasks produced by that worker. Combining this $Q_i = \pi_i Q$ (part (ii) of Lemma A.3), we find

$$\lambda_i^L = \pi_i Q. \tag{A.8}$$

Moreover, Q is related to Y as follows. Starting with equation (1), multiply both sides by $\lambda^{\frac{1}{\eta-1}}$, and bring this term inside the integral on the left-hand side. Substituting for $\lambda^{\frac{1}{\eta-1}}$ on that left-hand side using equation (A.1), rearranged as $\lambda^{\frac{1}{\eta-1}} = \left(\frac{Q(\tau)}{Y}\right)^{\frac{1}{\eta-1}} \frac{\lambda(\tau)}{\lambda}$, and simplifying algebra yields $\int_{\mathcal{T}} Q(\tau)d\tau = \lambda Y$.

As the left-hand side is the definition of Q and given $\lambda = 1$, it follows that $Q = Y$. Combining this with equation (A.8) we find that

$$\lambda_i^L = \pi_i Y. \tag{A.9}$$

Lastly, substitute for π_i in equation (A.9) using equation (A.7). Solving for λ_i^L, substituting into equation (A.6) and rearranging for Y yields equation (9). *Q.E.D.*

A.2 Matching block

A.2.1 Surplus recursions

This section derives recursions for the surplus values

$$S(x) = \Omega_1(x) - V_{f.0} - V_u(x), \tag{A.10}$$

and

$$S(x'|x,\xi) = \Omega_2(x,x',\xi) - \Omega_1(x) - V_u(x'). \tag{A.11}$$

Combining equation (A.10) with the expression for $\Omega(x)$ in equation (18) and the surplus sharing rules yields

$$\big(\rho+\delta\big)S(x) = f_1(x) - \rho\big(V_u(x)+V_{f.0}\big) + \lambda_{v.u}(1-\omega)\int \frac{d_u(\tilde{x}')}{u} S(\tilde{x}'|x,\tilde{\xi})^+ dH(\tilde{\xi})\tilde{x}'. \tag{A.12}$$

Similarly, for $S(x|x',\xi)$, and using the expression for $\Omega(x,x',\xi)$ in equation (17) gives

$$S(x|x',\xi)\big(\rho+2\delta\big) = f_2(x,x',\xi) - \rho\big(V_u(x)+V_u(x')+V_{f.0}\big) + \delta S(x) - (\rho+\delta)S(x'). \tag{A.13}$$

A.2.2 Population dynamics

For any type x, the measure of unemployment satisfies

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

The measure of exogenously separated workers of any type is equal to the measure of unemployed workers of that type finding new employment at either one-worker or two-worker firms.

For all x, the measure of one-worker matches follows

$$d_{m.1}(x)\bigg(\delta(x) + \lambda_{v.u}\int \frac{d_u(\tilde{x}')}{u}h(\tilde{x}'|x)d\tilde{x}'\bigg) = d_u(x)\lambda_u\frac{d_{f.0}}{v}h(x) + \delta(x)\int d_{m.2}(x,\tilde{x}')d\tilde{x}'. \tag{A.15}$$

Outflows from this state occur due to exogenous separation or because the one-worker firm meets and decides to hire a coworker of some type. Inflows occur when an unemployed worker of type x meets and gets hired by an unmatched firm or because a two-worker firm that has a type x as one of its employees loses the coworker.

Finally, for all (x, x'),

$$(\delta(x) + \delta(x'))d_{m.2}(x, x') = d_u(x)\lambda_u \frac{d_{m.1}(x')}{v} h(x|x') + d_u(x')\lambda_u \frac{d_{m.1}(x)}{v} h(x'|x). \tag{A.16}$$

The economic intuition parallels the aforementioned reasoning.

A.2.3 Stationary equilibrium

DEFINITION A.3: A stationary equilibrium consists of a production function $f(\cdot)$, a tuple of value functions, $\big(V_u(x), V_{f.0}, S(x), S(x|x', \xi)\big)$, together with a distribution of agents across states, $\big(d_{m.1}(x), d_{m.2}(x, x')\big)$, such that (i) $f(\cdot)$ is consistent with optimal task assignment (equation (6)), (ii) the value functions satisfy the HJB equations (15), (16), (A.12) and (A.13) given the distributions; and (iii) the stationary distributions satisfy the KFEs (A.15)-(A.16) given the policy functions implied by the value functions per equation (13).

A.2.4 Wage function

The value of employment for worker x given a coworker of type x' and match shock ξ is

$$\rho V_{e.2}(x|x', \xi) = w(x|x', \xi) - 2\delta\omega S(x|x') + \delta\omega S(x).$$

Combining with the surplus sharing rule (12) yields

$$w(x|x', \xi) = \rho V_u(x) + (\rho + 2\delta)\omega S(x|x', \xi) - \delta\omega S(x) \tag{A.17}$$

To simplify this expression further, substitute for $(\rho+2\delta)S(x|x', \xi)$ from equation (A.13)

$$\begin{aligned} w(x|x', \xi) &= \rho V_u(x) + \omega\left[f_2(x, x', \xi) - \rho\big(V_u(x) + V_u(x') + V_{f.0}\big) + \delta S(x) - (\rho + \delta)S(x')\right] - \delta\omega S(x), \\ &= \omega f_2(x, x', \xi) + (1 - \omega)\rho V_u(x) - \omega\rho\big(V_u(x') + V_{f.0}\big) - \omega(\rho + \delta)S(x'), \end{aligned}$$

and then substitute for $S(x')$ from equation (A.12) to obtain

$$w(x|x',\xi) = \omega\big(f_2(x,x',\xi) - f_1(x')\big) + (1-\omega)\rho V_u(x) - \omega(1-\omega)\lambda_{v.u} \int\int \frac{d_u(\tilde{x}'')}{u} S(\tilde{x}''|x',\tilde{\xi})^+ dH(\tilde{\xi}) d\tilde{x}''. \tag{A.18}$$

A.2.5 Coworker correlation in the model

Let $d_{m.2}(\hat{x},\hat{x}')$ denote the density of matches in terms of talent ranks.[A.1] Then the equilibrium joint density in teams is $\phi(\hat{x},\hat{x}') = \frac{d_{m.2}(\hat{x},\hat{x}')}{\int\int d_{m.2}(\hat{x},\hat{x}')d\hat{x}d\hat{x}'}$; the unconditional density in teams is $\phi(\hat{x}) = \int \phi(\hat{x},\hat{x}')d\hat{x}'$; and the distribution of coworker types conditional on type is $\phi(\hat{x}'|\hat{x}) = \frac{\phi(\hat{x},\hat{x}')}{\int \phi(\hat{x},\hat{x}')d\hat{x}'}$, the corresponding CDFs being $\hat{\Phi}(\hat{x})$ and $\hat{\Phi}(\hat{x}'|\hat{x})$. Then

$$\rho_{xx} = \frac{\displaystyle\int\int (\hat{x}-\bar{x})(\hat{x}'-\bar{x}) d\Phi(\hat{x}'|\hat{x}) d\hat{\Phi}(\hat{x})}{\displaystyle\int \big(\hat{x}-\bar{x}\big)^2 d\hat{\Phi}(\hat{x})}, \tag{A.19}$$

where $\bar{x} = \int \hat{x} d\hat{\Phi}(\hat{x})$ is the average worker type among those in teams.

A.3 Generalizing the two-worker model

The main text takes a team to comprise $n = 2$ workers. This assumption was driven by both analytical and computational constraints. First, it affords a closed-form expression for the production function $f(\cdot)$ when modeling correlated skills with a Gumbel copula. Second, in the context of the matching problem, it mitigates the problem of a combinatorially expanding state space that arises because under complementarities, hiring decisions depend on the *combination* of incumbent workers' types. This appendix outlines two approaches generalizing the production function to arbitrary $n \in \mathcal{Z}_{++}$.

Homogeneous ξ. The first, pragmatic approach assumes homogeneous bilateral distances, $\xi_{il} = \xi$, for any worker pair (i,l) among n workers; the case of workers' skill being

[A.1] I use the hat notation for simplicity, with the understanding that appropriate transformations due to the change of variables are applied. Moreover, the rank correlation aligns with the empirical approach of measuring discrete worker types, and avoids confounding shifts in assortativeness and in the marginal talent distribution.

independent ($\xi = 1$) is a special case. The production function then reads

$$Y = f(\mathbf{x}, \xi; \chi) = \left(\sum_{i=1}^{n} (l_i x_i)^{\frac{1}{\chi\xi+1}} \right)^{\chi\xi+1}, \tag{A.20}$$

where $\mathbf{x} = (x_1, ..., x_n)$. This implies an elasticity of complementarity that is symmetric and identical for all coworker pairs, and equal to $\gamma = \frac{\chi\xi}{1+\chi\xi}$. Note that equation (A.20) also incorporates heterogeneity in time endowments l_i to illustrate the isomorphism between talent and time endowment noted in Footnote 14. This first approach is very tractable and thus convenient, but the assumption of homogeneous horizontal distance is ad hoc.

Max-stable spatial model. A more rigorous approach involves explicitly constructing a max-stable spatial model for the skills process, generalizing the heuristic circular model in the main text. This approach exploits a spectral representation – using a Poisson point process to generate max-stable processes (e.g., De Haan, 1984). Dependence between extremes at different locations in the latent space – i.e., between workers – arises from shared spectral functions. Concretely, modeling skills in terms of a Brown-Resnick process, which uses Gaussian spectral functions, accommodates arbitrary n while preserving aggregation over the task space. While the bivariate margins of the Brown–Resnick process are of the Hüsler–Reiss type, which unlike the Gumbel copula does not yield a closed-form expression for $f(\cdot)$, in finite samples the Hüsler–Reiss copula is statistically indistinguishable from the Gumbel copula (Genest et al., 2011, Fig. 2). Thus, the latter is an analytically convenient and quantitatively innocent approximation. Appendix A.E.1 sketches more formally how such an approach could be developed.

A.4 Education and the rise in skill specificity

This section describes a theoretical result by which the sharp acceleration in secondary education since the 1980s[A.2] could generate an increase in skill specificity, i.e. the Fréchet inverse shape parameter χ. Intuitively, if more secondary education augments task-specific skills randomly, this trend translates into more dispersion in task-specific skills.[A.3]

[A.2] Barro and Lee (2013) show that the average years of secondary schooling in the German population aged 25+ increased from 1.32 years in 1985 to 6.9 years in 2010, having been relatively flat previously.

[A.3] I thank Ezra Oberfield for an insightful discussion motivating the reasoning developed here.

REMARK 5—Transformation of Fréchet shape parameter: Let Z be a Fréchet random variable (r.v.) with shape parameter $\theta > 0$ and scale parameter $x > 0$, and let $\{B_s\}_{s\geq 1}$ be a sequence of independent r.v.'s defined recursively as $B_s = \exp\left(-b_s/(\alpha\theta_{s-1})\right)$ where $\alpha \in (0,1)$, $\theta_0 = \theta$, $\theta_s = \theta_{s-1}\alpha = \theta\alpha^s$ for $ss \geq 1$, $\{b_s\}_{s\geq 1}$ are independent r.v.'s such that $\exp(b_s/\alpha)$ are i.i.d. positive α-stable r.v.'s. Assume Z and $\{B_s\}$ are independent. Define the r.v.'s $\{Z^{(s)}\}_{s\geq 1}$ recursively as $Z^{(0)} = Z, \quad Z^{(s)} = Z^{(s-1)} \times B_s, \quad s \geq 1$. Then for each $s \geq 1$, $Z^{(s)}$ is a Fréchet r.v. with scale x and shape $\theta_s = \theta\alpha^s$.

This result is a straightforward extension of Theorem 1 in Shanbhag and Sreehari (1977) and Boehm and Oberfield (2022, Footnote 11). Intuitively, if an individual starts a school year with a certain distribution of task-specific skills and leaves with a distribution where each task-specific skill is multiplied by an independent draw from a sufficiently fat-tailed distribution, with α controlling the dispersion in the multiplying draws, the more years of schooling they have, the more specific their skills will be.

APPENDIX B EMPIRICAL APPENDIX

B.1 Data sources and construction

Appendices O B.1.1 and B.1.2 describe the SIEED and the BIBB data sets. A description of how auxiliary labor market transition moments are constructed from the LIAB is relegated to Appendix A.F.2.

B.1.1 SIEED

This section provides details on the Sample of Integrated Employer-Employee Data (SIEED 7518) and how I process the data. Access is provided by the Research Data Center of the German Federal Employment Agency at the Institute for Employment Research (IAB). Compared to the more well-known LIAB, the SIEED dispenses with survey information, but comprises a larger sample and represents a broader period. This is especially advantageous in the analysis of time-series patterns in Section 4.2. A detailed description can be found in vom Berge et al. (2020).

The SIEED is based on administrative employer notifications and covers every worker at a random sample of ("panel") establishments as well as, crucially, the complete employment biographies of each of these workers, even when not employed at the establishments in the sample. For each worker, the data include information on demographics characteristics (e.g. age, gender, and education) and jobs (e.g. employer, wages, occupation, and industry). Throughout, I use the KldB-1988 2-digit classification for occupations and the WZ08 2-digit industry classification, relying on the harmonization over time provided by the IAB. To maximize sample coverage, I do not restrict myself to panel establishments, but instead require a minimum number of persons in every establishment-year cell (see below). The employment biographies come in spell format. I transform the dataset into an annual panel. Where a worker holds multiple jobs in a year, I define the job with the highest daily wage as the main episode. Nominal values are deflated using the Consumer Price Index (2015 = 100).

The main dataset is constructed based on the following selection criteria and steps. (1) I retain observations for individuals between 20-60 with a workplace in former West Germany (excl. Berlin), employed full-time and subject to social security. Among other things, spells associated with marginal part-time employment (recorded only from 1999 onwards) are excluded. (2) I drop observations in industries classified as agriculture and mining, utilities, finance and insurance, households as employers, as well as semi- or fully public industries such as social security and education (codes 31-9, 35-39, 64-68, 84-85, 87-88, 91, 94, 96-99). (3) I require non-missing observations for identifiers, wages, industry, occupation, age, tenure, gender and education. (4) I drop singleton person observations and focus on the largest connected set for each of 5 sample periods (1985-1992, 1993-1997, 1998-2002, 2003-2009, 2010-2017). (5) Using the resulting sample, I estimate residual wages and types, as described in detail below, and retain only observations for which these variables are not missing. (6) The final analysis sample is based on the additional restrictions that the establishment-year cell contains no fewer than ten observations; and that the industry cell contains at least 500 person observations and at least 25 unique employers in any of the years. These restrictions are important for analyses with employer and industry fixed effects as well as the industry-level analysis in Section 3.3.

Turning to wages, the earnings variable is top-coded at the "contribution assessment limit" ("Beitragsbemessungsgrenze") of the social security system. To impute right-censored wages, which is done after step (3), I follow standard practices, notably Card et al. (2013) and Dauth and Eppelsheimer (2020), by fitting a series of Tobit models to log daily wages, then imputing an uncensored value for each censored observation using the estimated parameters of these models and a random draw from the associated (censored) distribution. I fit 16 Tobit models (4 age groups, 4 education groups), after having restricted the sample per the above. I follow Card et al. (2013) 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, imputed wages are limited to 10 times the 99th percentile.

The empirical analyses use residualized wages as an input. To construct those, I proceed similarly to Card et al. (2013) and Hagedorn et al. (2017). Specifically, in the pooled sample I regress the (raw) log real daily wage of worker i in year t, $\tilde{w}_{it}$, on a person fixed effect and a time-varying characteristics index, X_{it} that comprises an unrestricted set of year dummies, quadratic and cubic terms in age interacted with educational attainment and a quadratic in job tenure.[B.1] The choice of covariates is informed by the observation that the theoretical model does not incorporate life-cycle factors or on-the-job learning nor aggregate productivity growth. Education itself is not included as a covariate, as it arguably represents time-invariant component of worker skill. As an input into the analysis, I then use $w_{it} = \exp\left(\tilde{w}_{it} - X_{it}'\hat{\beta}\right)$.

The final sample (1985-2017) includes 15,359,711 person-year observations for 2,018,314 unique persons, whose average age is 38 years. The mean unweighted establishment size is 29, the median being 14. The log average real (raw) daily wage in 2010 is 4.66.

B.1.2 Task complexity measure

This section summarizes how the task-complexity proxy is constructed. Appendix A.F.4 provides more details.

I draw on 5 waves of the Employment Surveys carried out by the German Federal Institute for Vocational Training (Bundesinstitut fuer Berufsbildung, BIBB), for the years 1985/86, 1991/92, 2006, 2012, and 2018. The survey interviews individual workers and features a recurring question about the tasks performed at work, e.g. in 2012: “Please think of your occupational activity as [...]. I will now give you a number of specific job tasks. Please tell me how often these job tasks occur in your work, whether they occur often,

[B.1] As the regression includes person and year fixed effects, I exclude the linear age term in light of the age-year-cohort identification problem. As in Card et al. (2013), age is normalized around 40.

sometimes or never." I follow the BIBB's guidance in standardizing the sample basis and focus on tasks that were repeatedly queried over time.

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, marketing, public relations
Other	repairing, buying, accommodating, caring, cleaning, protecting, measuring, operating, manufacturing, storing, writing, calculating	

TABLE B.1

CLASSIFICATION OF TASKS IN THE BIBB EMPLOYMENT SURVEYS

Notes. This table summarizes the classification of tasks into two groups: "Complex" and "Other."

Table B.1 summarizes how I classify tasks, which is guided by Spitz-Oener (2006) and Rohrbach-Schmidt and Tiemann (2013). Next, we can define an index capturing the importance of complex tasks for worker i in period p,

$$T_{ip}^{\text{complex}} = \frac{\text{\# of activities performed by } i \text{ in task category “complex” in sample year } p}{\text{total \# of activities performed by } i \text{ in sample period } p}$$

For example, if worker i performs five distinct activities and two of those belong to the category of complex tasks, then $T_{ip}^{\text{complex}} = 0.4$.

To merge the task-complexity measure into the SIEED I use occupational averages, computed using the German Classification of Occupations 1988 (KldB88). To link the KldB88 with waves using the KldB92, I rely on a crosswalk, which is of high quality as the two classifications are very similar.

The analysis uncovers several key findings, summarized next and decribed in further detail in Appendix A.F.4.

(i) The aggregate usage share of complex tasks in workers' activities has monotonically risen since 1985/86; the 1990s saw a particularly sharp increase.

(ii) This trend is prevalent across different levels of education. It is not driven by occupational employment effects either, instead the majority of the increase occurs within-occupation.

(iii) In the cross-section, the task portfolio of more educated individuals tends to be disproportionately skewed toward complex tasks compared to less educated individuals. The ranking of different occupations is intuitive and likewise reveals large variation in task shares.

(iv) Cross-sectional differences are robust to using measures of time spent on different tasks.

Altogether, these results provide reassurance that the time trend reported in Figure 3a is robust and not merely the result of composition effects, and that there is substantial cross-sectional variation.

B.2 Robustness checks: estimating the interaction coefficient β_c

This appendix reports a battery of robustness checks for the auxiliary regression estimating the interaction coefficient β_c that disciplines skill specificity χ in the indirect inference step. In parallel, it documents trends in coworker sorting using alternative worker type measures and sub-samples.

B.2.1 Baseline period: different controls

Table B.2 reports the estimates of β_c for the sample period 2010-2017 according to a variety of different specifications. It shows, firstly, that it makes little difference to $\hat{\beta}_c$ what combination of employer/industry/occupation FEs is included. Secondly, when workers are ranked within occupation, instead, $\hat{\beta}_c$ is only slightly lower, at 0.0059 (column (5)).

In addition, when the exercise of ranking workers and coworkers and estimating equation (20) is implemented separately for each 2-digit industry, the average value of $\hat{\beta}_c$ is similar, too, at 0.0061.

	(1)	(2)	(3)	(4)	(5)
Interaction coefficient ($\hat{\beta}_c$)	0.0068***	0.0064***	0.0063***	0.0061***	0.0059***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0008)
Employer FEs	No	No	Yes	Yes	Yes
Industry-year FEs	No	Yes	No	Yes	Yes
Occupation-year FEs	No	No	Yes	Yes	Yes
Type ranking	Economy	Economy	Economy	Economy	Occupation
Obs. (1000s)	3,912	3,912	3,912	3,912	3,606
Adj. R^2	0.787	0.800	0.809	0.816	0.769

TABLE B.2

AUXILIARY REGRESSION RESULTS (2010-2017 PERIOD)

Notes. Regression results based on equation (20), reporting the estimates of the coefficient on the interaction term, $\hat{\beta}_c$. The dependent variable is the (residualized) wage, in levels and divided by the year-specific average wage. Employer-clustered standard errors are given in parentheses. Observations are weighted by the inverse employment share of the respective type and (rounded) coworker type cell. Observation count rounded to 1000s. * p<0.1; ** p<0.05; *** p<0.01.

B.2.2 Years of schooling as a worker type measure

One concern with the auxiliary regression is that the dependent variable is that the period-t wage and the independent variables of interest, own and coworker type, are likewise a function of wages (i.e., a function of wages across all years comprised in the respective sample period). While this approach is model-consistent – with identification coming from variation in wages over time, while the regressors are invariant across years within sample periods – we may still worry about confounding effects.

The finding of positive and rising β_c is robust to using years of schooling as a non-wage proxy for worker talent types instead, similar to Nix (2020). Table B.3 shows the coworker correlation coefficient, $\rho_{xx'}$, and the interaction coefficient, β_c, under this alternative worker type measure. While the magnitudes are not directly comparable to those obtained under the

Period	Coworker sorting	Interaction coefficient ($\hat{\beta}_c$)
1985-1992	0.45	0.007***
1993-1997	0.48	0.005***
1998-2003	0.52	0.008***
2004-2009	0.55	0.009***
2010-2017	0.56	0.010***

TABLE B.3

SORTING AND INTERACTION COEFFICIENT: YEARS OF SCHOOLING

Notes. The entry in the first main column shows the correlation between own and average coworker years of schooling. The entry in the second main column is the point estimate for the interaction coefficient β_c in a version of regression (20) adapted to the use of years-of-schooling as a type measure. Years of schooling are imputed from completed education following Card et al. (2013). * p<0.1; ** p<0.05; *** p<0.01.

FE-based talent type measure, the time trends for both moments align with those reported in the main text.

In unreported results I also considered the alternative of estimating worker FEs and constructing worker types $\hat{x}_i$ that are invariant across sample periods. The implied time trends in $\rho_{xx'}$ and β_c are very similar to the baseline.

B.2.3 Log wage regression

In the auxiliary regression specification, the dependent variable is the (normalized) wage *level*, consistent with the theoretical model. A more conventional regression specification uses the log wage on the left-hand side and omits the interaction term,

$$\ln w_{it} = \beta_0 + \beta_1 \hat{x}_i + \beta_2 \hat{x}_{-it} + \psi_{j(i,t)} + \nu_{0(i,t)t} + \xi_{S(i,t)t} + \epsilon_{it}, \tag{B.1}$$

where types are treated as continuous variables.

Table B.4 reports the coefficient estimates for each of the 5 sample periods. The coefficient on the coworker type, β_2, has increased over time, suggesting coworker spillovers have become more important over time.

B.2.4 Moments by establishment size groups

A different concern regarding inference on β_c is that the variation in coworker quality exploited for identification may not be exogenous with respect to the error term ϵ_{it}. In

	1985-1992	1993-1997	1998-2003	2004-2009	2010-2017
Type	0.0930***	0.0940***	0.1045***	0.1124***	0.1138***
	(0.0003)	(0.0002)	(0.0004)	(0.0003)	(0.0004)
Coworker Type	0.0010	0.0034***	0.0051***	0.0081***	0.0073***
	(0.0006)	(0.0007)	(0.0008)	(0.0008)	(0.0007)
Observations (1000s)	3,366	2,274	2,410	3,398	3,912
Adj. R^2	0.777	0.805	0.805	0.838	0.817

TABLE B.4

AUXILIARY REGRESSION RESULTS BY SAMPLE PERIOD - LOG SPECIFICATION

Notes. All regressions include industry-year, occupation-year, and establishment FEs. Employer-clustered standard errors are given in parentheses. Observations are weighted by the inverse of the empirical match density for worker-type and coworker-type combinations, and rounded to the nearest thousand. Significance levels: * p<0.1; ** p<0.05; *** p<0.01.

the structural model, random variation naturally occurs due to search frictions. It could be argued, though, that this idiosyncratic variation averages out when constructing a representative coworker type for larger firms (cf. Hoxby, 2000).

To evaluate these concerns, I compute the coworker correlation measure of sorting, $\rho_{xx'}$, and estimate the interaction coefficient, β_c, separately for 4 different establishment size groups and, to study time trends, for 5 different sample periods. I perform this exercise both for the full sample and for the subsample of panel establishments for which the SIEED contains information on the entire workforce. Figure B.1 visualizes the results.

The results from this exercise are reassuring. Across establishment size groups, there is no evident sign of a bias, the magnitude of the estimates for β_c and sorting being similar across size bins. It can also be seen that trends are similar when using data for panel establishments only.

B.2.5 Non-linear coworker aggregation

In the main analysis, the representative coworker type for each person-year observation is constructed as the unweighted arithmetic mean of all coworkers' types, i.e., $\hat{x}{-it} = \left(\frac{1}{|S{-it}|}\sum_{k\in S_{-it}}\hat{x}_k\right)$. However, this approach ignores the non-linearity in aggregation *across coworkers* implied by the reduced-form production function (9): the power mean assigns disproportionate weight to low-type coworkers, particularly when χ and thus

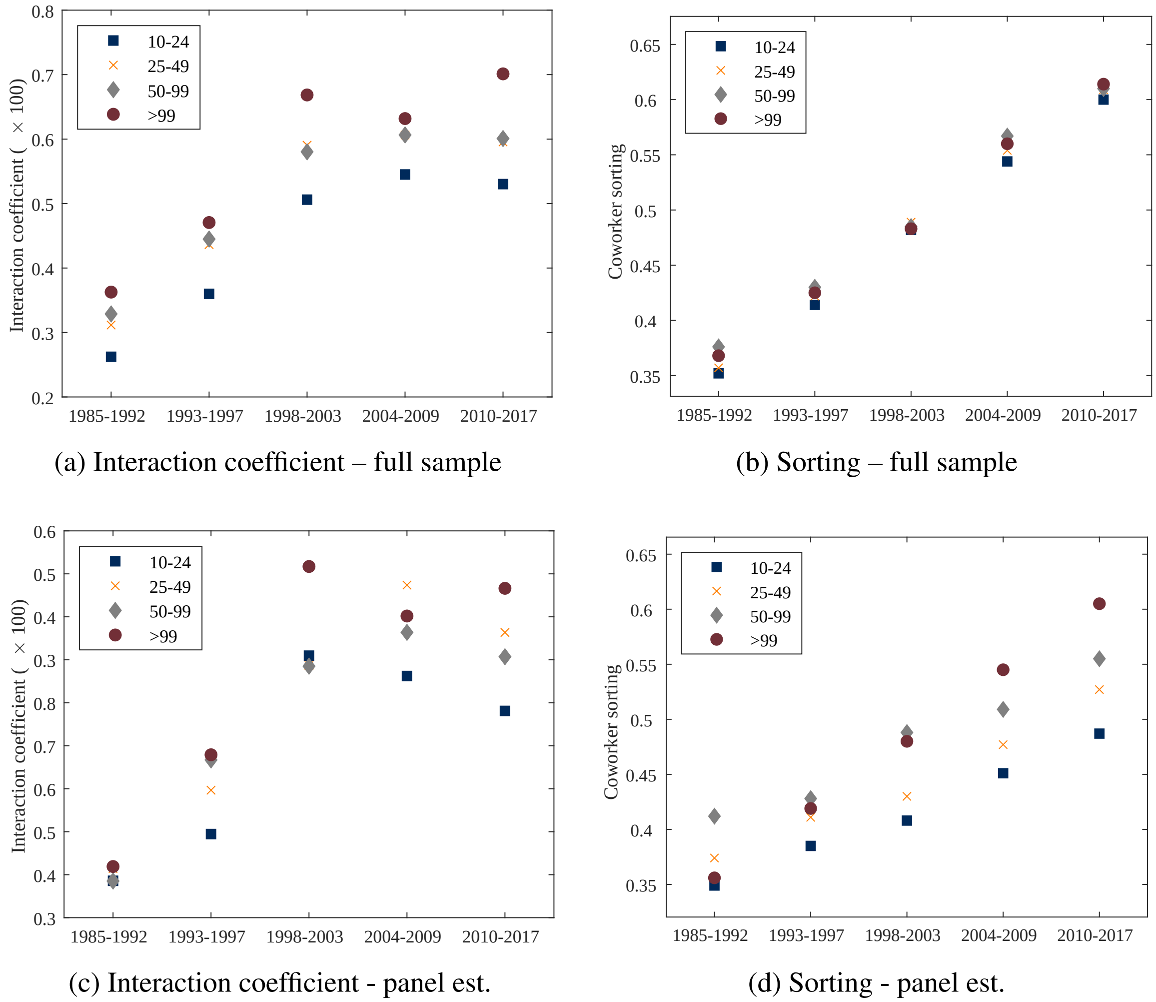


FIGURE B.1.—Auxiliary regression estimates and sorting by establishment size group

Notes. This figure indicates empirical estimates of β_c (panels (a) and (c)) and the coworker correlation (panels (b) and (d)) and when estimated separately for 5 sample periods and several establishment size groups. The underlying worker types are estimated from a pooled sample across all establishment size groups, as in the remainder of the paper. In panels (a) and (b), the estimation sample comprises all establishments, whereas panels (c) and (d) are based on the subsample of 'panel establishments'.

talent complementarities are large. Implementing the correct aggregation is practically infeasible, as it requires knowledge of χ, whose identification in turn relies on measures of coworker types and calibration of the full structural model.

Reassuringly, the bias from ignoring this non-linearity is small for the following reason. Let the correctly aggregated average coworker be $\tilde{x}{-}it = \left(\frac{1}{|S{-}it|}(\hat{x}k)^{1-\gamma}\right)^{\frac{1}{1-\gamma}}$, where γ

is the reduced-form elasticity of complementarity defined in Corollary 2. A second-order Taylor approximation around $\hat{x}{-}it$ shows that $\tilde{x}{-}it - \hat{x}{-}it \approx -\frac{1}{2}\gamma\frac{\sigma_x^2}{\hat{x}-it}$, where $\hat{\sigma}x'^2$ is the variance of coworker types. This shows that the unweighted average is upward biased in proportion to the product of γ and within-match dispersion of coworker talent. The bias remains small because precisely when χ is high, positive assortative matching ensures low dispersion in coworker talent.

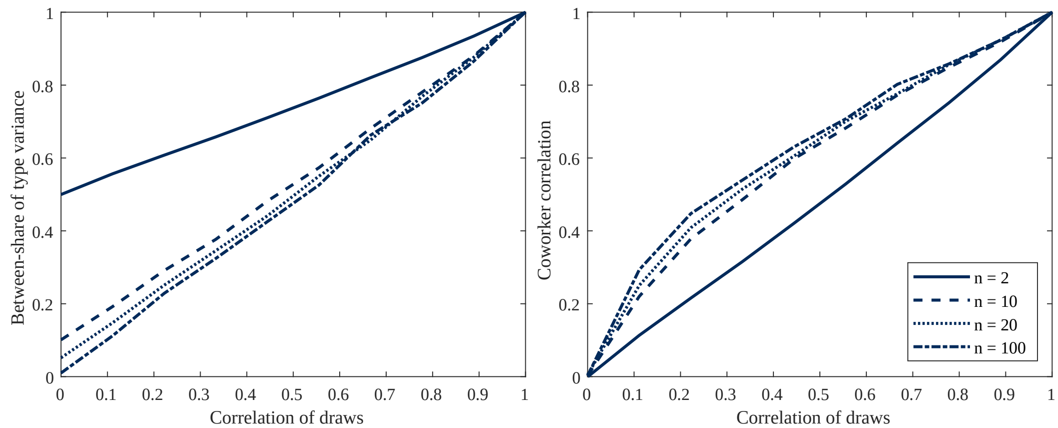
APPENDIX C QUANTITATIVE APPENDIX

C.1 Between-share adjustment procedure

This section describes an adjustment method used to correct the model-implied between-within firm variance decomposition for a statistical bias. Since the structural model features teams of size two, the level of the between-share of the log wage variance is upward biased when coworker sorting is less than perfect ($\rho_{xx'} < 1$) due to a mechanical statistical effect: the law of large numbers does not apply within production units. Even with random matching ($\rho_{xx'} = 0$), chance alone creates between-firm average wage differences that would vanish with large firms. This bias diminishes as sorting increases, disappearing entirely when $\rho_{xx'} = 1$, in which case all dispersion is across and none within units, regardless of team size. Figure C.1a illustrates these ideas graphically; its construction is described below.

The adjustment method I propose is based on a *statistical* model that can flexibly accommodate different degrees of coworker sorting as well as team sizes. Consider a random vector $X = (X_1, X_2, \ldots, X_n)'$ whose distribution is described by a Gaussian copula over the unit hypercube $[0, 1]^n$, with an $n \times n$ correlation matrix $\Sigma(\rho^c)$, which contains ones on the diagonal, while the off-diagonal elements are all equal to a parameter ρ_c. Formally, the Gaussian copula with parameter matrix $\Sigma(\rho_c)$ is $C_\Sigma^{\text{Gauss}}(x) = \Phi_R(\Phi^{-1}(x_1), \ldots, \Phi^{-1}(x_n))$, where Φ^{-1} is the inverse cdf of a standard normal and Φ_R is the joint cdf of a multivariate normal distribution with mean vector zero and covariance matrix equal to $\Sigma(\rho_c)$. In our context, n may be interpreted as the average team size. Each vector of observations drawn from the distributions of X, $x_j = (x_{1j}, x_{2j}, \ldots, x_{nj})'$, describes the types of workers in that team, indexed by j.

In this setup, we can derive an analytical formula for the population between-team share of the variance of types as a function of n and ρ_c. Since the marginals of the



(a) Between-share of type var.

(b) Coworker sorting

FIGURE C.1.—Illustration of the between-share adjustment method

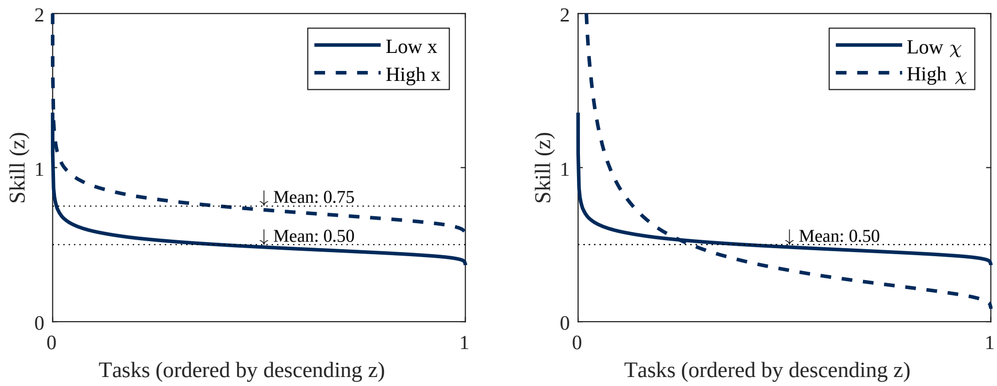
Notes. This figure illustrates how the coworker correlation coefficient (left panel) and the between-firm share of the variance of types (right panel) vary with the correlation parameter in the Gaussian copula. The different lines represent different "team sizes," that is, varying lengths n of the vector X. Workers are binned into ten deciles. The results are based on one million draws.

Gaussian copula are continuous uniforms over $[0,1]$, the variance of the union of all draws is $\frac{1}{12}$. The team mean $\bar{x}j = \frac{1}{n}\sum i = 1^n x_{ij}$ has variance $\frac{1}{n^2}\left(\frac{n}{12} + n(n-1)(\frac{\rho_c}{12})\right)$ given the specified correlation structure. Taking the ratio yields the between share: $\sigma^2_{x,\text{between-share}}(\rho_c, m) = \frac{1}{m}\left(1 + (m-1)\rho_c\right)$. With empirical average size $\hat{n}$, the correction factor is: correction-factor $= \frac{1}{2}\left(1+\rho_c\right) - \frac{1}{\hat{n}}\left(1 + (\hat{n}-1)\rho_c\right)$. I set $\hat{n} = 20$ and $\rho_c = 0.38$, the empirical coworker correlation in the first sample period.[C.1]

Two potential concerns merit discussion. First, ρ_c differs from the coworker correlation $\rho_{xx'}$. To compare them, consider M samples from X with leave-out-mean $\bar{x}_{-i,j} = \frac{1}{n-1}\sum_{k\neq i} x_{k,j(i)}$ and coworker correlation $\rho_{xx'} = \text{corr}(x_i, \bar{x}{-}i)$. Figure C.1b confirms that $\rho_{xx'}$ and ρ_c track closely, though $\rho_{xx'} > \rho_c$ for larger n and intermediate ρ_c. Second, the adjustment concerns variance of types rather than wages. Simulations using the structural wage function confirm that adjustment factors for types and wages are very similar.

APPENDIX D ADDITIONAL FIGURES AND TABLES

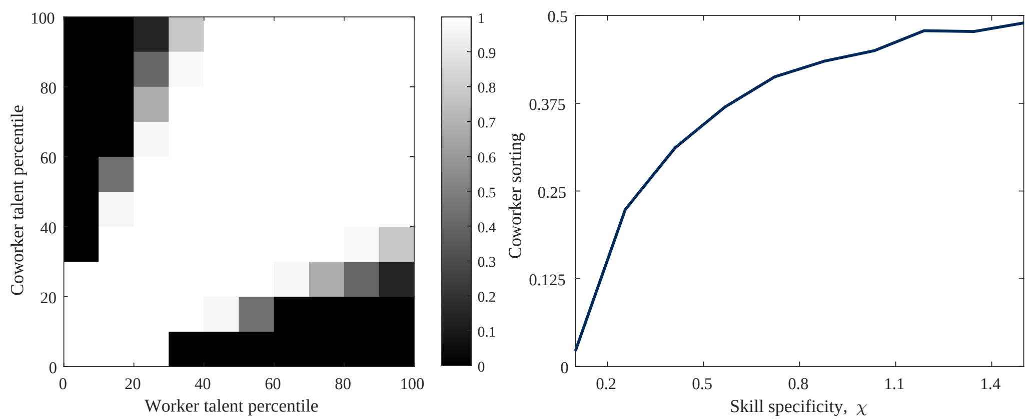
[C.1] Basing the correction on the earlier sample implies a bigger downward adjustment, which avoids over-stating the degree of between-firm inequality that the model can endogenously generate. Note also that the exact value of $\hat{n}$ does not matter much, as the magnitude of the bias diminishes rapidly as $\hat{n}$ grows.



(a) x_i: talent (b) χ: skill specificity

FIGURE D.1.—Parametrization of a worker's task specific skills

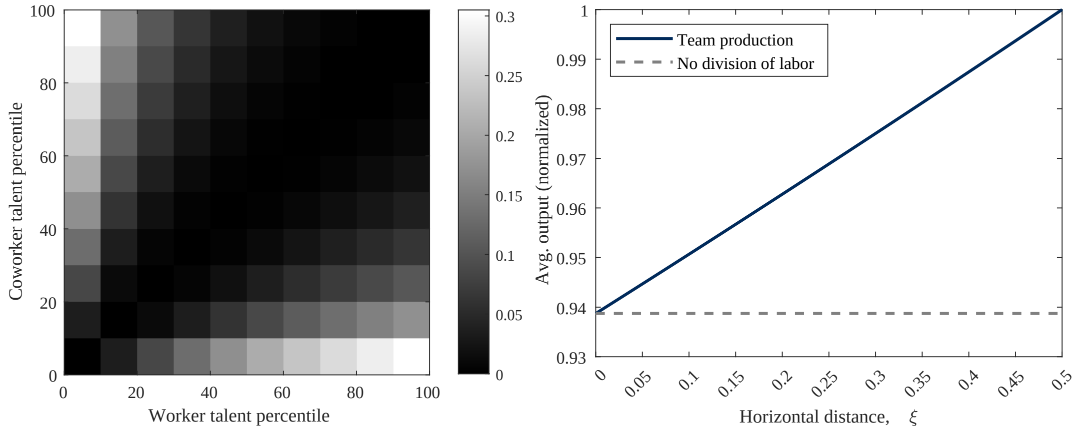
Notes. This figure illustrates the properties of the Fréchet distribution. Panels a(a) and (b) plot realizations of $z_i(\tau)$ against tasks, ordering the otherwise unordered tasks by descending z. For this illustration, $\eta = 2$, so that $\iota = (\Gamma(1-\chi))^{-1}$, and hence $\mathbb{E}[Z_i] = x_i$ for $\chi < 1$.



(a) Matching probabilities (b) Skill specificity and sorting

FIGURE D.2.—Equilibrium matching and comparative statics: skill specificity fosters sorting

Notes. The left panel shows the conditional matching probabilities; parameter values are taken from Section 3.4. The right panel shows how coworker sorting varies with χ.



(a) Talent complementarity (b) Horizontal distance

FIGURE D.3.—Properties of the production function

Notes. In the left panel, each cell indicates for every pair (x,x'), the value of $\frac{f(x,x|\xi=1)+f(x',x'|\xi=1)-2f(x,x',\xi=1)}{f(x,x|\xi=1)+f(x',x'|\xi=1)}$. The right panel shows $\int\int \frac{f(x,x',\xi)}{f(x,x',1)}dxdx'$ as a function of ξ.

	Coworker sorting	
Period	Within-economy type ranking	Within-occupation type ranking
1985-1992	0.38	0.32
1993-1997	0.44	0.37
1998-2003	0.51	0.42
2004-2009	0.58	0.45
2010-2017	0.64	0.48

TABLE D.1

COWORKER SORTING OVER TIME

Notes. This table indicates the correlation coefficient between a worker's estimated type and that of their average coworker, computed separately for five sample periods.

Parameter	Description	’85-’92	’93-’97	’98-’03	’04-’09	’10-’17
δ_0	Sep. rate, constant	0.014	0.016	0.017	0.019	0.015
δ_1	Sep. rate, scale	-0.981	-1.009	-1.014	-0.941	-0.841
χ	Skill specificity	0.280	0.463	1.044	1.091	2.466
a_0	Production, constant	0.445	0.438	0.355	0.301	0.267
a_1	Production, scale	1.229	1.243	1.389	1.554	1.562
$\bar{b}$	Unemp. flow utility, scale	0.703	0.700	0.699	0.638	0.652
λ_u	Meeting rate	0.161	0.118	0.127	0.154	0.228

TABLE D.2

ESTIMATED PARAMETERS – BASELINE

Notes. This table summarizes the parameters estimated under the baseline specification for each of the 5 sample periods.

Parameter	Description	’85-’92	’93-’97	’98-’03	’04-’09	’10-’17
δ_0	Sep. rate, constant	0.014	0.015	0.016	0.017	0.013
δ_1	Sep. rate, scale	-0.946	-0.956	-0.936	-0.834	-0.682
χ	Skill specificity	0.205	0.381	0.386	0.613	1.079
a_0	Production, constant	0.449	0.445	0.381	0.312	0.266
a_1	Production, scale	1.233	1.245	1.407	1.577	1.600
$\bar{b}$	Unemp. flow utility, scale	0.707	0.699	0.697	0.638	0.644
λ_u	Meeting rate	0.156	0.115	0.114	0.145	0.225

TABLE D.3

ESTIMATED PARAMETERS – WITHIN-OCCUPATION RANKING OF WORKER TYPES

Notes. This table summarizes the parameters estimated for each of the 5 sample periods when workers are ranked within two-digit occupations.

Parameter	Description	'85-'92	'93-'97	'98-'03	'04-'09	'10-'17
δ_0	Sep. rate, constant	0.014	0.016	0.017	0.019	0.015
δ_1	Sep. rate, scale	-0.981	-1.009	-1.014	-0.941	-0.841
χ	Skill specificity	0.181	0.269	0.479	0.520	1.175
a_0	Production, constant	0.463	0.474	0.423	0.410	0.358
a_1	Production, scale	1.195	1.168	1.278	1.355	1.400
$\bar{b}$	Unemp. flow utility, scale	0.708	0.703	0.700	0.639	0.645
λ_u	Meeting rate	0.156	0.112	0.116	0.140	0.218

TABLE D.4

ESTIMATED PARAMETERS – WITHIN-INDUSTRY ANALYSIS

Notes. This table summarizes the parameters estimated for each of the 5 sample periods when the targeted variance of wages and β_c are constructed as the averages across two-digit industries.

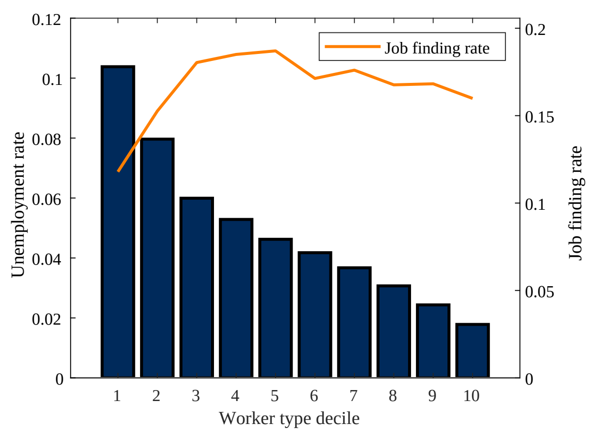


FIGURE D.4.—Job finding and unemployment rates by worker talent type

Notes. This figure shows unemployment and job finding rates by talent type in the estimated model.

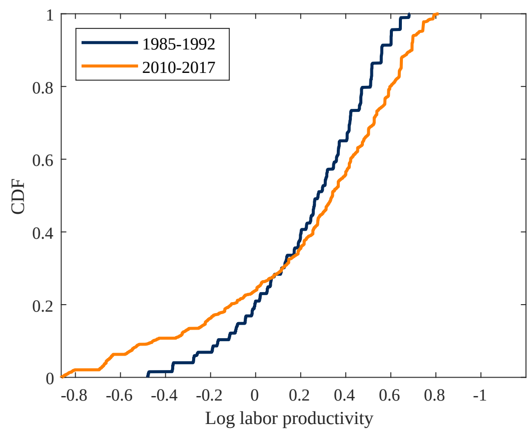


FIGURE D.5.—Model predicts increased firm-level productivity dispersion

Notes. This figure plots the cumulative distribution of labor productivity implied by the model, comparing the calibrated versions for 2010-2017 and 1985-1992.

Co-editor [Name Surname; will be inserted later] handled this manuscript.