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

Lukas B. Freund§
University of Cambridge

December 2022

Abstract

Modern production frequently involves teamwork among employees specialized in different tasks. I develop a model of teams in which firms assign tasks to workers who are heterogeneous in their overall quality and whose efficiency varies across different tasks. In addition to productivity gains, the division of labor generates coworker complementarities: the marginal productivity of one employee's quality is increasing in other team members' quality. This interdependence is stronger when variation in worker-task specific efficiencies is high. In frictional labor markets, coworker complementarities carry macroeconomic implications for both productivity and inequality. Coworker quality mismatch lowers team productivity, leading employers to search for workers of similar quality. In equilibrium, firms with "superstar teams" pull away in terms of productivity and pay. I validate the model's key mechanisms using administrative micro data. Paralleling a shift in the nature of tasks, a theory-informed measure of coworker complementarities has doubled since 1990. A structural estimation exercise suggests that this rise explains between one quarter and one half of the increase in the between-firm share of wage inequality in Germany (1990-2010).

Keywords: skill complementarity, firms, inequality, matching, teams, technology

*I am immensely grateful to Vasco Carvalho for his guidance, support and feedback. I thank Pontus Rendahl and Wouter Den Haan for frequent conversations. This research was partly conducted while visiting Princeton University, and I thank Ezra Oberfield and Gianluca Violante for their hospitality and advice. For helpful discussions and comments I furthermore thank Florin Bilbiie, Giancarlo Corsetti, Maarten De Ridder, Matthias Doepke, Mayara Felix, Luis Garicano, John Grigsby, Joe Hazell, Tomer Ifergane, Bill Janeway, Hanbaek Lee, Ben Moll, Adrian Ochs, Jane Olmstead-Rumsey, Tommaso Porzio, Ricardo Reis, Richard Rogerson, Ludwig Straub, Tom Winberry and Christian Wolf. I am indebted to Chiara Criscuolo and Peter Gal for facilitating the analysis of Portuguese microdata as well as fruitful conversations. I gratefully acknowledge financial support from Gates Cambridge Trust (BMGF OPP1144), the Keynes Fund (JHVB) and the Procter Fund.
This study uses the Linked Employer Employee Data longitudinal model 1975-2019 (LIAB LM7519) from the IAB. Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and subsequently remote data access under project number fdz2188.
First version: 8 December 2022.

§*Email*: lukas.beat.freund@gmail.com. *Web*: www.lukasfreund.com.

1 Introduction

A classical literature highlights the division of labor as an important source of economic prosperity. Consistent with this description, modern production processes frequently involve teamwork between employees who are specialized in different job tasks. Firms coordinate this division of labor and thus mobilize the vast knowledge required to efficiently produce goods and services. Indeed, the nature of work has shifted away from routine jobs toward tasks that are more complex and which tend to have greater training requirements (Figure 1a). The importance of team production appears to have grown concomitantly.[1] Yet, macroeconomic models of growth or inequality commonly make limited room for heterogeneity of workers and jobs within firms, or for their interdependence.[2]

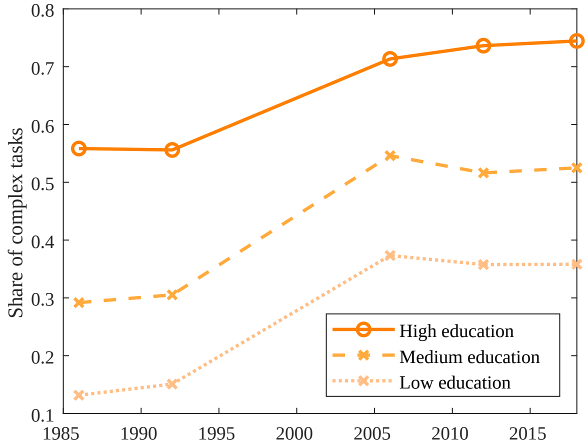
In this paper, I develop a parsimonious theory of production with a team of heterogeneous workers, and argue that it reveals a distinct feature of the division of labor that carries macroeconomic significance. That feature is coworker complementarity in production: the marginal productivity of one employee's quality is increasing in other team members' quality. The model shows that such complementarities naturally arise from the division of labor, and that their strength varies with production requirements. Specifically, when tasks are more complex, requiring workers to specialize, then coworker complementarities are stronger. Intuitively, if workers' abilities vary across tasks, then total output can be substantially dragged down by one worker performing poorly, as other team members may not know how to efficiently execute the tasks that worker is responsible for. As a result, the division of labor is not only associated with familiar productivity gains. The resulting complementarities also imply that quality mismatch across team members lowers productivity. This has consequences for whom employers are willing to hire in the first place.

Through the lens of this model, the transformation of work happening at the micro level has macroeconomic implications for productivity and the evolution of wage inequality. A shift towards more complex tasks leads to stronger coworker complementarities, worsening productivity losses from coworker mismatch. Firms respond by hiring workers of (more) similar quality in equilibrium, even if this involves costly search. As a consequence, gaps between firms in terms of workforce quality, productivity and pay widen. A rich empirical literature documents that this "firming up of inequality" (Song *et al.*, 2019) is indeed observed across for many rich economies.[3] In particular, the rise in wage inequality over the past few decades is primarily a between-firm phenomenon (Figure

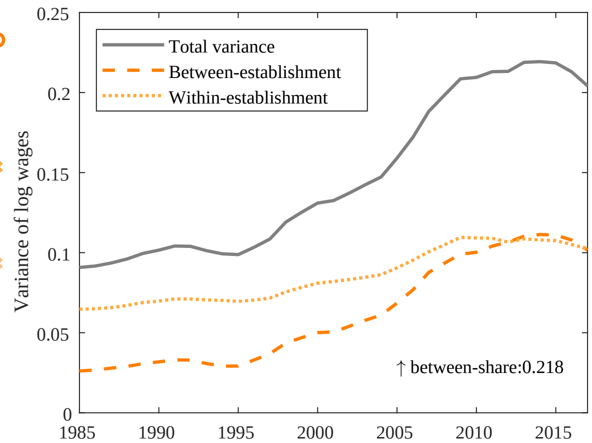
[1] See, for instance, Bresnahan *et al.* (2002); Jones (2009); Bloom and Reenen (2011); Lin (2011); Deming (2017).

[2] An important exception is the literature on knowledge-based hierarchies, following Garicano (2000). Garicano and Rossi-Hansberg (2015) survey how this framework has generated insights into a wide variety of economic phenomena, including the distribution of income.

[3] Key references include Card *et al.* (2013); Song *et al.* (2019); Criscuolo *et al.* (2021). Appendix A.4 recaps the rising share of wage inequality accounted for by between-firm differences across countries; performs a more detailed wage variance decompositions using micro data from Germany and Portugal; and studies labor market sorting patterns In parallel, firm productivity has become increasingly dispersed since the 1980s, even within narrowly defined industries and particularly in skill-intensive sectors (Andrews *et al.*, 2019; Decker *et al.*, 2020; De Ridder, 2021).



(a) Task complexity



(b) Wage inequality

Figure 1: Micro and macro trends in Germany

Notes. Panel 1a depicts the average share of complex, or non-routine abstract, tasks in individual workers' set of activities, for four points in time and distinguishing between three education groups. It is constructed from repeated, cross-sectional waves of the German BIBB Employment Survey. See Appendix D for details. In Panel 1b, the solid, dashed and dotted lines depict, respectively, the total variance of log (residual) wages and the between-establishment and within-establishment components. I report the centered three-year moving average. Appendix A.4 contains details.

1b); an important factor being that labor markets are increasingly characterized by coworker sorting, so that the most productive workers collaborate in some firms, segregated from less productive employees who cluster at other employers. The paper thus contributes a theoretical framework that offers a unified and parsimonious interpretation of empirically observed, structural shifts within and between firms.

The paper proceeds in three main steps. In step one I build a model of team production and embed it into a general equilibrium, frictional labor market environment. A sequence of theoretical results – concerning aggregation, characterization, and measurement – sharply summarizes how the task content of production maps onto complementarities; what the implications are at the macro level; and how to bring the model to the data. In step two I validate key model mechanisms and estimate the evolution of coworker complementarities, using matched employer-employee micro data. I find that in Germany, coworker complementarities have almost doubled in magnitude since 1990. In step three I quantify the macroeconomic effects of coworker complementarities. I structurally estimate the model, leveraging the estimated micro moments to impose discipline. A key take-away from counterfactual exercises is that stronger complementarities can account for one quarter to one half of the rise in the between-establishment share of wage inequality in Germany (1990-2010).

Elaborating on the first step, I start by setting out a micro-founded, tractable model of team production. A firm employs a team of employees to whom it assigns a continuum of tasks. These workers are ex-ante heterogeneous in terms of their overall quality type (vertically) and, conditional on that type, their efficiency varies across different tasks (horizontally). The key parameter of interest, labelled "task complexity" for short, controls the dispersion in worker-task specific productivities and thus captures a salient feature of the nature of work: Are tasks such that a high-quality worker is equally good – and thus more efficient than a low-quality worker – across all tasks; or are workers' skills specific to particular tasks. The greater this parameter, the more pronounced is comparative advantage and, hence, the more productivity hinges on who performs which tasks inside a firm.[4]

Under the optimal organization of production, wherein each worker is assigned those tasks in which they have a comparative advantage, the strength of coworker complementarities hinges on this task complexity parameter. I derive an *aggregation* result which maps workers' overall quality types onto team output, without having to explicitly keep track of the task assignment. The resulting team production function reveals that the elasticity of complementarity between coworkers' qualities is constant and increasing in the task complexity parameter. Stronger complementarities imply a greater weight on the quality of the least-capable team members in determining team output. Therefore, it is more costly in terms of output if high-quality employees are paired with low-quality coworkers. Intuitively, the former then have to spend more time with tasks at which they are relatively inefficient; coworker quality mismatch is thus tied to task mismatch.

To study matching patterns – who does, in fact, work with whom – and wage inequality, I embed this conception of team-based production into a continuous-time version of an equilibrium labor market sorting model à la Herkenhoff *et al.* (2022), wherein firms hire workers into teams in a labor market characterized by random search. Search costs mean that workers and firms agree on a range of mutually acceptable matches (Shimer and Smith, 2000). Differences in the strength of complementarities generate alternative equilibrium assignments of workers to firms.

For a stylized version of the model I provide *characterization* results that sharply summarize how the degree of complementarity shapes the sorting of workers into teams and the resulting distribution of wages. Crucially, coworker complementarities incentivize the formation of teams composed of workers of similar quality, translating at the macro-level into more pronounced segregation of workers and, concomitantly, into greater between-firm differences in productivity and wages.

The final theoretical result addresses a *measurement* problem and reveals an empirically feasible strategy for disciplining the model with micro data. In most production environments we do not have measures of individual output. Directly estimating coworker complementarities in production is accordingly almost impossible. Using the structural model, I show how they can be inferred indirectly. The key is that coworker complementarities in output are proportional to complementarities in wages. Given information on wages and mobility of workers across coworker groups, the latter

[4] Tractability is preserved by borrowing methodology from trade models à la Eaton and Kortum (2002). Worker-task specific efficiencies are draws from a Fréchet distribution. The average value of a draw is proportional to the worker's (one-dimensional) quality type. The task complexity parameter is the inverse of the Fréchet shape parameter.

can be measured using matched employer-employee panel data.

In a second step, and using this measurement result, the paper validates the key model mechanisms empirically and estimates coworker wage complementarities. To do so, I draw on rich, administrative micro data from Germany as well as – for robustness – from Portugal. The main empirical findings are as follows. First, coworker wage complementarities have approximately doubled from 1990 to 2010, paralleling the growing importance of complex tasks depicted in Figure 1a. Second, in the cross-section, occupations performing more complex tasks exhibit stronger coworker wage complementarities. In turn, coworker sorting is more pronounced in occupations with stronger complementarities. Lastly, industries that likely feature relatively high coworker complementarity also tend to be characterized by more coworker sorting and between-firm inequality.

In step three, I calibrate the model to fit micro- and macro-moments of the German economy in 1990 and in 2010. The structural estimation targets the evidence on coworker complementarities obtained from the micro data but deliberately leaves between-firm inequality or coworker matching patterns untargeted. The model fits the data well also along these untargeted dimensions.

Using the estimated model, I evaluate the macroeconomic implications of coworker complementarities. A first exercise considers the realized productivity costs of coworker quality mismatch. I find that misallocation along this dimension carries non-trivial costs: Coworker mismatch is estimated to have lowered per-capita output by about 2% in 1990, rising to 3% in 2010 due to strengthened coworker complementarities. The productivity losses would have been more severe had coworker sorting not increased over this time period.

Finally, I quantify the contribution of strengthened coworker complementarities to the firming-up of wage inequality in Germany. The estimated model predicts a 0.177 point increase in the between-establishment share of the variance of log (residual) wages, close to the 0.156 rise in the data. The baseline analysis implies that if coworker complementarities had, counterfactually, remained at their 1990-level, then the between-share of wage inequality would have increased only by half as much, Sensitivity analyses push the contribution of complementarities to just below 25%. Thus, a rise in coworker complementarities – as estimated from the micro data – can account for between one quarter and one half of the total rise in the between-share; the remainder is attributed to mechanical effects of skill-biased technological change.

RELATED LITERATURE. The paper connects with three broad literatures that are concerned, respectively, with organizational models of firm production, the theory of labor market matching, and empirical analyses of wage inequality and productivity.

From a theoretical perspective, this paper has long roots. In the spirit of Lucas (1978) and Rosen (1982), the organization of production among and the distribution of wages across ex-ante heterogeneous individuals is jointly determined by the equilibrium assignment of individuals to firms and tasks within firms. Different production functions generate alternative equilibrium assignments (cf. Garicano and Hubbard (2012)).

In particular, the paper relates to organizational models of the firm, which peer into the black box of the production function and study the implications of within-firm division of labor.[5] I build on Kremer (1993) and Neffke (2019) who argue that coworker interdependencies should be particularly significant in environments that are more "complex" (Kremer, 1993) or knowledge-intensive (Neffke, 2019). [6] In similarity to Garicano (2000), I do not assume a firm-level production function featuring complementarities but instead emphasize that they emerge endogenously from the optimal organization of production when coworkers are heterogeneous.[7] At the technical level, the theoretical team production block combines elements from task-assignment models à la Acemoglu and Autor (2011) and Acemoglu and Restrepo (2018) with an assumption taken from Eaton and Kortum's (2002) seminal trade model that facilitates a tractable analysis of assignment with a continuum of tasks and a discrete number of producers. Deming's (2017) study of social skills provided inspiration for an analogy to trade but the setup in that paper does not yield an aggregative team production function, nor does it focus on coworker complementarities or equilbrium matching. I contribute to this literature a micro-founded theory of coworker complementarities and show how their interaction with labor market frictions – the aforementioned papers assume frictionless labor markets – shapes macroeconomic outcomes. The paper also offers a strategy to measure complementarities, tests the theory in micro data, and provides a quantitative assessment.

As in the canonical, search-frictional heterogeneous-agent model à la Shimer and Smith (2000), the assignment of workers to firms in my model is shaped by complementarities between production partners' qualities.[8] In the standard model, however, the strength of production complementarities is an exogenous parameter, so that it is difficult to explain what underlying forces drive changes in sorting. I instead treat complementarities as endogenous by explicitly modelling the organizational problem of assigning tasks to workers. Furthermore, instead of considering complementarities between firm and worker attributes in one-worker-one-firm matches, this paper examines complementarities across coworkers' attributes in multi-worker firms.

[5] In addition to the papers discussed in the main text, important examples include Becker and Murphy (1992), Kremer and Maskin (1996), Garicano and Rossi-Hansberg (2006), and, more recently, Porzio (2017) and Bloesch *et al.* (2022).

[6] In Kremer's (1993) model, complementarities are stronger when the production process is complex in the sense of comprising a large number of tasks. I sever the tight link between team size and complementarities by endogenizing the number of tasks assigned to each team member. This setup makes it possible to study the implications of changing complementarities in a frictional general-equilibrium environment.

[7] Garicano's (2000) seminal paper centers on complementarities between managers and their supervisees, whereas my model is most naturally interpreted as capturing the interactions between workers within a given "layer."

[8] Selected, important contributions in the area include Eeckhout and Kircher (2011), Hagedorn *et al.* (2017), Lindenlaub (2017), Lopes de Melo (2018), Bagger and Lentz (2019), Bilal *et al.* (2022) and Elsby and Gottfries (2022).

The setup of the general-equilibrium block closely follows Herkenhoff *et al.* (2022). Their model is the first where (frictional) sorting is between workers within teams rather than between workers and firms.[9] While their primary focus lies on coworker learning, in a recently revised manuscript, Herkenhoff *et al.* (2022) likewise note that stronger supermodularity of the production function can qualitatively explain increased labor market segregation. I contribute a theoretical explanation for such a change in the production technology, derive how to measure such complementarities in micro data and use them to discipline the estimation, and provide a quantitative assessment of how much complementarities have indeed changed and what share of the "firming up" of inequality this mechanism explains.[10] Herkenhoff *et al.*'s (2022) model is richer in other dimensions, notably by including dynamic human capital spillovers. Their study highlights an important implication of stronger coworker sorting which this paper is silent on: it reduces the opportunities for low human capital workers to learn from better coworkers. Herkenhoff *et al.*'s (2022) analysis underscores the importance of understanding when and why coworker sorting occurs. In contributing such an explanation, I view my paper as complementary to theirs.

Turning to the empirics of wage inequality, different strands of the literature highlight structural shifts within and between-firms, respectively. One strand studies how technological shifts in the nature and organization of work within firms have altered returns to skills (Krusell *et al.*, 2000; Acemoglu and Autor, 2011). This literature documents how employees are increasingly performing complex, non-routine tasks.[11] These are typically associated with greater skill and training demands (e.g., Lin, 2011); they also frequently involve teamwork and put a premium on social skills (e.g. Deming, 2017). In parallel, but separately, a large number of papers, including notably Card *et al.* (2013) and Song *et al.* (2019), foregrounds the large and growing role of firms in explaining wage inequality.[12] This paper contributes a structural model that shows how the transformation of work, by amplifying coworker complementarities, helps rationalize rising between-firm inequality.

[9] Other papers building on that setup include Bodker (2019) and Hong (2022). Both consider interactions between worker-worker and worker-firm complementarities, the latter also studies learning. Relative to both, I contribute and test a theory of team complementarities and study aggregate productivity and rising between-firm inequality over time.

[10] Chade and Eeckhout (2020) provide an alternative, theoretical approach to analyzing assortative matching with teams. Their study shows how the interaction of production complementarities with externalities can give rise to stochastic matching. A version of their model in which knowledge spillovers in a firm's downstream market affect how each firm chooses the skill composition of its team likewise gives rise to the prediction that increased coworker complementarities tilt wage inequality toward the between-firm component. The differences between my paper and theirs parallel those listed in the discussion of Herkenhoff *et al.* (2022).

[11] On the historical evolution of the task content of production, see among many others, Autor *et al.* (2003), Spitz-Oener (2006), Atalay *et al.* (2020) and Autor *et al.* (2022). In Appendix D I corroborate these findings drawing on repeated, cross-sectional microdata collected in the German BIBB Employment Survey.

[12] Also see, for instance, Barth *et al.* (2016); Alvarez *et al.* (2018); Wilmers and Aeppli (2021); Criscuolo *et al.* (2021); Håkanson *et al.* (2021); Sorkin and Wallskog (2021) and Kleinman (2022).

2 Theory: team production and matching

This section develops the theoretical model and answered three successive questions. One, what are the origins and determinants of coworker complementarities? Two, what are the implications of a given degree of coworker complementarities for matching patterns (who works with and for whom), productivity, and between-firm inequality? And three, how can we measure coworker complementarities in the data?

2.1 Microfoundations for production complementarities in teams

I start by considering a single production unit, that is a firm who employs a team of ex-ante heterogeneous workers. I study the optimal assignment of tasks to these workers and derive how production depends on their quality types. The team's composition is exogenous in this section; the following section then considers hiring decisions, as well as wage-setting, and how those vary based on the production function derived here.

2.1.1 Setup

A team consists of a fixed number N of workers who perform tasks required for the production of a single good. The underlying production technology is owned by the firm. The employer allocates workers' time across tasks in order to maximize production. Each worker has a type $x_i \in [0,1]$, $i = 1, \ldots, N$, which denotes its rank in the worker quality distribution, and is endowed with one unit of time that they supply inelastically. It is notationally convenient to denote the set of team members by $\mathcal{S} = \{1, \ldots, N\}$.

FINAL GOODS PRODUCTION. The final good (or service) is produced from a unit mass of tasks $\tau \in \mathcal{T}$ (e.g., $\mathcal{T} = [0,1]$) according to a constant elasticity of substitution (CES) aggregator,

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

where $q(\tau)$ measures the amount of task τ used in production and $\eta > 0$ is the elasticity of substitution across tasks.

TASK PRODUCTION. Tasks are produced by workers using their fixed endowments of time and task-specific efficiencies ("skills"). In classic Ricardian fashion, the amount of task τ produced by worker i – the relation of which to $q(\tau)$ will be discussed shortly – is linear in efficiency units of labor,

$$y_i(\tau) = a_1 z_i(\tau) l_i(\tau), \tag{2}$$

where $z_i(\tau)$ is i's efficiency in producing task $\tau \in \mathcal{T}$, the parameter a_1 controls the sensitivity of task output to worker efficiency ($a_1 > 0$), and $l_i(\tau)$ is the time dedicated by i to task τ. With this notation, the time constraint for worker i is

$$1 = \int_{\mathcal{T}} l_i(\tau) d\tau. \tag{3}$$

TASK-SPECIFIC EFFICIENCIES. Workers are heterogeneous along two dimensions; vertically, as captured by their quality type x_i, and horizontally, meaning that conditional on x_i, a given worker may be better in some tasks than in others (similar to Grigsby (2022)). The task-specific efficiencies for worker i, $\{z_i(\tau)\}_{\tau \in \mathcal{T}}$, are treated as the realizations of a Fréchet-distributed random variable, drawn independently for each worker i. This assumption parallels trade models in the tradition of Eaton and Kortum (2002) and parsimoniously captures both vertical and horizontal differentiation among workers. Thus, for all $z \geq 0$, the distribution of efficiencies for worker i is

$$G_i(z) := Pr(z_i(\tau) \leq z) = \exp\left(-\left(\frac{z}{\iota x_i}\right)^{-1/\chi}\right). \tag{4}$$

A worker's type x_i determines the *scale* of the worker-specific distribution. The (inverse) shape parameter, $\chi \in [0, 1]$, determines the degree of dispersion and is identical for all workers. Finally, $\iota := \Gamma(1 + \chi - \eta\chi)^{\frac{1}{1-\eta}}$ is a convenient scaling term, where Γ denotes the Gamma function.[13]

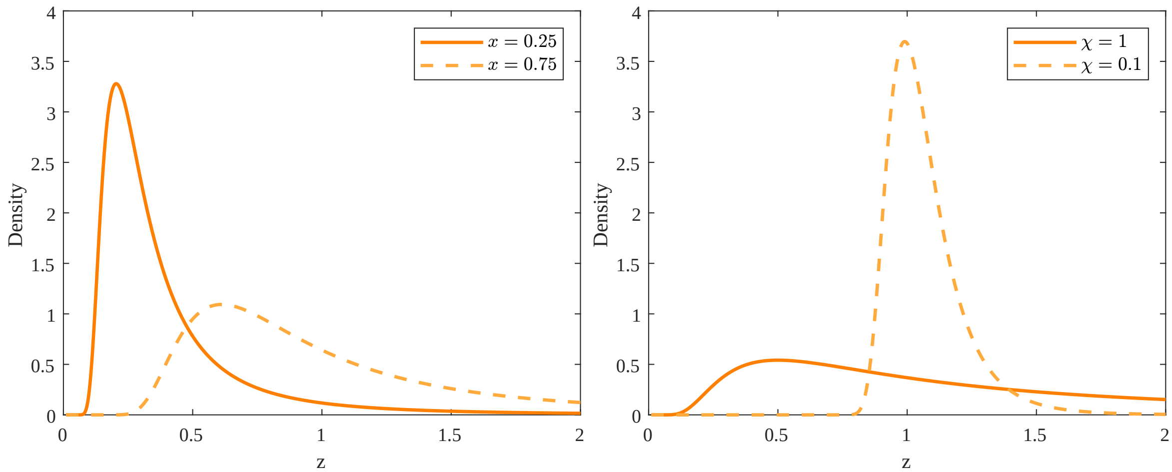
Two remarks are in order to unpack this description, which is at the heart of the model. First, equation (4) implies that workers (may) differ along both *absolute* and *comparative* advantage lines. Each worker can produce any task, but they are potentially heterogeneous in both their *average* task-specific production efficiency and, conditional on that average, in the distribution of efficiencies *across* different tasks. Variation in workers' quality types, x, connotes absolute advantage, insofar as the expectation of $z_i(\tau)$ is greater when x_i is higher. When $\chi > 0$, there is also dispersion in worker-task specific efficiencies, i.e., for a given value of x_i, a worker is better in some tasks than in others. Figure 2 illustrates how the distribution (density) of task-specific productivities varies with x and χ, respectively.[14]

Second, χ is the key parameter with respect to which I conduct comparative statics exercises. Formally, it controls the joint distribution of relative productivities for any given task τ across coworkers and, therefore, the strength of comparative advantage.[15] In terms of interpretation, χ

[13]This normalization ensures that varying χ or η does not mechanically change production levels. I assume that $1 + \chi - \eta\chi > 0$, which will ensure a well-defined cost index, introduced below.

[14]Edmond and Mongey (2021) likewise treat the shape parameter of the Freéchet distribution as measuring the strength of comparative advantage, while the scale parameter is a worker-specific of skill.

[15]To be precise, in the case of a univariate Fréchet distribution, χ controls two conceptually distinct distributions: the distribution of relative productivity between any two tasks for a given individual; and the joint distribution of relative



(a) Worker-specific scale parameter

(b) Technological (inverse) shape parameter

Figure 2: Illustration of properties of Fréchet distribution

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

captures the nature of production processes, which comprises, in reduced-form, the nature of both tasks and worker skills. As it shapes the dispersion of worker-task specific productivities, an increase in χ could, in principle, originate in shifting task requirements (e.g., fewer tasks which different workers are similarly productive in) or team members' skill acquisition (e.g., workers may concentrate their training on a narrower set of tasks to minimize fixed learning costs, cf. Rosen (1983) and Alon (2018)). In the spirit of Kremer (1993) and Garicano (2000), I label χ the "task complexity."[16]

ROLE OF FIRMS. Completing the model description is an account of how team production differs from multiple individuals producing the final good separately. When workers collaborate inside a team, rather than producing on their own, division of labor is possible. Formally, but very simply, the amount of any task available for final good production, $q(\tau)$, is the sum over task production by *all* team members,

$$q(\tau) = \sum_{i=1}^{N} y_i(\tau). \tag{5}$$

productivities for any given tasks across individuals and therefore the strength of comparative advantage. What matters here is the latter, i.e., that for any task, χ determines the slope of the relative productivity schedule for any pair of workers. The tight connection can be severed by using a multivariate Fréchet distribution, as Lind and Ramondo (2018) fruitfully do in the trade context, but the simpler version suffices for present purposes.

[16]In Kremer's (1993) model, wherein each worker performs one task, "complexity" refers to the number of distinct tasks, respectively team members. In Garicano (2000), greater complexity indicates the extent to which "unexpected" problems are confronted by the team.

Underlying this assumption is that in modern economies, an important role of firms lies in coordinating the collaboration between workers with specialized knowledge (cf. Becker and Murphy, 1992; Garicano, 2000). In practice, a firm not merely incorporates the intellectual property rights, or a "recipe", for a particular product; nor is a firm merely the sum of machines and tools with which workers are equipped. Much of the knowledge required for production is intangible and embedded in individuals with limited time to learn and work, as opposed to that knowledge being codified and tradeable on markets. Hence its mobilization and efficient use requires individuals to specialize and collaborate. In such environments, an important role of firms is to coordinate this process and assign tasks to workers with a comparative advantage in performing them.[17]

ORGANIZATIONAL OPTIMIZATION PROBLEM. The role of the firm as a coordinator is reflected in its optimization problem. The firm chooses total 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 total production Y, subject to the constraints (1)-(5). The associated Lagrangean is

$$\begin{aligned}\mathcal{L} = Y &+ \lambda\left[\left(\int_{\mathcal{T}} q(\tau)^{\frac{\eta-1}{\eta}}d\tau\right)^{\frac{\eta}{\eta-1}} - Y\right] \\ &+ \sum_{i=1}^{N}\left\{\lambda_i^L\left(1-\int_{\mathcal{T}} l_i(\tau)d\tau\right) + \int_{\mathcal{T}}\lambda_i(\tau)\left(a_1 z_i(\tau)l_i(\tau) - y_i(\tau)\right)\right. \\ &+ \left.\int_{\mathcal{T}}\tilde{\lambda}(\tau)\left(\sum_{i=1}^{N} y_i(\tau) - q(\tau)\right)d\tau + \int_{\mathcal{T}}\bar{\lambda}_i(\tau)y_i(\tau)d\tau\right\},\end{aligned}$$

where λ, λ_i^L, $\lambda_i(\tau)$, $\tilde{\lambda}(\tau)$ and $\bar{\lambda}_i(\tau)$ are Lagrange multipliers. The first four 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. The final multiplier applies to a non-negativity constraint in task production.

2.1.2 Solving the organizational problem

To solve for the optimal production plan we can proceed in two basic steps. First, we derive the demand for tasks for a given set of shadow prices (treating those as known). Then we determine these shadow prices given the distribution of task-specific productivities across workers.

The first-order-condition (FOC) with respect to $q(\tau)$ is

$$\tilde{\lambda}(\tau) = \lambda\left(\int_{\mathcal{T}} q(\tau)^{\frac{\eta-1}{\eta}}d\tau\right)^{\frac{\eta}{\eta-1}-1} q(\tau)^{-1/\eta} \tag{6}$$

[17] Rivkin and Siggelkow (2003, 292), quoted in Dessein and Santos (2006) write: "The [qualitative management] literature is unified in what it perceives as the central challenge of organizational design: to divide the tasks of a firm into manageable, specialized jobs, yet coordinate the tasks so that the firm reaps the benefits of harmonious action."

Substituting for Y from equation (1) and defining $Q(\tau) := \tilde{\lambda}(\tau)q(\tau)$, we obtain

$$\frac{Q(\tau)}{\lambda Y} = \left(\frac{\tilde{\lambda}(\tau)}{\lambda}\right)^{1-\eta}. \tag{7}$$

This is the planner analogue to the standard iso-elastic demand implied by the CES aggregator.

Now integrate on both sides and use that the shadow value of all tasks used is related to total production (as derived in Appendix B.1.1), which yields

$$Q := \int_{\mathcal{T}} Q(\tau)d\tau = \lambda Y, \tag{8}$$

and, hence, $Y = \frac{Q}{\lambda}$, to obtain an expression for the shadow cost index:

$$\lambda = \left(\int_{\mathcal{T}} \tilde{\lambda}(\tau)^{1-\eta} d\tau\right)^{\frac{1}{1-\eta}}. \tag{9}$$

Who should produce which tasks, and what does that imply for the shadow costs faced by the firm, i.e., for $\{\tilde{\lambda}(\tau)\}_{\tau\in\mathcal{T}}$? To answer this question, we first observe that the FOC w.r.t. $y_i(\tau)$ is

$$\bar{\lambda}_i(\tau) + \tilde{\lambda}(\tau) = \lambda_i(\tau). \tag{10}$$

so that $\tilde{\lambda}(\tau) = \lambda_i(\tau)$ 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 the shadow value of a task be equal to the minimum shadow cost of producing it across all team members,

$$\tilde{\lambda}(\tau) = \min_{i\in\mathcal{S}} \left\{\lambda_i(\tau)\right\}. \tag{11}$$

Worker i produces task τ if they attain this minimum.

Next, we characterize the term inside $\{\cdot\}$. If $y_i(\tau) > 0$, straightforward manipulation of the FOC for $l_i(\tau)$ yields

$$\lambda_i(\tau)\frac{y_i(\tau)}{l_i(\tau)} = \lambda_i^L, \tag{12}$$

To relate the shadow price of producing the task to the price cost of the (labor) input used to produce it, substitute for $l_i(\tau)$ in the task-level production function and rearrange to find

$$\lambda_i(\tau) = \frac{\lambda_i^L}{a_1 z_i(\tau)}. \tag{13}$$

Thus, the shadow price of task τ produced by worker i is the ratio of the shadow value of i's time and i's efficiency in producing that task. Using this result in equation (11) yields

$$\tilde{\lambda}(\tau) = \min_i \left\{ \frac{\lambda_i^L}{a_1 z_i(\tau)} \right\}. \tag{14}$$

The next step is to derive what share of tasks is performed by each worker and at what (shadow) cost final goods can be produced.[18] The key results are summarized in the following lemma.

Lemma 1. *Suppose that workers' task-specific efficiencies are independently Fréchet-distributed. Then:*

(i) *The shadow cost index is*

$$\lambda = a_1^{-1} \left(\sum_{i \in \mathcal{S}} \left(\frac{x_i}{\lambda_i^L} \right)^{1/\chi} \right)^{-\chi}. \tag{15}$$

(ii) *The fraction of tasks for which i is the least-cost provider is*

$$\pi_i := Pr\{\lambda_i(\tau) \leq \min_{k \in \mathcal{S} \setminus i} \lambda_k(\tau)\} = \frac{(x_i/\lambda_i^L)^{1/\chi}}{\sum_{k \in \mathcal{S}} \left(x_k/\lambda_k^L\right)^{1/\chi}}. \tag{16}$$

(iii) *The shadow value of all tasks used in final goods production that were produced by worker i, defined as $Q_i := \int_{\mathcal{T}} \tilde{\lambda}(\tau) y_i(\tau) d\tau$, is a fraction π_i of the total shadow value of tasks used:*

$$Q_i = \pi_i Q. \tag{17}$$

Proof. Appendix section B.1.2. □

Using these results, and after normalizing the shadow price of final goods output to unity ($\lambda = 1$), we can characterize the optimal organization of production and its implication for coworker complementarities.

[18] That this class of problem, with arbitrarily many potential producers of a continuum of tasks, is tractable when the task-specific efficiencies are Fréchet-distributed is a key insight of Eaton and Kortum (2002).

2.1.3 Characterizing the optimal organization of production

We can develop intuition for the properties of the optimal organization of production in three steps. The first property restates a canonical result (e.g. Garicano, 2000; Duranton and Puga, 2004): The optimal organization of production features complete division of labor whereby each worker is assigned those tasks in which she has a comparative advantage. Inspecting equation (14) reveals that any task is produced by whichever worker is most efficient in producing that task *relative* to the (task-invariant) shadow cost of their time.[19]

It is equally intuitive that better workers perform more tasks, other things equal. The share of tasks produced by worker i is $\pi_i = \left(x_i^{\frac{1}{1+\chi}}\right)\left(\sum_{k\in\mathcal{S}}(x_k)^{\frac{1}{1+\chi}}\right)^{-1}$, so that $\frac{\partial \pi_i}{\partial x_i} > 0$. Clearly, when team members are differentiated only in their task-specific productivities and $x_i = x\ \forall i \in \mathcal{S}$, it follows that $\pi_i = 1/N$.

Most importantly, insofar as χ exceeds zero, the optimal organization of production renders the value of any worker's time a function not only of her own endowments but also of her environment. In our simple model, "her environment" simply refers to the number and abilities of her coworkers. Thus, notice that the shadow value of all tasks produced by worker i is

$$Q_i = N^{1+\chi} a_1 x_i^{\frac{1}{1+\chi}} \tilde{x}_i^{\chi}, \tag{18}$$

where $\tilde{x}_i = \frac{1}{N}\sum_{k\in\mathcal{S}} x_k^{\frac{1}{1+\chi}}$. Intuitively, and other things equal, the more an organization permits a worker to focus on those tasks in which she is particularly efficient, the greater the total value of tasks she can produce. For a given team size, the better the coworkers of individual i, the more worker i can focus on the tasks in which she is relatively more efficient. Such coworker interdependencies are more marked when χ is high. For suppose, to the contrary, that a worker is equally productive in all tasks ($\chi = 0$). Then it does not matter which tasks she performs, neither for her productivity nor for total team output. It is when individuals specialize in particular tasks, or when comparative advantage is pronounced, that their productivity will depend on the quality of her coworkers (cf. Neffke, 2019). Crucially, as is clear from equation (18), the absolute gain from having better coworkers is greater for high-quality workers. Put simply, more output is at stake.

The following aggregation result summarizes how output depends on the team composition.

Proposition 1 (Aggregation result)**.** *Team output Y can be written as a function of members' quality*

[19] Strictly speaking, we would need to specify a tie-breaking rule for tasks where multiple workers have the same shadow cost. But those tasks have mass zero anyway.

types, $f : [0,1]^N \to \mathbb{R}_+$*,*

$$f(x_1, \cdots, x_N) = \underbrace{N^{1+\chi}}_{\text{efficiency gains}} \underbrace{\left(\frac{1}{N}\sum_{i=1}^{N}(a_1 x_i)^{\frac{1}{1+\chi}}\right)^{1+\chi}}_{\text{complementarities}}. \tag{19}$$

Proof. See Appendix Section B.1.3. □

This result shows that conditional on a value of χ, knowing the number and quality types of team members is sufficient to pin down output under the optimal organizational plan.[20]

The first term summarizes the benefits from team production relative to each worker producing all tasks themselves, which are increasing in χ. A hypothetical scenario in which firms do *not* implement any division of labor serves as a useful benchmark. Supposing that each worker produces all tasks (in proportions minimizing their respective shadow costs of producing the final good), it is straightforward to show that total output is simply $\sum_{i=1}^{N} a_1 x_i$. These efficiency gains from the division of labor echo a long-standing theoretical literature (e.g., Smith, 1776; Becker and Murphy, 1992; Garicano, 2000) and cohere with recent empirical evidence (Chaney and Ossa, 2013; Tian, 2021; Adenbaum, 2022; Boehm and Oberfield, 2022; Kuhn *et al.*, 2022; Minni, 2022).[21]

The second term takes the form of a CES function. The following corollary sharply summarizes coworker interdependencies arising from the efficient organization of production.

Corollary 1 (Tasks and coworker complementarities)**.** *The elasticity of complementarity,* $\gamma = \frac{\chi}{1+\chi}$*, is constant and increasing in the value of* χ*.*

Intuitively, the more *horizontally* differentiated coworkers are at the micro-level of *tasks*, the more limited the scope for substitution is in terms of workers' *absolute advantage* at the team-level.[22]

[20]Notice that the elasticity of substitution across tasks, η, does not show up in equation (19). Technically, as noted also by Eaton and Kortum (2002, Footnote 18), this holds as long as we maintain that $1 + \chi - \eta\chi > 0$, 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, to be sure, a tight implication of the Fréchet, which serves to sharply bring into relief that coworker complementarities do not hinge on the assumption, maintained by Kremer (1993), that tasks combine in a Leontief fashion. On this point, also see Jones (2011). In a more general setting, the strength of coworker complementarities would also be influenced by the value of η.

[21]A subtlety warrants comment. Heterogeneity in task-specific productivities, or comparative advantage, can be "natural" or "acquired" (Grossman and Helpman, 1990). The former is traditionally associated with Smith (1776), whose famous pin factory examples points to productivity gains when workers spend more time on each task, for example due to learning-by-doing or saved costs from switching tasks. The model instead ascribes these efficiency gains to exogenous comparative patterns among individuals. This Ricardian specification affords greater tractability, but the general reasoning in this paper is entirely consistent with acquired comparative advantage. In such a world, as considered for instance by Costinot (2009), greater task complexity (χ) would represent higher fixed costs of learning a task or greater scope for learning-by-doing, equally giving rise to gains from the division of labor.

[22]This aggregation result builds on the seminal paper of Houthakker (1955) and, more recently, Acemoglu and Restrepo (2018), Martinez (2021) and Ocampo (2022). A similar micro-foundation for an aggregate CES production

This relationship resonates with the empirical findings of Jäger and Heining (2022). Simply put, even the most talented worker may lose much time when having to perform a task that they have not trained in.

The flip-side of such coworker complementarities is that productivity is lowered if team members vary in their quality type. One way of seeing this is to observe that in statistical terms, the constant-elasticity term in equation (19) corresponds to the generalized mean – also known as Hölder or power mean – of the underlying x_i values (cf. Jones, 2011).[23] Indeed, this term reduces to the arithmetic mean when $\gamma = 0$, the geometric mean when $\gamma = 1$, and the harmonic mean when $\gamma = 2$. The Leontief case, where by team-output is determined by the minimum value of x_i, obtains for $\gamma \to \infty$; the reverse, "superstar" case (Rosen, 1982) holds when $\gamma \to -\infty$. By the power mean inequality, the general property is that higher values of χ, and hence of γ, put greater weight on the ability of the least-capable team member.[24] Greater values of χ are, therefore, not only associated with pronounced efficiency gains from team production. Once within-firm worker and job heterogeneity is allowed form, we find that larger values of χ also render *coworker mismatch* – dispersion in team members' ability – more costly. In short, the composition of a team becomes more important.[25]

2.2 How complementarities shape team composition in equilibrium

In a second step, we want to understand the implications of coworker complementarities for how firms choose whom to hire. Thus, instead of looking at one team of a given composition, suppose now that there are *many* ex-ante heterogeneous workers and many potential employers, whose matching decisions are endogenous. Which combinations of workers would different multi-worker firms hire if production operates as described in the preceding section – and how do matching patterns and wage distribution respond to changes in χ? To answer this question, I will embed the

function is independently derived in Dvorkin and Monge-Naranjo (2019).

[23] Of course, the microfoundations strictly speaking impose an upper bound on γ because $\chi \leq 1$. Task complementarities may not be too strong, lest the gross substitutability condition, which assures uniqueness of the optimum, be violated (cf. Alvarez and Lucas, 2007).

[24] Proposition 1 offers a microfoundation for the reduced-form model of the relationship between individual and team outcomes assumed by Ahmadpoor and Jones (2019) in their study of team production and matching in science and invention. They postulate that team output corresponds to $\beta_N \left[\frac{1}{N}\sum_{i=1}^{N} x_i^{\zeta}\right]^{1/\zeta}$, where β_n "captures impact benefits associated with teamwork" while the second term measures by how much these benefits diminish as the gap between team members' abilities grows.

[25] An implicit premise of the model here is that the division of labor incurs no output losses due to coordination frictions. Yet there is a compelling case that implementing the division of labor may limit the amount of time available for task production because of communication requirements (Becker and Murphy, 1992; Deming, 2017). Freund (2022) generalizes the model presented here and incorporates coordination frictions. I show, among other things, that the qualitative association between χ and coworker complementarities exists unless coordination is so costly that it is optimal to implement no division of labor whatsoever. This generalization also implies other predictions that are of independent interest. For instance, the importance of organizational quality for team productivity is increasing in χ.

aggregative team production model into general-equilibrium environment where employers and employees match on a search-frictional labor market.

Before doing so, it is instructive to pause briefly and consider, what matching patterns would emerge if we assumed a frictionless labor market instead (as in Kremer (1993), for instance)? The answer is that for any value of χ strictly greater than zero the equilibrium (and socially optimal) allocation would feature *pure positive assortative matching*. As Corollary 1 tells us, workers' qualities are complements in production unless $\chi = 0$. And when a worker's marginal productivity is increasing in coworkers' quality, a feasible Pareto improvement exists unless all team mates of a worker of type x are likewise of type x. The reason is simple. Even if both a low-x and a high-x are more productive in the presence of a better coworker, complementarities mean that the benefit is differentially larger for the high type. Hence, in this setting, the joint distribution of workers' and coworkers' talents would be degenerate and the matching solution is characterized by a function $\mu(x) = x$, where $\mu : [0, 1] \to [0, 1]$ maps a worker's type into the average type of her coworkers. Moreover, in such a frictionless matching environment an increase in χ would have no impact on matching patterns (unless $\chi \leq 0$ initially).

Yet, this prediction of pure positive sorting is clearly counterfactual – the correlation between coworkers' types depicted in Figure 1b would need to equal one. I will therefore allow for what is probably the most widely studied friction leading to stochastic matching, namely search costs (Shimer and Smith, 2000). In the presence of search frictions, some degree of "mismatch" between team members' types is tolerated, since waiting is costly. How much mismatch there will be in equilibrium crucially depends on the strength of complementarities and, thus, on χ.

2.2.1 Environment: agents, production & search technology, distributions

I embed the team production function into a version of Herkenhoff *et al.*'s (2022) matching model. To focus on the key mechanism of changing production complementarities, I abstract from learning and from on-the-job search. These simplifications permit a tight characterization of the equilibrium properties in a stylized version of the model. They also make it possible to derive a novel measurement result that is informative about how to discipline the strength of complementarities in the model using micro data.[26]

Time is continuous. All agents are infinitely-lived, risk-neutral, and maximize the present value of payoffs, discounted with a common rate $\rho \in (0, 1)$. There is a unit mass of workers, denoted $d_w = 1$, who are either employed (e) or unemployed (u). As before, workers are ex-ante heterogeneous with respect to their productivity types, $x \in [0, 1]$. As production is increasing in each argument, we may

[26]My notation is different and seeks to align my setup as closely as possible with the well-understood, standard model of matching between workers and *one-worker* firms, especially Hagedorn *et al.* (2017).

consider x a worker's rank in the underlying productivity distribution. Hence, the distribution of worker types is uniform. For ease of exposition, I will often elliptically refer to a "worker x" instead of a "worker of type x."

There is a mass d_f of firms that are either vacant (v) or producing (p). I write down the model in a general form, allowing firms to be ex-ante heterogeneous in their productivity, too. So, a firm's type is $y \in [0, 1]$. Allowing for ex-ante firm heterogeneity in this way is straightforward and underscores that formal results proved here do not hinge on such an assumption; it also facilitates robustness exercises. In the baseline quantification, though, I study how large differences across firms are even when shutting down ex-ante firm heterogeneity.

Production can take place with a single employee or with a team. The maximum team size is normalized to $N = 2$, i.e., there are sharply decreasing returns. This assumption is evidently unrealistic and made for sake of tractability. Nevertheless, it does not prevent us from studying how coworker complementarities shape between-firm inequality. Denote the production function of a single-worker firm $f(x, y) : [0, 1]^2 \longrightarrow \mathbb{R}_+$, short for $f(x, y, \emptyset)$, and that of a two-worker firm is $f(x, y, x') : [0, 1]^3 \longrightarrow \mathbb{R}_+$. Throughout, I refer to the coworker type by x'. The order of arguments in any reference to multi-worker firms – wherein we first reference one worker, then the firm and then the other worker – may seem a bit awkward at first but will be convenient later. The value of the aggregation result in Proposition 1 is that we do not need to keep track of worker-task specific efficiencies or assignments – that happens "in the background" and we care only about the implications for coworker complementarities.

POPULATION COMPOSITION. The measure of producing matches consisting of a firm of type y and one worker of type x is $d_m(x, y)$ and $d_m(x, y, x')$ is the corresponding measure of matches with an additional coworker x'.[27]

The following adding-up property holds for workers of type x:

$$d_w(x) = d_u(x) + \int d_m(x, \tilde{y})d\tilde{y} + \int \int d_m(x, \tilde{y}, \tilde{x}')dyd\tilde{x}', \tag{20}$$

where $d_u(x)$ is the measure of unemployed workers of type x. The aggregate unemployment rate is $U = \int d_u(\tilde{x})d\tilde{x}$.

Similarly for firms,

$$d_f(y) = d_v(y) + \int d_m(x, y)d\tilde{x} + \frac{1}{2}\int \int d_m(\tilde{x}, y, \tilde{x}')d\tilde{x}d\tilde{x}'. \tag{21}$$

The division by 2 in the last term avoids double-counting since $d_m(x, y, x')$ is symmetric in x and

[27] We can think of $d_m(x, y)$ as shorthand for $d_m(x, y, \emptyset)$.

x'.

TIMING. Within any increment of time, each worker may be exogenously separated from their employer at an exogenous Poisson rate δ. Next, unmatched workers and all firms engage in random search and matching decisions. Then, production and surplus sharing happen.[28] Every unemployed worker and all firms engage in search. Unemployed workers contact a firm at an exogenous Poisson rate M_u. Hence, for any firm the probability of meeting any unemployed worker is $M_f = M_u U$; the probability of meeting a worker $x \in \tilde{X} \subset [0,1]$ is $M_f \frac{\int_{\tilde{X}} d_u(x)dx}{U}$. Upon meeting, the types x and y are perfectly observed, but the task-specific productivities are revealed to the firm only when the hiring decision is made.

2.2.2 Definition and sharing of joint surplus

SURPLUS DEFINITIONS. The joint value of a pair (x, y) is

$$\Omega(x,y) = V_p(x,y) + V_e(x,y), \tag{22}$$

where $V_e(x,y)$ is x's value of being employed at y while $V_p(x,y)$ is the value of y producing with x.

The surplus of a match between a worker x and a firm y that has no other employees is

$$S(x,y) = \Omega(x,y) - V_v(y) - V_u(x), \tag{23}$$

where $V_u(x)$ is the value of unemployment for x; and $V_v(y)$ is the value of y's vacancy.

The joint value of a firm y with team (x, x') is

$$\Omega(x,y,x') = V_p(x,y,x') + V_e(x,(y,x')) + V_e(x',(y,x)), \tag{24}$$

with $V_e(x,(y,x'))$ denoting the value of x being employed at y together with a coworker x'.

Then the ("marginal") surplus from a firm y that has employee x' hiring a type x worker is:[29]

$$S(x,(y,x')) = \Omega(x,y,x') - \Omega(x',y) - V_u(x). \tag{25}$$

This surplus will rise in the output generated by x and y, yet fall in both sides' option values. Notice, furthermore, that this surplus is *not* symmetric in the two worker-related arguments even if $\Omega(x,y,x')$ is, because it matters for both outside options whether type-x is the potential employee

[28]To make the point of this paper, it is sufficient to consider a world in which there are no endogenous separations nor is it permissible to replace a worker with a better matched candidate. I have explored these extensions and may add them back if doing so proves relevant.

[29]Since this is a discrete choice, the use of the term "marginal" is loose.

or type-x' is (unless $x = x'$).

SURPLUS SHARING. Wages are continuously renegotiated. Negotiation takes place over the entire surplus and the firm treats each employee as marginal, i.e., their outside option is unemployment. Bargaining takes the generalized Nash form. The bargaining power of workers is $\omega \in [0, 1]$.

The wage $w(x, y)$ of a worker of type x employed at a firm of type y satisfies

$$(1-\omega)\big(V_e(x,y) - V_u(x)\big) = \omega\big(V_p(x,y) - V_v(y)\big), \tag{26}$$

In a two-worker firm, the wage $w(x, (y, x'))$ of a worker of type x being employed at firm of type y with a coworker of type x' satisfies

$$(1-\omega)\big(V_e(x,(y,x')) - V_u(x)\big) = \omega\big(V_e(x',(y,x)) + V_p(x,y,x') - V_e(x',y) - V_p(x',y)\big). \tag{27}$$

This surplus sharing protocol is consistent with the assumptions set out by Bilal *et al.* (2022), such that all worker and firm decisions are characterized by their joint surplus, which in turn only depends on individual types.

2.2.3 Equilibrium conditions

We consider a stationary equilibrium. This equilibrium is characterized by a set of Hamilton-Jacobi-Bellman equations (HJBs) that describe agents' optimal matching strategies together with Kolmogorov-Forward equations (KFEs) that pin down the stationary distributions.

OPTIMALITY CONDITIONS. Starting with unmatched agents, the asset value of an idle firm y satisfies

$$\rho V_v(y) = (1-\omega) M_f \int \frac{d_u(\tilde{x})}{U} S(\tilde{x}, y)^+ d\tilde{x}, \tag{28}$$

where notation is eased by letting, for any r, $r^+ = \max\{r, 0\}$ indicate the optimal choice. The discounted value thus corresponds to the weighted conditional expectation of its share of the surplus of a match with an unemployed worker (i.e., conditional on the surplus being at least zero) times the unconditional probability of meeting any unmatched worker.

The value of such an unemployed worker x follows

$$\rho V_u(x) = b(x) + M_u \omega \left[\int \frac{d_v(\tilde{y})}{V} S(x, \tilde{y})^+ + \int \frac{d_m(\tilde{x}', \tilde{y})}{V} S(x, (\tilde{y}, \tilde{x}'))^+ d\tilde{x}' \right] d\tilde{y}. \tag{29}$$

Notice that we need to take into account the worker's flow value from home production, $b(x)$, as well as differentiating between the worker meeting an unmatched firm *or* a one-worker firm.

What about the *matched* agents? To understand the model mechanics, it is most instructive to consider the HJBs for the joint values of firm and worker(s), starting with the joint value of a firm y with a pair of workers x and x':

$$
\begin{aligned}
\rho\Omega(x,y,x') = f(x,y,x') + \Big[&\delta(-\Omega(x,y,x') + \Omega(x,y) + V_u(x')) \\
&+ \delta(-\Omega(x,y,x') + \Omega(x',y) + V_u(x))\Big].
\end{aligned} \tag{30}
$$

The discounted value contains the flow value of production, and at a rate δ either x or x' leaves the firm. Notice that in case of separation, the change in values is equal to the marginal surplus $S(x',(y,x))$ if x' leaves and equal to $S(x,(y,x'))$ if x is separated.

Crucially, the joint value of a firm y with employee of type x, which satisfies

$$
\begin{aligned}
\rho\Omega(x,y) = f(x,y) + \delta\big[-\Omega(x,y) + V_u(x) + V_v(y)\big]& \\
+ M_f \int \frac{d_u(\tilde{x}')}{U} \max\{-\Omega(x,y) + V_e(x,(y,\tilde{x}')) + V_p(x,y,\tilde{x}'),0\}.d\tilde{x}'&
\end{aligned} \tag{31}
$$

The discounted value contains the flow value of production. At rate δ the match is exogenously destroyed. But with (unconditional) probability M_f, the firm meets an unemployed worker. The conditions under which such a meeting leads to a match are of critical interest to us: What *type* of worker is a firm y with employee x willing to hire (and conversely, which job would the candidate x' be willing to accept)? The provisional answer is that a match is formed if the sum of values accruing to x and y when teaming up with x' exceed their joint value if x' is not hired; noting that the latter value comprises, amongst others, the option value of waiting to meet a worker other than x' who is a better match. Since $\big(-\Omega(x,y) + V_e(x,(y,x')) + V_p(x,y,x')\big)$ is proportional to $S(x',(y,x))$ – as Appendix B.2.2 shows – an equivalent statement is that a worker x' will be hired if the marginal surplus from doing so is positive. (That is, as in Bilal *et al.* (2022), hiring decisions are fully characterized by the joint marginal surplus.)

POPULATION DYNAMICS. In stationary equilibrium, the inflows and outflows into different states balance each other. These flows are characterized by a set of KFEs and they interact with the value functions because hiring decision functions are pinned down by the surplus values:

$$h(x,y) = \mathbf{1}\{S(x,y) > 0\}, \tag{32}$$

$$h(x,(y,x')) = \mathbf{1}\{S(x,(y,x')) > 0\}. \tag{33}$$

These policy functions describe, respectively, whether a vacant firm y is willing to hire a worker x;

and whether a firm y that already employs worker x' is willing to hire a worker x.[30] The bargaining protocol ensures that decisions are mutually agreed upon, i.e., they are privately efficient. As the KFEs are straightforward but clunky, they are collection in Appendix B.2.1.

2.2.4 Equilibrium definition

In stationary equilibrium, agents' values must be consistent with the distributional dynamics to which they give rise, and vice versa. Appendix B.2.2 derives the recursions for the surplus values, see specifically equations (B.10) and (B.13), so that the optimality conditions are fully described in terms of the values of unmatched agents and surpluses. A formal definition follows.

Definition 1. *A stationary search equilibrium is a tuple of value functions $(V_u(x), V_v(y), S(x,y), S(x,(y,x')))$ together with a distribution of agents across states, $(d_m(x,y), d_m(x,y,x'))$ such that (i.) the value functions satisfy the HJBs* (28), (29), (B.10) *and* (B.13) *given the distribution; and (ii.) the stationary distributions satisfy the KFEs* (B.7)-(B.8) *given the policy functions implied by the value functions according to equations* (32)-(33).

This equilibrium needs to computed numerically because of a non-trivial general-equilibrium interaction: Agents' expectations and matching decisions must conform with the distribution to which they give rise, yet as that distribution evolves, so do agents' expectations over future meeting probabilities and, hence, their optimal actions.

2.2.5 Elucidating the mechanisms

The primary margin of interest in the model concerns the matching decisions of a firm y which already employs a worker of type x – we may want to think of x as the "representative employee" of the firm – and an unmatched worker of type x' after they met.[31] These agents need to decide whether to either match, in which case the team (x,x') produces and the output is shared, or instead to wait for another production partner with whom the respective parties can generate a larger surplus.

This tradeoff between match quality and search costs is shaped by the strength of production complementaries. For as we have seen, mismatch is more costly when χ and hence the elasticity of complementarity γ is high. This means, for instance, that if x is a good type and the potential coworker x is a bad type, then the surplus from such a match would be very low, because strong complementarities mean that the value of tasks produced by x, while high in potential, will be

[30] I impose that matching only takes place when the relevant surplus is strictly greater than zero; since the indifference case occurs for a measure zero of agents, this assumption has no bearing on the result.

[31] Especially when firms are ex-ante homogeneous, or at least very similar in productivity, the decision of an unmatched firm and an unmatched worker is quite trivial, as there is little variation in surplus across potential matching combinations. Indeed, in the numerical solution of the model we generally have that $h(x,y) = 1$ for any (x,y).

dragged down by the low productive capacity of x'. On the other hand, if χ is low, x will simply perform a greater share of tasks, without that lowering the average efficiency with which she performs these tasks. Consequently, to avoid high costs of mismatch in a high-χ economy, similar types are more likely to be matched together than distant types, i.e., there is positive assortative matching. Such coworker sorting translates into greater differences across firms in the quality of their workforce, productivity, and average wage.

To characterize this mechanism more sharply, in Appendix B.3 I consider a stylized version of the model. It embodies the same logic as the full model but can be solved by hand. Specifically, I derive closed-form results for matching decisions and the distribution of worker types as well as wages across firms.[32] Here I briefly sketch that stylized model and summarize key findings. The stylized model adapts the setup of Eeckhout and Kircher (2011) to the context of matching into teams. As in that paper, it is assumed that if agents decide against a match upon meeting, they incur a fixed search cost and, in return, get matched with their optimal match partner. It turns out that in this environment, the matching decision is characterized by a simple threshold rule. A firm who is already employing a worker of type x' will hire a worker of type x *if* the absolute distance between the two types, $|x - x'|$, is lower than a threshold value s^*. This threshold is *decreasing* in the strength of complementarities.[33]

The following statement summarizes the predictions of the simplified model. Appendix B.3 derives them formally.

Corollary 2 (Coworker complementarities and distributional outcomes)**.** *Other things equal, stronger coworker complementarities are associated with a lower value of the equilibrium threshold s^*, leading to the following outcomes.*

(i) *Coworker sorting is stronger, i.e., the correlation between coworkers' types is higher. Formally, the Pearson correlation is $\rho_{xx} = (2s^* + 1)((s^*)^2 - 1)^2$.*

(ii) *The average coworker type is lower for types of quality below s^*, greater for types above $1 - s^*$, and unchanged for intermediate types. Formally, the expected coworker type is*

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

[32] Numerical exercises confirm that the predictions of the full and the stylized model are qualitatively very similar.

[33] Similarly, Coles and Francesconi (2019) point out that in search-and-matching models with ex-ante heterogeneous agents, "much can be learned about the structure of steady state equilibria from considering the partial equilibrium conditions [for an exogenously given population of unmatched agents] in isolation." The same reasoning applies here, except that I do not adopt a "clones assumption" (Burdett and Coles, 1999) but, instead, use Eeckhout and Kircher's (2011) trick to gain tractability.

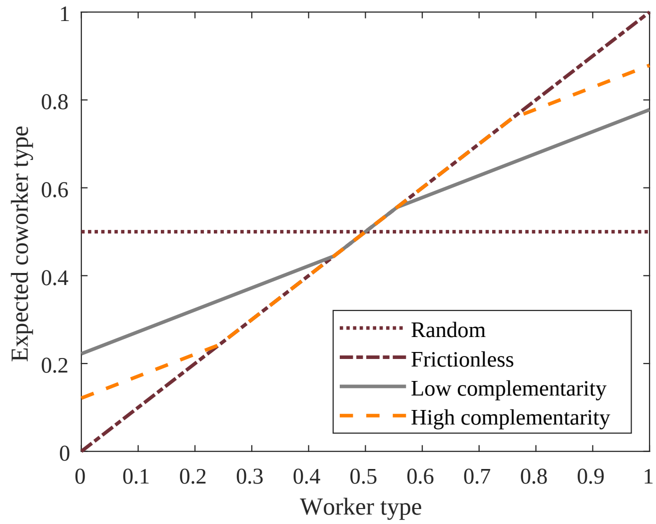


Figure 3: Illustration of matching patterns in the stylized matching model

Notes. This figure plots, for each worker type x, the expected coworker type, $\hat{\mu}(x)$. Each line represents an alternative equilibrium assignment, as summarized by the threshold s^*. Specifically, "Random" corresponds to $s^* = 1$; "frictionless": $s^* = 0$; "Low complementarity": $s^* = 0.25$; "High complementarity": $s^* = 0.125$.

(iii) *The between-firm share of the variance of wages is higher. (A closed-form but aesthetically less than appealing result can be found in the appendix.)*

Figure 3 graphically illustrates the equilibrium matching decisions through the lens of result 2(ii). For any worker type (horizontal axis), it plots the *expected* coworker type, $\hat{\mu}(x)$, with each line representing a different equilibrium. Consider first the case of random matching, which obtains when types are perfect substitutes. Then, as the dotted line illustrates, $\hat{\mu}(x)$ is the same for any x. Next, consider the 45-degree line. It shows how, in the absence of any search frictions, there is pure assortative matching, i.e., $\hat{\mu}(x) = x$. The remaining two lines represent in-between cases and illustrate the *nonlinearity* implied by result 2(ii). First consider the solid line. Search frictions mean that worker types below the threshold s^* are sometimes paired up with coworkers better than themselves. The opposite applies to high types, that is those above $1 - s^*$. For types in between those two kink points, the average coworker type corresponds to their own type. Now consider the orange line, which illustrates an environment with stronger complementarities and, hence, a lower threshold. For a worker in the middle of the distribution, the matching set shrinks symmetrically in both directions. But for the very best types, stronger complementarities can only raise the minimum type with whom they get matched in equilibrium, so that the *average* coworker type is higher than under weak complementarities. The converse applies to low types.

In summary, when teams are formed in a search-frictional labor market, then a strengthening of complementarities implies that the matching equilibrium increasingly features "superstar teams" composed of the best workers, on the one hand, and "laggard teams," on the other hand (Andrews *et al.*, 2019; Autor *et al.*, 2020).[34] From a theoretical perspective, this link between complementarities and superstar dynamics is at least curious. Following Rosen (1981), greater *substitutability* is typically associated with "superstar" phenomena, as it implies, for given inputs, extra weight on the highest-quality input.[35] As discussed at the end of Section 2.1, the opposite occurs with stronger complementarities. When input choice, that is in our case workers, are is treated as endogenous, the consequence is that high-quality workers are more likely to matched together.

2.3 Measuring coworker complementarities in theory

We now have a qualitative understanding of the macro-distributional implications of coworker complementarities and a conjecture that they ought to have risen over time alongside a rise in task complexity. This section provides the methodological ground for taking the theory to the data and, eventually, toward quantitative application: a theory-guided and empirically feasible approach to measuring complementarities using standard matched employer-employee panel data.

A preliminary step is to identify workers' quality types. Provided that the wage function is monotonically increasing in the worker type, x, and if there is sufficient labor market mobility, this can be accomplished using data on wages and matches. In addition to widely used reduced-form approaches à la Abowd *et al.* (1999), a theory-consistent approach is the non-parametric ranking method proposed by Hagedorn *et al.* (2017). Appendix B.2.4 proves that the wage function $w(x,(y,x'))$ is indeed monotonically increasing in x and, hence, Hagedorn *et al.*'s (2017) approach extends to the present, theoretical environment.

If each worker's quality type is known, how can we measure the degree of coworker complementarity in production, $\frac{\partial^2 f(x,y,x')}{\partial x \partial x'}$, which guides coworker sorting? Except in circumscribed contexts (Mas and Moretti, 2009; Adhvaryu *et al.*, 2020) we do not have measures of individual *output* while establishment or firm output is buffeted by many variables other than worker quality.

Under our surplus sharing rule (27), a worker appropriates a constant fraction (corresponding to their bargaining power) of the marginal surplus generated by matching with the employer and its existing employees. After some algebra (detailed in Appendix B.2.3), we can write the implied wage

[34] A similar non-linearity is also operative in the standard model of workers and firms matching into pairs. To the best of my knowledge, this insight is novel.

[35] This point occasionally prompts the comment that in Kremer (1993), greater complexity is not only associated with stronger complementarities but also a more right-skewed distribution of income for a given, symmetric distribution of qualities. Note, however, that this result hinges on the assumption that the production function also features increasing returns to the quality of the workforce as a whole; it does not hold under decreasing or, as in this model, constant returns.

of x employed at y with x' as

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

Intuitively, the wage pays the worker an ω-share of the increase in production from x being added to the (y,x') coalition; plus a fraction of x's outside option; *minus* compensation of (y,x') for their foregone share of a surplus from potentially hiring some other worker.

The key insight is that to identify complementarities, we do not need to identify the full production function in levels. It is sufficient to identify the cross-partial derivative (i.e., partial identification). Differentiating twice with respect to x and x' respectively yields the following insight.

Corollary 3 (Measuring coworker complementarities in theory)**.** *The strength of coworker complementarities in production is proportional to the strength of coworker complementarities in wages:*

$$\frac{\partial^2 f(x,y,x')}{\partial x \partial x'} = \frac{1}{\omega}\frac{\partial^2 w(x,(y,x'))}{\partial x \partial x'}. \tag{34}$$

The intuition is straightforward. In a situation of bilateral monopoly (due to the presence of search frictions), a fraction of the surplus accrues to the worker. The curvature of this surplus is shaped by the production function. The level of the wage is also influenced by the bargaining partners' respective outside options. Holding constant these outside options, if the surplus increases by some increment Δ, then the worker appropriates a fraction ω of that Δ-increase. Hence, wage changes are informative about changes in production. Viewed differently, when considering the cross-partial derivative, the outside options drop out. That of the employee is foregone irrespective of whether x is matched with a bad or a good coworker. That of the pair (y,x') to which x is added is foregone irrespective of whether they hire x or a different alternative. Hence, building on existing work in the literature on identifying the worker bargaining parameter ω, we can use measurements of $\frac{\partial^2 w(x,(y,x'))}{\partial x \partial x'}$ to quantitatively discipline $\frac{\partial^2 f(x,y,x')}{\partial x \partial x'}$.

3 Empirical evidence

Taking stock, we now have all the theoretical tools needed to evaluate the macroeconomic implications of coworker production complementarities. The conjecture is that structural shifts in the nature of work have strengthened such complementarities (Corollary 1 and Figure 1a). Through the lens of the model, the macroeconomic implications include greater coworker sorting and widening gaps in average wages paid across firms (Corollary 2). Appendix A.4 documents – in line with the existing literature – that these macro-distributional trends are also observed in the data.

In this section, I first discuss empirical evidence on changes in the nature of work and how they map onto the key model parameter χ (Section 3.1). To provide a quantitative indication of how complementarities have indeed evolved over time, I then implement the measurement approach implied by Corollary 3 using German micro data (Section 3.2). The key finding is that the magnitude of wage complementarities has increased about twofold in the two decades since 1990. In Section 4, these micro-moments impose discpline when estimating the model and performing quantitative analyses. Before that, I exploit cross-sectional variation to subject the predictions of the model structure to further empirical scrutiny (Section 3.3).

3.1 Empirical proxy for χ: task complexity

The key technology parameter of interest in the model is χ, which controls the degree to which worker-task specific efficiencies are dispersed.[36] Formally, χ thus describes a property of the joint distribution of worker-task-specific efficiencies. This object is not directly measurable in the data and could, as noted earlier, move with changes in either task requirements or workers' skill acquisition. In this paper I do not take a firm stance on the relative importance of or aim to disambiguate between these two sources. In practice, as a proxy for χ I use the extent to which different tasks are abstract and non-routine, or "complex" for short, meaning that they require cognitive skills and cannot be performed by following explicitly programmed rules (cf. Autor *et al.*, 2003). I next discuss how empirical evidence, existing literature and intuition all point to the task requirements of production, thus defined, as a plausible proxy.

An important advantage of the task complexity proxy is that we have both time-series and cross-sectional evidence on its evolution. Thus, as noted in the introduction, the average complexity of workplace task reported by respondents in a long-running, large-scale survey in Germany has indeed consistently increased since the mid-1980s. At the same time, we can use cross-sectional variation in task complexity for comparison with alternative proxies of χ. In the first instance, Online Appendix D reports that there is extensive variation in the complexity of tasks performed by different occupations. Managerial and high-skill professional groups tend to score highly, while manual and routine service jobs rank on the opposite end of the spectrum. Reassuringly, this task-complexity based ranking resembles an occupation-level measure of within-firm task differentiation across positions, constructed by Bloesch *et al.* (2022) using Burning Glass Technologies US vacancy

[36] An intuitive label of χ is "task specialization," but note that this concept is distinct from the extent to which the worker-task assignment exhibits division of labor. In the model presented here, though, according to the optimal assignment the measure of tasks performed by more than worker is *always* equal to zero, and χ merely pins down the gains from this division of labor relative to "autarky." In Freund (2022), I present a more general model in which, indeed, division of labor can be incomplete due to coordination costs and, other things equal, a higher value of χ induces the firm to allow for less task overlap across workers.

data. They report that managerial and professional jobs are the most differentiated, followed by technicians; service and manual jobs have the lowest differentiation scores.[37]

Intuitively, there are several reasons why task complexity should positively correlate with χ. Heterogeneity in task-specific productivities, or comparative advantage, can be "natural" or "acquired" (Grossman and Helpman, 1990) but in either scenario variation is more likely when tasks are complex. In the former case, as Martellini and Menzio (2021, 340) state, matter-of-factly: everyone is likely to be similarly productive in routine jobs "almost by definition", whereas non-routine occupations exhibit more differentiation (also see Mollick (2011)). If we view task-specific productivities as acquired, dispersion in task-specific productivities naturally emerges when there are large fixed costs of knowledge acquisition and, hence, increasing returns to specialization (Rosen, 1983; Alon, 2018) or when there is significant scope for learning-by-doing.[38] In this spirit, Caplin *et al.* (2022) find that the time needed to reach maximal productivity is highest for management occupations or those that "require high knowledge" or management occupations.

Finally, and more loosely, an increase in χ, which per Proposition 1 goes along with greater benefits from team production, is also consistent with the rising importance of teamwork in the economy. This rise of teams has been extensively documented for scientific production (Jones, 2021; Pearce, 2022). More broadly, complementing the transformation of the task content of production, firms have reallocated skilled workers into flexible, team-based settings (e.g., Bresnahan *et al.*, 2002; Bloom and Reenen, 2011).[39] One technological factor is automation leading to the replacement of humans in routine tasks, leaving them to handle the more complex problems (Acemoglu and Restrepo, 2018). But, indeed, the notion that there are secular forces putting upward pressure on specialization is a long-standing one (Smith, 1776; Neffke, 2019). Thus, Jones (2021) connects Einstein's notion of "inevitable specialization" to the rise of teams: "[For] the greater the stock of knowledge in an area, the narrower the expertise of the individual investigator becomes, and the greater the role of teamwork in attacking broad problems."[40]

[37]Consistent with the idea that the combination of position-specific skills and production complementarities generates hold-up power for individual workers, Bloesch *et al.* (2022) find that managers exhibit both the highest hold-up power, while low-wage service, manual, and administrative jobs have the smallest measure.

[38]Consider two equally talented high-school students. It takes them similar amounts of time to solve a given mathematical problem. Yet, after they each spent five years at university acquiring the skills demanded of an electrical or a mechanical engineer, one would find it very difficult to perform a task considered easy by the other, and vice versa. Both types of engineering would be required in, say, the design of a modern air plane.

[39]Empirically, the rise of team production has been studied most carefully in the context of scientific production (Wuchty *et al.*, 2007). This partly reflects the powerful incentives for teamwork in a context where the cumulativeness of knowledge pushes individuals toward specialization (Jones, 2009). In addition, though, this empirical context simply permits more straightforward measurement and teamwork is arguably of much broader relevance (Boning *et al.*, 2007; Weidmann and Deming, 2021). Survey evidence documenting a secular rise in the role of teamwork is available, for instance, in the US (Lazear and Shaw, 2007) and the UK (Wood and Bryson, 2009).

[40]This idea is by no means limited to the economics literature. Building on the seminal sociological work of Thompson (1967), Raveendran *et al.* (2020) write: "[As] task complexity and ambiguity increase within organizations, self-selection of roles, expertise, and job duties has become increasingly prevalent, making businesses and firms another area in

3.2 The evolution of coworker wage complementarities in Germany

This section uses information on matches and wages from the longitudinal version of the German Linked-Employer-Employee dataset (LIAB) to measure the evolution of coworker wage complementarities since the mid-1980s.

3.2.1 LIAB Data

The LIAB contains the complete employment biographies of individuals employed at a representative panel of establishments repeatedly interviewed in the IAB Establishment Panel. Relevant variables include a worker's establishment and average daily wage alongside a rich set of other characteristics, including employment status, age, gender, tenure, occupation, and education, among others. Individual records originate in labour administration and social security data processing. On account of these advantages, the LIAB is a widely used micro-dataset (including for the study of coworker learning, see Jarosch *et al.* (2021)).

For the analysis, I convert the data into an annual panel, deflate wages to be measured in 2015 terms, and follow literature standards in imputing top-coded observation (Gartner, 2005; Dustmann *et al.*, 2009; Card *et al.*, 2013). I restrict the sample in two steps. First, I select full-time employed individuals at establishments in West Germany, aged 20-60, with real wages of at least 10 Euros. Then I restrict attention to the largest connected set, as we will rely on worker mobility for identification, and subsequently drop observations in establishment-year cells containing fewer than ten worker observations. Appendix A.1 provides further information on how I process the LIAB as well as summary statistics.

3.2.2 Mapping theory and data

The mapping between the parsimonious theoretical model and the complexity of real-world data warrants further explanation. The theoretical production unit consists of only the employer and (at most) two workers, whom I labelled a "team." In the data, I take a worker's coworker group to consist of all other individuals in the same establishment-year cell. In the LIAB, establishments are distinguished by industry, location and ownership. As such, the employer unit is narrower than a firm, as employees across different establishments of the same firm could be geographically separated, for instance.[41] Still, the median team size is around 20 – which is rather larger than two. The theoretical team size restriction was imposed on grounds of tractability rather than realism. As argued also in Herkenhoff *et al.* (2022), though, in empirical practice we can collapse all coworkers

which the process [of intensifying specialization] unfolds on a regular basis."

[41] For ease of language, in the remainder of the text I use the terms "establishment" and "firm" interchangeably.

into a “representative coworker” and, thus, create a mapping between data and theory.[42]

Second, since the model neither incorporates wage determinants such as life-cycle dynamics or tenure effects nor features aggregate productivity dynamics, we require a measure of residual wages that nets out these factors. To obtain this measure, I regress the log real daily wage of individual i in year t, $\tilde{w}_{it}$, on an individual fixed effect and an index of time-varying characteristics, X_{it}, which includes year dummies, a cubic in age and a quadratic in job tenure.[43] As an input into the reduced-form analysis, as well as in the estimation, I then use $w_{it} = \exp(\ln(\tilde{w}_{it}) - X_{it}'\hat{\beta})$. The dispersion in these residual wages accounts for approximately 65% of the variance in raw wages in 2010.[44]

3.2.3 Implementing the measurement result

The first step toward a measure of coworker wage complementarities is to estimate each worker's time-invariant quality type. I implement two alternative approaches that have been proposed in the relevant literature. One, I estimate two-way fixed effect wage regressions in the spirit of Abowd *et al.* (1999, AKM) – using dimensionality reduction techniques to mitigate a well-known incidental parameter bias problem (Andrews *et al.*, 2008; Bonhomme *et al.*, 2019) – and rank worker's by their worker fixed effect, i.e., I create an ordinal measure that corresponds to the quality x in the theoretical model. Two, and as foreshadowed in Section 2.3, I implement a version of the non-parametric ranking algorithm proposed in Hagedorn *et al.* (2017). Details on both approaches are relegated to Appendix A.3. The correlation between the resulting two alternative rankings is 0.7. I use the AKM-based measure as a baseline and report robustness checks in Section A.4.5.1.

Given such a ranking I compute worker and coworker quality types. To align the empirical analysis with the structural model, which I solve after discretizing the worker type space with ten grid points, I bin the workers according to their decile rank. Worker i's time invariant type (decile rank) is denoted $\hat{x}_i$. The binning procedure helps reduce noise.[45] Next, given worker ranking and

[42] I have experimented with an alternative definition of coworkers, comprising those workers who in the same establishment and also the same two-digit occupation (cf. “Team Definition 2” in Jarosch *et al.* (2021)). Some results for this alternative definition are shown in Section 3.3, but I prefer the baseline approach. The team production model tells us precisely that coworker complementarities will exist when workers are *differentiated* in their task-specific productivities, which is more likely to obtain across different occupations who nevertheless contribute to the same final output.

[43] As the regression includes worker 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), I normalize age to be flat at 40. I intentionally do not include establishment fixed effects here. The reason is that the log-linear model with employer fixed effects attributes any correlations in output between workers at the same employer, insofar as they are reflected in wages, to a common employer component, even if these correlations are due to coworker complementarities.

[44] See Appendix A.4.3 for variance decompositions using several, alternative measures of residual wages.

[45] To be precise, I compute an individual's decile rank in the year-specific distribution, since the number of workers in the sample varies across years (different from the model) and I want to ensure that the type distribution is uniform in each year. Having said that, the binning procedure also implies that a worker's *decile* rank generally does not change across years. I omit a time subscript to clearly signify that $\hat{x}_i$ denotes a characteristic that is time-invariant modulo changes in the sample composition.

information on employment relationships I compute, for each person-year, the average quality of coworkers. Define the set of i's coworkers in year t as $S_{-it} = \{k : j(kt) = j(it), k \neq i\}$, where $j(it)$ is the identifier of i's employer in period t. Then $\hat{x}_{-it} = \frac{1}{|S-it|} \sum_{k \in S_{-it}} \hat{x}_k$ is the average type of i's coworkers in year t.

With residual wages, worker and coworker quality measures at hand, we can quantify the cross-partial derivative $\frac{\partial^2 w(x,(y,x'))}{\partial x \partial x'}$, i.e., how the slope of the wage function with respect to the (representative) coworker type varies with the own type. There are two different routes to compute this object. The first option is to construct a non-parametric wage function and numerically compute the cross-partial derivative, for instance using finite-difference methods (Appendix C.2 elaborates). The attractiveness of this approach lies in its close alignment with the theoretical model; I will rely on it to inform the model estimation.

We might worry, however, about additional threats to identification that are salient in the data even as the theoretical model abstracts from them, including time-varying shocks at the industry- or occupation level that correlate with wages. An alternative approach that addresses such concerns but imposes more parametric structure is to estimate the following reduced-form wage regression:

$$w_{it} = \beta_0 + \beta_1 \hat{x}_i + \beta_2 \hat{x}_{-it} + \beta_3 (\hat{x}_i \times \hat{x}_{-it}) + \psi_{j(it)} + \nu_{o(i)t} + \xi_{s(i)t} + \epsilon_{it}, \tag{35}$$

where $\psi_{j(it)}$ denotes employer fixed effects (FE), $\nu_{o(i)t}$ are occupation-year FEs, $\xi_{s(i)t}$ are industry-year FEs.

The coefficient of interest is β_3. The question we ask with this regression is the following. Conditional on your own type and your employer's type, having a better coworker might benefit your productivity and, hence, wage. How does the magnitude of this benefit vary with your own type? If the true wage process follows equation (35), then β_3 measures the coworker wage complementarities of interest; for differentiating w_{it} with respect to $\hat{x}_i$ and $\hat{x}_{-it}$ yields β_3. In terms of interpretation, and since I treat $\hat{x}$ as continuous in the estimation, β_3 indicates how much more the real wage of an individual i rises with a one-decile increase in quality of coworkers compared to an individual i' whose rank is one decile lower than that of i.

The regression specification leans on the peer effects literature and existing analyses of coworker wage effects.[46] As is common, identifying variation comes from both movers, i.e., changes in coworker quality for individuals who switch employer, and from stayers, i.e., changes in coworker quality induced by other employees joining or leaving the coworker group. To address the reflection problem (Manski, 1993) I use a pre-determined, model-consistent measure of coworker quality and seek to identify so-called exogenous peer effects, as opposed to simultaneously estimating

[46] See, in particular, Arcidiacono *et al.* (2012); Cornelissen *et al.* (2017); Barth *et al.* (2018); Cardoso *et al.* (2018); Hong and Lattanzio (2022).

types and peer effects.[47] Controlling for the worker's own type accounts for the potential selection of high types into coworker groups with a high average type; such coworker sorting is, of course, precisely what we are hypothesizing here. Finally, I control for unobserved time-invariant employer heterogeneity as well as shocks at the occupation-year or industry-year level using a rich set of fixed effects.

Guided by the structural model, the analysis here distinctively zeroes in on *coworker wage complementarities*, as opposed to *coworker wage effects* in a broader sense, which the literature usually seeks to quantify. Through the lens of the structural model, if having a better coworker confers vast wage (and, ultimately, production) benefits to a worker, but the absolute increase is the same regardless of the beneficiary's own type, then this provides no incentive for coworker sorting. What matters is, instead, whether the benefit from greater coworker quality is differentially larger for high types. Meanwhile, a common specification along the lines of

$$\ln(w_{it}) = \zeta_0 + \zeta_1 \hat{x}_i + \zeta_2 \hat{x}_{-it} + \text{controls} + \text{error},$$

where the dependent variable is in logs, *imposes* a constant (semi)-elasticity, so that the absolute gain is greater for high types. Through the lens of our structural model, the log-specification would be misspecified. Put differently, the specification in equation (35) marries lessons from the peer effects literature regarding threats to identification with the emphasis of the structural literature on worker-firm matching that it is important to take account of non-monotonicities in payoffs (e.g., Eeckhout and Kircher, 2011).

3.2.4 Result

Figure 4 plots, in orange squares, the point estimates for the interaction coefficient of interest, $\hat{\beta}_3$, alongside 95% confidence intervals, after ranking workers and then estimating regression (35) separately for 5 sample periods. The key takeaway is that the magnitude of β_3 has increased over time, approximately doubling from the first to the final sample period. Appendix A.4.5.1 reports that $\hat{\beta}_3$ evolves very similarly when worker types are estimated non-parametrically.

The same figure overlays, in black circles, the task complexity measure. The evident co-movement of this proxy of χ and the estimated coworker complementarities coheres well with the predictions of the model. To provide a sense of magnitudes, if β_3 is equal to 0.1, this means that the real hourly wage increase from a one decile improvement in the average coworker quality is 0.5 euros greater for a worker who is themselves in the top decile compared to a worker in the fifth decile.

Complementing these regression-based results, Table 1 also reports the average cross-partial deriva-

[47] In that respect, my approach is similar to Nix's (2020) use of education as indicator of coworker quality. The present approach also allows for coworker effects through unobservables.

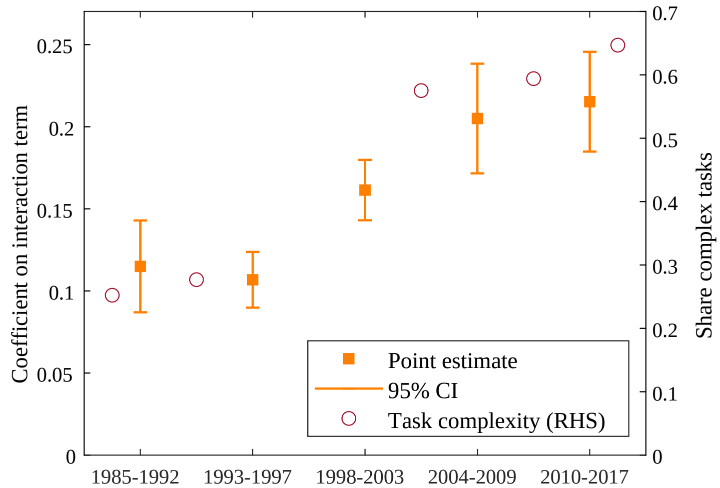


Figure 4: Empirical estimates of coworker wage complementarities across sample periods

Notes. This figure reports the point estimate for the coefficient β_3, alongside confidence intervals. Workers are ranked and then regression (35) is estimated on five subsamples of the data. Standard errors are clustered at the employer-level. Magnitudes are reported in terms of hourly wages (using a scaling factor of 7.5 working hours per day), to aid comparison with results shown for Portugal in the following section. The share of complex tasks in workers' activities, depicted in circles, is computed as described in the notes for Figure 1a. 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.

tive computed non-parametrically (see Appendix C.2 for the methodology). The estimates are similar to those obtained from the regression approach, only smaller in magnitude, and have likewise increased substantially over time. Jointly, these analyses establish the following result.

Empirical result 1. *The magnitude of coworker wage complementarities has increased over time.*

Sample period	$\hat{\beta}_3$	$\widehat{\frac{\partial w^2(x,(y,x'))}{\partial x \partial x'}}$
1985-1992	0.115	0.046
1993-1997	0.107	0.050
1998-2003	0.161	0.095
2004-2009	0.205	0.104
2010-2017	0.215	0.157

Table 1: Empirical estimates of coworker complementarities: comparison of methods

Notes. This table reports, in the first main column, the same point estimate for the coefficient β_3, also described in Figure 4; and in second main column, the non-parametric approximation of the cross-partial derivative of the wage function with respect to own type and average coworker type. Magnitudes are reported in terms of hourly wages (using a scaling factor of 7.5 working hours per day).

3.3 Validating the model mechanisms

The estimates of coworker complementarities will serve as key moments to discipline the quantitative model. Yet, any counterfactual results generated by that model will also hinge on its remaining elements. Hence, if the structural model is to be relied upon, it would be reassuring if the following two, model-implied relationships also held in the cross-section:

(1) Coworker complementarities are stronger when tasks are more complex.[48]

(2) Greater coworker complementarities are associated with more positive assortative matching among coworkers.

More implicitly, the model is premised on the following, third relationship, for instance by abstracting from product market frictions and questions of market power.

(3) Differences across firms in the quality of their workforce are a key source of between-firm inequality in productivity and pay.

The second goal of this Section 3 is to subject these three relationships to empirical scrutiny. Exploiting cross-sectional variation across occupations and industries each one of them finds support in the data, as well as in the existing literature.

3.3.1 Portuguese micro-data

To inspect these relationships in the data, I expand the analysis by supplementing the LIAB data with a second set of micro data, namely the Portuguese Quadros de Pessoal/Relatório Único (QP, henceforth) matched employer-employee panel dataset. All results reported for Portugal are drawn from joint work with Criscuolo and Gal (Criscuolo *et al.*, 2023).[49] The motivation to undertake this effort is that the LIAB data posses several undesirable properties that limit further probing of the predictions of the model. In particular, a non-negligible fraction of wage observations are imputed due to top coding; information on all employees is available only for a subset of establishments; the sample is large but does not cover the full population; and while information on establishments' productivity can be constructed, information on sales and non-labor inputs are based on surveys and, thus, measurement error may be severe.[50]

[48]This premise is not necessary for the argument that coworker complementarities are behind rising between-firm inequality. But if the relationship between tasks and complementarities predicted by the production function microfoundations put forth does obtain, then this lends further credibility to the argument at hand, as well as offering a natural interpretation of the structural forces at work.

[49]I thank Chiara Criscuolo and Peter Gal, and the OECD's Directorate for Science, Technology, and Innovation (STI) more broadly, for facilitating these analyses and the usage of the results in this paper.

[50]In Appendix A.4 I show that in Portugal, too, measures of worker segregation have increased over time. Different from Germany, the between-firm share of wage inequality is flat, however, this is because falling firm-specific pay premia compensate the rise in sorting and segregation. Moreover, coworker wage complementarities also have also increased over time, just as we found in the LIAB.

The QP is an annual mandatory census of all employers in Portugal. It provides detailed information on workers' employment status, hourly wage (incl. bonus payments), gender, education, age, tenure and ISCO-08 2-digit occupation. By merging in balance sheet and income statement information from the Informação Empresarial Simplificad, we can construct a measure of value-added per worker to study dispersion in firm-level productivity. These data are unusually rich; they are also of high quality (cf. Card *et al.*, 2016). Importantly for our purposes, the QP data comprises the universe of Portuguese employers and employees, wages are not top-coded, and there is little measurement error.[51] Appendix A.2 provides a more detailed description, including of the sample selection criteria, which resemble those applied to the LIAB.

3.3.2 Main findings

Using the 2010-2017 subsample of these data, I conduct three sets of analyses to further probe the model-implied connections between tasks and complementarities, complementarities and sorting, sorting and between-firm inequality. While distinctively reduced-form and correlational, each analysis exploits a different margin of cross-sectional variation: across occupations, hierarchical layers and industries.[52] I proceed by successively summarizing each main result and then elaborating on the underlying analysis as well as related literature.

Empirical result 2 (Occupations). *Occupations performing more non-routine abstract tasks exhibit stronger coworker wage complementarities. In turn, coworker sorting is more pronounced in occupations exhibiting stronger coworker wage complementarities.*

A rich literature in labor economics conceives of different occupations as bundling different types of tasks. Using a (cardinal or ordinal) ranking of the task content of different occupations to facilitate interpretation, we can then study variation across different occupations in the strength of coworker complementarities as well as coworker sorting.

First, if we take an occupation's reliance on non-routine, abstract (NRA) tasks to be a proxy for χ (as in Deming (2017)), Proposition 1 leads us to expect a higher NRA score to predict stronger complementarities. To this end, I use task indices constructed by Mihaylov and Tijdens (2019) to measure the degree to which each 2-digit occupation involves non-routine abstract (NRA) tasks.[53]

[51] Why not use Portugal as the primary datasource, instead of Germany? One reason is that I want to connect my analyses to previous results in the literature, which have often treated Germany, being the largest European economy, as a reference case. Portugal, on the other hand, has experienced several quite idiosyncratic macroeconomic shocks and developments over the past three decades. I accordingly focus on cross-sectional analyses. Much more practically, the time frame for the analysis of the Portuguese data was limited.

[52] Throughout, as a measure of a worker's type I use their rank in the distribution of worker fixed effects as estimated from an AKM model. I have also verified that repeating the analysis using the subsample containing only one gender yields very similar results.

[53] The index by Mihaylov and Tijdens (2019) is attractive because it is constructed from occupation-specific lists of tasks provided by ISCO-08. Therefore, no selection among many potential task scales is involved, which can be

Thanks to the comprehensive coverage of QP, we can run a version of regression (35) for each occupation. Figure 5 plots the estimated, occupation-specific coefficient on the interaction term between own quality and coworker quality (vertical axis) against the occupation-level NRA score (horizontal axis). We observe a positive relationship. The estimated interaction term is close to zero for those occupations with the lowest NRA score, whereas it is around 0.2 when tasks are primarily non-routine abstract.

Theoretical result 2 furthermore predicts that worker in occupations featuring greater coworker wage complementarities are also matched in a more positively-assortative pattern. That is indeed exactly what we observe when plotting that relationship in the second panel of Figure 5. In sum, variation across occupations thus lends support to relationships (1) and (2).

These findings are consistent with existing results reported in the literature (while zooming in on the theory-relevant measure of coworker complementarities). Thus, Neffke (2019) studies detailed Swedish micro data and suggest that coworker effects are important in many knowledge-intensive jobs (e.g. health care, engineering) and professional occupations (e.g., lawyers); and in skill-intensive (e.g. R&D) an crafts-based industries (e.g. construction). Jäger and Heining (2022) find that when a high-skilled or specialized worker dies, their coworkers in other occupations experience wage decreases.

Empirical result 3 (Hierarchical layers). *Coworker wage complementarities are (weakly) monotonically increasing in the layer of a firm's internal hierarchy.*

The richness of the Portuguese data facilitates an alternative operationalization of who an employees' coworkers are: those who work at the same hierarchical level of a firm. From 2010 onward, we can use a variable that assigns workers according to a consistent definition into seven different, vertifically differentiated layers, grouped by similarity in the complexity of tasks and skills required. These range from top executives to non-skilled.[54]

This operationalization is appealing on at least two counts. First, the definition of coworkers aligns well with the model in Section 2.1, wherein team members do not manage each other but they are potentially differentiated in terms of their expertise in different tasks. Second, layers are naturally hierarchical, with higher layers performing more complex tasks (a mapping that emerges naturally in Garicano (2000)-style models of knowledge-based hierarchies). If we interpret these hierarchical layers as an ordinal proxy for χ, we would expect the cross-partial derivative

controversial, as discussed by Acemoglu and Autor (2011). Moreover, ISCO-08 is also applied in Portugal over the sample period. I aggregate the measures to the 2-digit level.

[54] Per Decree-Law 380/80, firms should indicate for each employee the qualification level indicated in the relevant Collective Agreement. If this is not available, firms should select the qualification level of the worker. Table B-1 in Mion and Opromolla (2014) provides details on tasks and skills. Also see Caliendo *et al.* (2020). I exclude apprentices/intern-s/trainees.

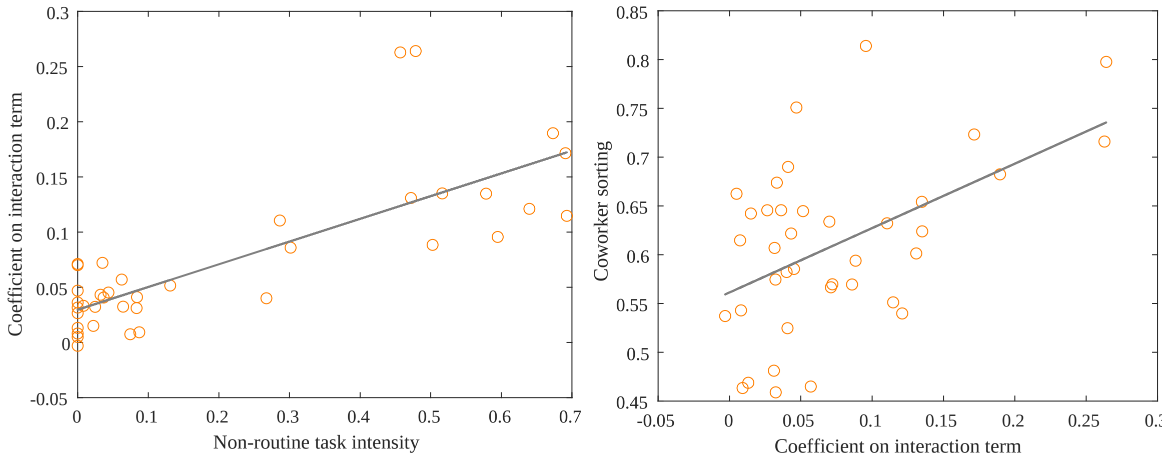


Figure 5: Complementarities and sorting across different occupations

Notes. The left panel plots for each ISCO-08 2-digit occupation a measure of coworker wage complementarity against the non-routine task intensity of that occupation. the former measure is the estimated coefficient on the interaction term β_3 when estimating a version of equation 35 separately for each occupation. The right panel plots the occupation-level coworker correlation coefficient, i.e., the correlation between a worker's fixed effect and the average fixed effect of their coworkers against the strength of coworker wage complementarity. Throughout, to avoid overlap across occupations, the set of employee i's coworkers here includes in all those employees in the same firm-year-occupation cell. Moreover, the data point for CEOs (ISCO-08-2d code 11) is omitted from the plots. While the associated observations align with the pattern shown but the magnitudes are substantially larger and thus make the visualization less clear for the remaining occupations.

To test this hypothesis, I therefore run the following variation of the regression specification in (35), for each $l = 1, \ldots, 7$.[55] Figure 6 shows the estimated value of β_3 alongside 95% confidence bands for each layer. The point estimates are almost monotonically increasing in the hierarchical layer, ranging from effectively zero for those classified as non-skilled to above 0.2 for top executives. This finding thus lends further support to relationship (1).[56]

Empirical result 4 (Industries). *Industries with production processes featuring relatively high skill complementarity also tend to be characterized by greater coworker sorting and between-firm inequality.*

To classify industries, I reconstruct an industry-level proxy for complementarity proposed by Bombardini *et al.* (2012), using O*NET data on occupation and industry-level occupational employment weights from QP.[57] The basic idea behind Bombardini *et al.*'s (2012) approach is that an industry is more likely to feature skill complementarity if many workers are in occupations for which the

[55] I deliberately estimate the regression separately for each layer, as opposed to pooling samples and including a double interaction between the interaction term and l, to avoid restricting potential varriation in β_1 and β_2 across l. The downside is that we do not exploit variation from switchers across layers within firms, as a result.

[56] Appendix Figure A.6 shows that a similar pattern can be observed in Germany across broad occupation groups.

[57] Clearly, we could characterize industries also on the basis of occupational task indices, as considered before. I adopt Bombardini *et al.*'s (2012) approach as a baseline in order to test the theory against a variety of proxies for χ, including ones used in the extant literature for related purposes.

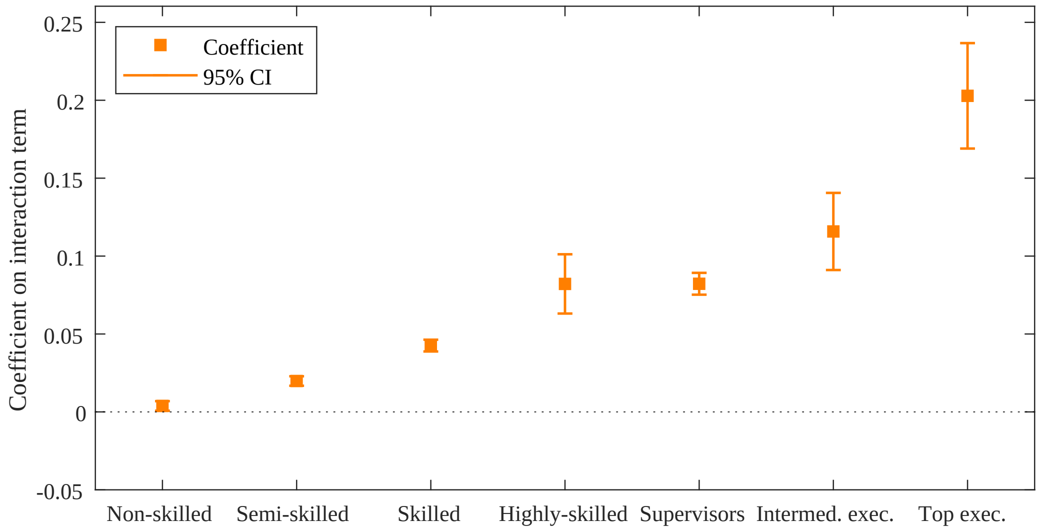


Figure 6: Complementarities across different hierarchical layers

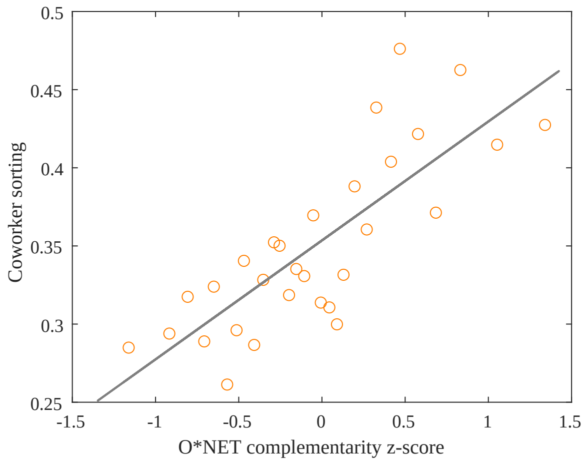
Notes. This figure reports the point estimate for the coefficient β_3, alongside confidence intervals, when regression (35) is estimated separately for each layer, as described in the text.

following four characteristics are rated important: teamwork, impact on coworker output, communication, and contact.[58] Figure 7 shows that the higher an industry scores according to the proxy measure, the greater tends to be the degree of coworker sorting as well as the dispersion of productivity and wages across firms.

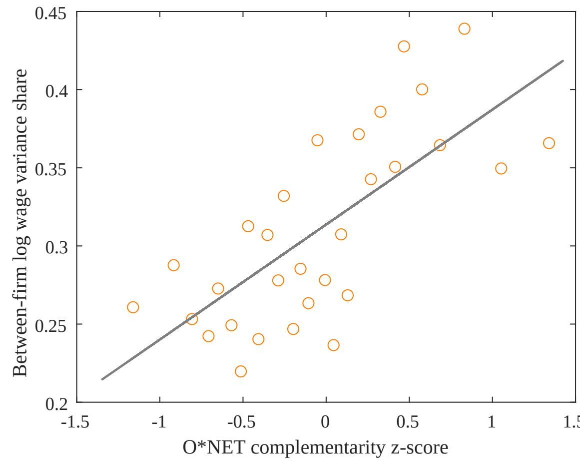
This finding provides further reassurance that prediction (2) of the model is consistent with correlations in the data. Industry variation is also highly consistent with premise (3), i.e., the importance of workforce composition. Thus, Appendix Figure A.7 reports that across 4-digit industries, measures of between-firm inequality in productivity and wages are increasing in the degree of coworker sorting. Indeed, the relationship between coworker sorting and the between-firm wage of the variance of log wages is almost linear with a slope equal to one. That is, the industries with the highest measure of coworker sorting (around 0.55) have, on average, a between-firm share of the variance of log wages equal to 0.55; industries featuring a coworker sorting coefficient around 0.15 have a between-share of close to 0.15.

Again, these results are in line with a literature that has underscored the important role of differences in workforce composition across firms in explaining between-firm wage inequality. In addition to the aforementioned Card *et al.* (2016) and Song *et al.* (2019), Håkanson *et al.* (2021) use direct measures of workers skills from military enlistment tests in Sweden and find that between-firm

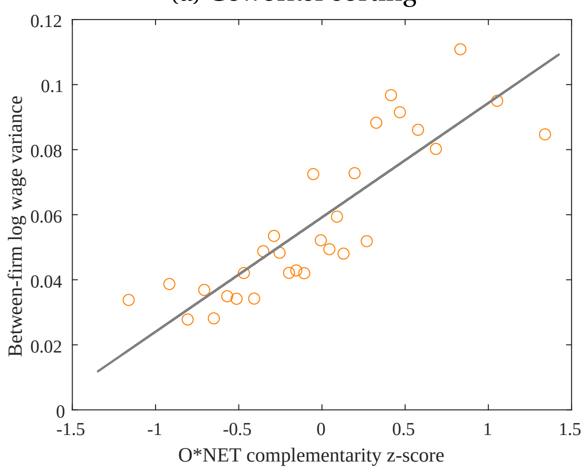
[58]That is, I first construct the occupation-level "CTIC" meaures based on the average score the O*NET experts assigned, on a Likert scale from one to five, and taking the mean across the four different dimensions. Then, after cross-walking to the ISCO-2d classification available in the QP data (using publicly available crosswalks), construct an employment-weighted mean score for each 4-digit. The reader is referred to Bombardini *et al.* (2012) for details on the questions selected from O*NET.



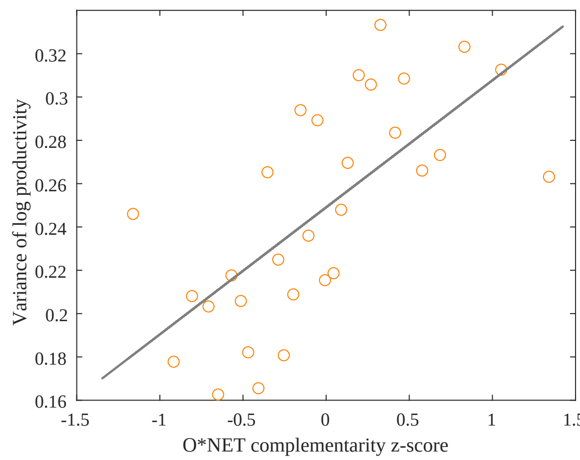
(a) Coworker sorting



(b) Between-share of wage inequality



(c) Between-firm wage inequality



(d) Productivity dispersion

Figure 7: Binscatter graph of sorting and between-firm inequality across industries

Notes. Each panel plots, at the NACE-4-digit industry level, the moment indicated in the respective plot against the industry-level proxy for complementarity, which is constructed per description in the main text. Observations are grouped into 30 bins. To limit the influence of outliers, I drop the bottom and top 5% industries in terms of each moment of interest. The linear regression line is fitted based on unweighted observations.

skill inequality is a key driver behind between-firm wage inequality. Sorkin and Wallskog (2021) find, meanwhile, for the U.S., that productivity dispersion, between-firm earnings inequality, and coworker sorting are successively higher for more recent cohorts of firms in the U.S (also see Berlingieri *et al.* (2017)).

4 Structural estimation

Informed by theoretical and empirical results, I next quantitatively evaluate the macroeconomic implications of coworker complementarities. I start by calibrating the model to match micro- and macro-moments characterizing the German economy in 2010. The estimates of coworker wage complementarities at the micro-level, summarized in Table 1, together with Corollary 3, serve to discipline the elasticity of complementarity, γ, and thus indirectly the task complexity parameter χ. This section first outlines the methodology, then discusses the estimation results. I conclude by re-estimating key parameters, including γ, for 1990.

4.1 Baseline calibration: 2010

4.1.1 Methodology

DISTRIBUTIONAL AND FUNCTIONAL FORM ASSUMPTIONS. The estimated version of the model abstracts from ex-ante heterogeneity in firm types (i.e., the distribution of y is degenerate, concentrated on $y = 1$). The objective is to narrowly focus on the dynamics of coworker complementarities and examine to what extent they are capable of endogenously giving rise to between-firm inequality without *assuming* permanent differences ex-ante. Throughout, the model is solved after discretizing worker types, x, into ten equally spaced grid points.

Next, the functional form of the production function(s) is directly informed by the microfoundations set out in Section 2.1. For quantitative purposes, I introduce three generalizations. First, I introduce a constant factor, a_0, such that even the lowest-ranked employee produces a strictly positive amount of output.[59] Thus, a firm with a single worker x produces $f(x) = a_0 + a_1 x$. Second, and turning to multi-worker firms, while Proposition 1 implies an upper bound for the elasticity of complementarity, $\overline{\gamma} = \frac{1}{2}$ (attained when $\chi = 1$), in the estimation I leave γ, unrestricted. Third, the degree to which labor productivity is greater when working in teams than working alone is

[59] In model, even the least productive worker, who according to equation (19) barely contributes to output, searches for employment. This is likely not true in the data, where those individuals would be out of the labor force. I interpret the intercept term a_0 as a stand-in for formally modelling this participation/selection margin.

controlled by a parameter a_2, so that

$$f(x, x') = 2\, \underbrace{a_2 \left[a_0 + a_1 \left(\frac{1}{2}(x)^{1-\gamma} + \frac{1}{2}(x')^{1-\gamma} \right)^{\frac{1}{1-\gamma}} \right]}_{\text{avg. labor productivity}}.$$

Lastly, the value of home production is proportional to the output produced alone, i.e., $b(x) = b_1 \times f(x)$.

EXTERNALLY SET, NORMALIZED OR ESTIMATED OFFLINE. In a first step, the following parameters are set outside the estimation. Regarding the parameters introduced in the preceding section, the flow value of home production is equal to 70% of the output produced without coworkers, similar to Herkenhoff *et al.* (2022)

Furthermore, as explained in Section 2.1, the production parameter a_2 has to be greater than unity for team production to be beneficial at all; the exact value turns out not to be crucial for present purposes, and I set $a_2 = 1.1$. This means that the average labor productivity in a team with two workers x and x' is 10% greater than that of a worker whose rank corresponded to the power mean of x and x', i.e., $\left(\frac{1}{2}(x)^{1-\gamma} + \frac{1}{2}(x')^{1-\gamma} \right)^{1/(1-\gamma)}$, if he or she produced without coworkers.[60]

Moving on to preferences, the discount rate ρ is set to 0.01. The implied annual interest rate of almost 13% is counterfactually high, as is common in this type of model; a high discount factor effectively proxies for concavity in the utility function, which is abstracted from for sake of tractability. Moving to labor markets, the exogenous separation rate, δ, fits a monthly separation rate of 0.01 (Hobijn and Sahin, 2009; Jung and Kuhn, 2014).[61] The bargaining parameter $\omega = 0.50$ follows standard values in the literature (see, e.g., Herkenhoff *et al.* (2022)).

INTERNALLY ESTIMATED. The remaining four parameters are estimated using indirect inference by matching moments. The estimated values of these parameters, summarized in the vector

[60] I have experimented with a specification that allows a_2 to endogenously vary with γ, as would be implied by the aggregation result in Proposition 1. The returns to scale in labor productivity term is decreasing in the team size and, so, the maximum team size of 2 required for tractability in the search environment could lead to an overstatement of this benefit, since one should really think of x in the one-worker unit as representing the average quality of the existing workforce. One plausible approach is to let $a_2 = \left(\frac{\bar{N}+1}{\bar{N}}\right)^{\frac{1}{1-\gamma}-1}$, where $\bar{N}$ denotes the average team size in the economy, so that a_2 measures the scale benefit *relative* to that. As a baseline, I opted to instead let a_2 be exogenous in order to cleanly isolate the effect of γ as a measure of complementarities. Using the above formula Setting $a_2 = 1.1$ is consistent with a median team size of around 10 and $\gamma = 0.5$.

[61] As documented by Jung and Kuhn (2014), average worker flows in and out of unemployment are substantially lower in Germany than in the U.S.

Parameter	Description	Value	Source
ρ	Discount rate	0.01	External
ω	Worker bargaining weight	0.50	External
b_1	Home production proportionality	0.7	External
a_2	Team benefit	1.1	External
δ	Separation rate	0.01	Offline estimation
γ	Elasticity of complementarity	0.638	Internal estimation
a_0	Production constant	3.85	Internal estimation
a_1	Skill sensitivity	20.78	Internal estimation
M_u	Meeting rate	0.304	Internal estimation

Table 2: Model parameters for 2010

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

$\psi = \{\gamma, a_0, a_1, \lambda_u\}$, minimize the following objective function (cf. De Ridder, 2021).

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

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

I next describe the moments used. Notice that ψ chiefly includes the production function parameters; all three will be informed by moments constructed using the same measure of residual wages used for the empirical analysis in section 3. Also included among the estimands is the job arrival rate, since the mapping between worker's meeting rate M_u and empirically observed job finding rates is mediated by the shape of agents' matching sets, which in turn are endogenous to the production function parameters. While the parameters in ψ are jointly estimated, each is closely informed by one of these moments, as explained next. Importantly, all moments relating to wages are constructed based on the subset of workers employed in teams.

For our purposes, the most important moment is a measure of wage complementarities, which following Result 3 directly informs the strength of *production* complementarities. To impose minimal linearity assumptions, I target the non-parametrically constructed average cross-partial derivative as indicated in Table 1, but compute them specifically for the year 2010.

Next, the values of a_0 and a_1 are guided by the average log wage and the total variance of log wages, respectively. Notice, in particular, that the dispersion of wages will be decreasing in a_0 and increasing in a_1. Finally, λ_u targets an unemployment rate of 5%.

To close this section, two final remarks are in place. First, following Bilal *et al.* (2022), Appendix

Parameter	Targeted moment	2010			1990		
		Value	m	$\hat{m}$	Value	m	$\hat{m}$
γ	$\widehat{\frac{\partial w^2(x,x')}{\partial x \partial x'}}$	0.638	0.109	0.109	0.383	0.049	0.049
a_0	Avg. log wage	3.85	2.66	2.66	7.20	2.73	2.73
a_1	Var. log wage	20.78	0.209	0.209	15.97	0.102	0.102
M_u	Unemp. rate	0.304	0.05	0.05	0.273	0.05	0.05

Table 3: Estimated parameters and targeted moments

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

C.1 conducts two exercises that jointly support the notion that ψ and each of its components are well-identified. Second, the estimation approach deliberately eschewed targeting the key macro-distributional moments of interest, that is, the degree of coworker sorting and the between-share of wage inequality. Instead, I leave these moments untargeted. They will serve as key yardsticks to assess the model's ability to explain differences across firms in workforce composition and pay.

4.1.2 Results

Table 2 summarizes the baseline parameterization of the model, including the internally estimated parameters. The key parameter γ in 2010 is inferred to be 0.638, pointing to non-trivial complementarities. Table 3 shows the model fit. The model is capable of matching the targeted moments perfectly, ψ being exact-identified.

I now turn to an evaluation of the model in terms of untargeted moments. Starting with matching patterns, the theoretical coworker correlation slightly undershoots its empirical counterpart but is of similar magnitude (0.43 vs. 0.56) and indicates substantial, positive coworker sorting. Figure 8 plots, for each type, the average coworker type (after binning into deciles) according to both the model and the data. Noting that in the estimation, the only moment that informs the degree of coworker sorting is the average coworker wage complementarity estimated in the micro data, the model fit is good.

What about between-firm inequality? As indicated before, the estimation exercise assumes that there is no permanent, ex-ante heterogeneity across firms. Insofar as there are differences across firms in the stationary equilibrium, in terms of wages but also labor productivity, then those emerge purely due to ex-post heterogeneity in firms' workforce composition; nor do they reflect any intrinsic feature of firms. In Section 5, we take a detailed look at the model-implied decomposition of wage dispersion into between- and within-firm components. In brief, this analysis reveals that

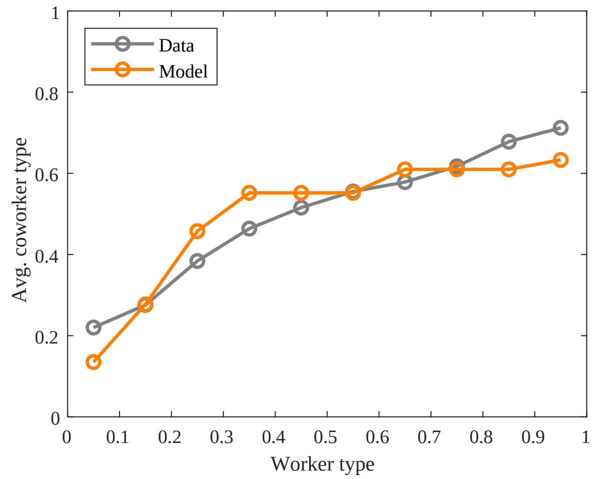


Figure 8: Average coworker: data vs. model (untargeted)

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

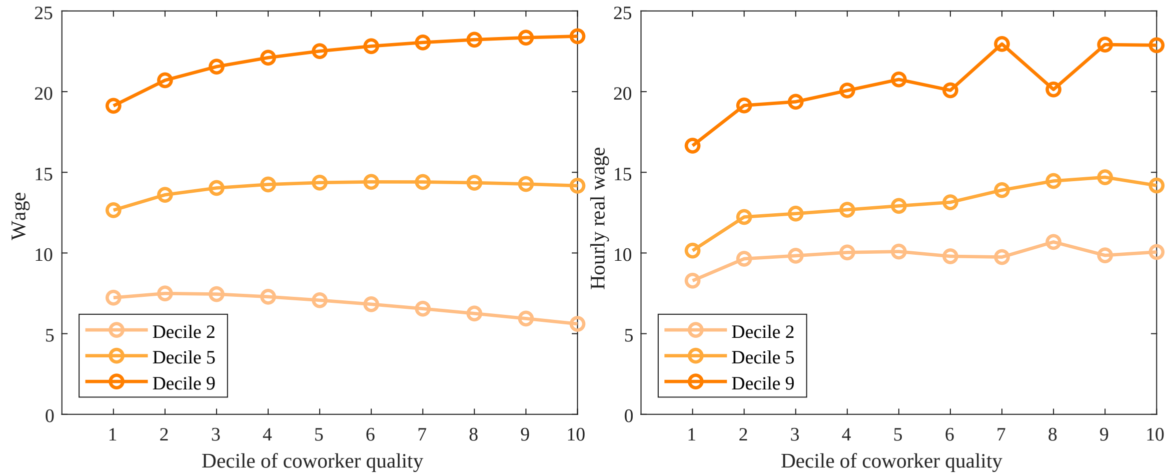
the stationary equilibrium of the theoretical model fits the data very well along that dimension. While I put less emphasis on the productivity dimension, it warrants highlighting that the estimated model generates substantial ex-post dispersion in firm productivity as well. For reference, Syverson (2004) reports an average 90:10 labor productivity ratio of 4:1 (the same ratio is about 2:1 when considering total factor productivity). The model prediction is a 90:10 labor productivity ratio of 2.94:1, thus coming quite close to the data even without assuming ex-ante heterogeneity among firms.

Continuing the analysis, Figure 9 depicts the wage function, plotted for the second, median and ninth deciles of the worker productivity distribution, in the model and when estimated non-parametrically in the 2010 LIAB sample.[62] The pictures broadly align, though certainly not perfectly. In particular, the model implies that for the bottom decile, the wage is actually decreasing in the coworker type, whereas it is flat in the data. On the flipside, the gain accruing to a top worker from climbing up the coworker ladder is understated by the model. What underlies this discrepancy is that the model overstates the weigh of outside options in influencing wages. A low-type has to compensate the firm for the foregone opportunity from pairing its other worker with a better match.[63] The divergence is of limited significance in practice, because, by dint of strong coworker sorting, there are very few matches between low and high types in any case.

Overall, therefore, the calibrated model fits the data describing the German economy in 2010 quite well, in spite of its parsimonious structure.

[62] I show the second and the ninth decile, instead of the bottom and the top, to minimize potential measurement error due to censoring.

[63] Allowing for replacement hiring would improve the capacity of the model to fit the data in this respect.



(a) Model

(b) Data (2010)

Figure 9: Non-parametric wage functions: model vs. data

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

4.2 Re-estimation for 1990

Finally, we can re-estimate the parameter vector ψ by targeting the same empirical moments as in the preceding step, but measured in 1990. The final column in Table 3 indicates the estimated parameter values and compared the targeted moment in model and data.

As hypothesized, the estimated elasticity of complementarity is lower, at 0.383, than the value of 0.638 we had estimated for 2010. In addition, the rise in a_1 from 1990 to 2010 alongside a decline in a_0 is consistent with skill-biased technological change, which renders match output more sensitive to team quality over time.

5 Quantitative analysis

Using the estimated model we can perform a sequence of quantitative analyses. I start by putting a number on the equilibrium loss in productivity resulting from the interaction of coworker complementarities and search frictions. I then quantify the contribution of strengthening complementarities to increased between-firm wage inequality in Germany.

5.1 Productivity costs of coworker quality mismatch & induced reallocation

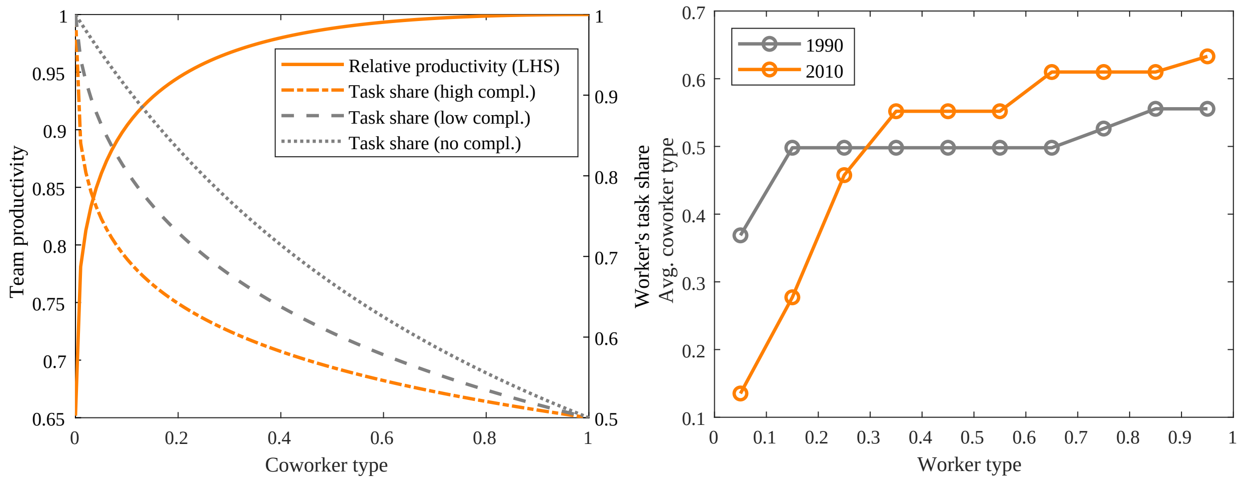
This paper started out with the observation that the division of labor is a prominent source of efficiency gains. Section 2.1 showed, however, that once we allow for heterogeneity in coworkers' quality, the division of labor naturally gives rise to complementarities in these qualities. As a result, team productivity is disproportionately influenced by the least capable team member. If workers are hired from frictionless labor market, the equilibrium allocation of workers to teams coincides with the efficient allocation; any team consists of workers of the same quality and there is no productivity loss due to coworker complementarities. When match quality has to be weighed against search costs however, the same is not true.

The efficient benchmark points to the following experiment to quantify the productivity costs due to coworker mismatch. Taking as given the distribution of worker types employed at multi-worker firms, how much output would they produce if, counterfactually, if we re-jigged the assignment of workers into teams so as to feature pure positive sorting?

Evaluating production at the factual and counterfactual match distribution, at the parameter values estimated for 1990, implies that per capita output would be 1.78% higher if there were no coworker quality mismatch. Performing the same counterfactual exercise for 2010 indicates only slightly higher costs of mismatch, at about 2.09% of output per capita. To put this number in perspective, the estimated gain from eliminating mismatch in 2010 is of similar magnitude to that reported by Hagedorn *et al.* (2017, Fig. 5(b)) in their assessment of output losses due to worker-firm complementarities in a search environment.

To disentangle the underlying mechanism, it is instructive to proceed in two steps, first holding constant the worker-firm assignment and then considering changes therein. Thus, note first that stronger complementarities will worsen the cost of mismatch for any given any allocation of workers into teams. Figure 10a offers an interpretation through the lens of the microfoundation set out in Section 2.1. The solid line plots the productivity of a team composed of one worker of the highest quality type depending on the coworker's quality type (horizontal axis), under the 2010 calibration, *relative* to the productivity of the same team but with the elasticity of complementarity at the 1990 level. We see that the team produces less output in 2010 unless paired up with a worker of the same quality, and the productivity cost is greater if the coworker's quality type is relatively low.

The team production model connects these *quality mismatch* costs with *task mismatch*, in the sense that the high-quality worker is performing tasks that she is relatively less productive in and would not have to do if teamed up optimally. The dashed lines indicate the task assignment patterns. If the comparative advantage parameter χ is zero (dotted line) or relatively low (1990; dashed), the firm makes the high-quality worker extend the share of tasks that she is performing when paired with a lower quality coworker. As long as comparative advantage is not particularly important, such changes in task assignment are of little consequence for productivity. By contrast, such substitution along the extensive, task assignment margin is harder under the 2010 calibration, as the high-quality worker's task-specific productivity declines rapidly when assigned to tasks she is not specialized in.



(a) Tasks and mismatch

(b) Changes in coworker sorting

Figure 10: Mismatch costs and model-implied coworker sorting

Notes. The left-hand panel illustrates the share of tasks assigned to a worker of type $x = 1$ depending on the quality of the coworker, x', as well as the productivity of this team under γ^{2010} relative to γ^{1990}. For 2010 I set aside the theoretical upper bound on χ. The right-hand panel depicts, for every worker type, the average coworker type, for the 1990 and 2010 calibrations, respectively.

As a consequence, the low quality of the coworker functions exerts a greater influence on the overall productivity of the team.

In a second step, we take into account that the observed rise in coworker sorting means that the distance between factual and efficient matching distribution has diminished over time. The right-hand panel of Figure 10b illustrates this labor market reallocation as captured by the estimated model. Through the lens of the structural model, the increased segregation of worker types by quality has occurred precisely the market allocation mechanism – with wages serving as price signals – pushes towards greater sorting when complementarities are stronger. The model thus highlights that the rise in coworker sorting and ensuing, widened gaps between firms in productivity and average wages need not reflect worsening frictions such as product market power; in this model, they are equilibrium outcomes that limit productivity losses due to strengthening for a given degree of input market frictions.[64] Indeed, the estimation implies that absent any worker reallocation, productivity in 2010 would be 2.68% lower.[65] What are the distributional implications of this reallocation?

[64]To be very clear, the model does not feature product market frictions or many other features that could render the increase in coworker sorting an inefficient outcome. The paper can, thus, not say much about the relevance of these features except noting that they are not strictly speaking needed to explain the observed changes.

[65]To compute this figure, I take the estimated production function and unconditional distribution of types in multi-worker for 2010 but suppose, counterfactually, that each worker type's conditional coworker type distribution is as in 1990. This approach avoid confounding changes in the unconditional distribution of quality types with changes in coworker matching patterns.

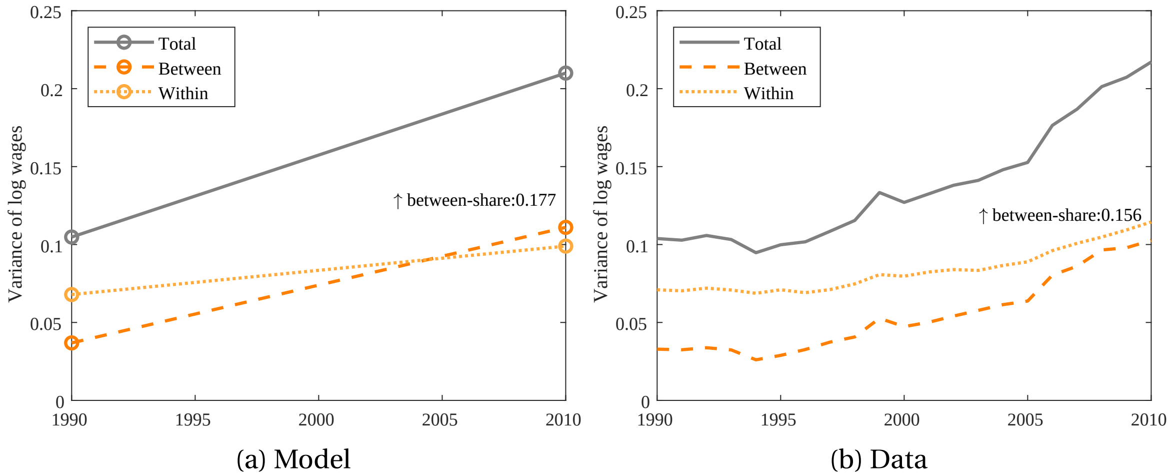


Figure 11: Between-within decomposition of the variance of log wages: model vs. data

Notes. This figure shows the between-within decomposition of the variance of log wages as predicted by the model (left panel) and in the data (right panel). All model-implied moments are constructed based on the subset of workers employed in teams and the between-within decomposition is corrected for statistical bias as described in the main text and Appendix C.3. The empirical moments are computed on the basis of the same residualized wages used in the estimation of the model.

5.2 Rising between-firm inequality

In parallel to starker labor market segregation, the estimated model predicts an increase in the between-firm share of wage inequality. Figure 11 depicts the model-implied decomposition of the variance of log wages at the stationary equilibrium when computed using the two sets of parameters, once for 1990 and once for 2010. The data counterpart, computed at a yearly level, is displayed alongside and illustrates the "firming up of inequality" (Song *et al.*, 2019) in the German economy. Yet, in marked contrast to an abundant empirical literature documenting this pattern empirically across many countries, there is a scarcity of structural models providing a quantitative explanation. This uses the estimated model to provide such an account.[66]

Before we evaluate the picture emerging from Figure 11, it bears emphasizing that for the model I report a between-within decomposition that corrects for the statistical upward-bias in the between-share of wage inequality that arises because in the theoretical model, the team size is normalized to two (so that the law of large numbers does not apply within production units). Appendix C.3 explains the method I propose and adopt to correct this statistical bias. I apply the same correction factor for both 1990 and 2010, so the predicted change is unaffected in any case. Insofar as one

[66]Inequality between firms is, per se, arguably less of a concern than between individuals. As discussed by De Loecker *et al.* (2022), it may reflect, however, an underlying market friction or be a source inequality between persons. For instance, increased between-firm inequality may reflect improved or worsened factor allocation, depending on the driving force. Dynamically, worker segregation could foster wage inequality and lower productivity growth, if learning from coworker is important (Jarosch *et al.*, 2021; Herkenhoff *et al.*, 2022) or if worker mobility facilitates technology diffusion from leading firms to laggards (Akcigit and Ates, 2019).

might also care about the level of either component, the model predictions would be misleading without statistical correction.

Figure 11 indicates that the between-within decomposition implied by the model fits the data quite well at both points in time – which is noteworthy since the estimation only targeted the *total* variance of log wages. It thus also correctly predicts the "firming up of inequality" that characterizes the German economy over the sample period. In terms of magnitudes, the model slightly overpredicts the increase in the between-firm share (0.177 vs. 0.156).

How much of the rise in the between-firm share of wage inequality predicted by the model is due to a strengthening in coworker complementarities? Ex ante, the answer is not obvious, as noted also by Song *et al.* (2019). Even with a stable distribution of worker types across firms, insofar as that distribution exhibits positive sorting, technological change that amplifies the return to skill should mechanically lead to greater between-firm inequality: the remuneration of already highly-paid employees, who tend to cluster together, would rise relative to less well paid employees in other firms.[67]

	Δ data	Δ model	Implied % Δ model due to $\uparrow \gamma$
Baseline	0.156	0.177	-
Scenario 1: only $\uparrow \gamma$	0.156	0.0935	52.9
Scenario 2: fix γ (1990)	0.156	0.0942	46.7
Within-occupation ranking	0.156	0.1237	-
Scenario 1b: only $\uparrow \gamma$	0.156	0.0293	23.7
Scenario 2b: fix γ (1990)	0.156	0.0948	23.3

Table 4: Counterfactual evaluation of rising between-firm inequality

Notes. This table summarizes the results of the counterfactual exercises described in the main text. Δdata is the empirical percentage point change in the between-share of the variance of log wages in the LIAB. The second column is the model counterpart. The percent figure in the final column is rounded to nearest integer. Denoting by m_j the between-share, in scenario 1, that final column value is computed as $100 \times \frac{m_j(\gamma^{2010},\psi_{-1}^{1990})-m_j(\psi^{1990})}{m_j(\psi^{2010})-M_j(\psi^{1990})}$, where $\psi_{-1} = \{a_0, a_1, M_u\}$. In scenario 2, the final column value is equal to $100 \times \left(1 - \frac{m_j(\gamma^{1990},\psi_{-1}^{2010})-m_j(\psi^{1990})}{m_j(\psi^{2010})-m_j(\psi^{1990})}\right)$. The lower half of the table is based on a ranking of workers within their respective two-digit occupation.

Table 4 answers this question by considering the counterfactual, "What would the model-predicted between-firm share of wage inequality have been in 1990 if the elasticity of complementarities, γ, had been at its 2010 level?", and computing the implied difference in the between-share from 1990 to 2010 due to γ having increased from 0.383 to 0.638 ("Scenario 1"). This exercise suggests that increased coworker complementarities, by themselves, can explain about half of the rise in

[67]Through the lens of an AKM-framework, increased return to skill could magnify both the worker-firm sorting and the worker-worker segregation components. Song *et al.* (2019) present a calculation suggesting that, respectively, 9% and 35% of the increase in sorting and segregation they document for the U.S. are due to these mechanical effects.

the between-firm share of wage inequality predicted by the model, which in turn corresponds to the empirically observed increase. As there is no unique decomposition, Table 4 also considers an alternative counterfactual. In Scenario 2, the elasticity of complementarity is fixed at its 1990 level. We see that the aforementioned effect of rising returns to skill ($a_1 \uparrow$) nonetheless pushes up the between-share but the rise is only half as large as in the baseline. The implied increase in the between-firm share from 1990 to 2010 due to γ is slightly smaller than that coming out of Scenario 1 but of similar magnitude.

A further exercise serves as a more conservative robustness check. While the structural model does not explicitly incorporate the reality of multiple, distinct occupations, we may be concerned that the empirically measured increase in coworker complementarities, or the predicted increase in between-firm inequality, is confounded by an increasing concentration of high-skill occupations in the same establishments. Appendix sections A.4.3 and A.4.4 establish that between-firm wage inequality and coworker sorting have increased even after accounting for occupational heterogeneity, though the magnitudes are somewhat smaller, consistent with the analyses of Card *et al.* (2013) and Handwerker (2020). Suppose, in particular, that instead of ranking workers economy wide, we classify them within their respective two-digit occupations. Given this new ranking of workers, I re-estimate regression (35) for each of the five sample periods. I find that the coefficient on the interaction term still increases, but the difference between the first period (0.065) and the last period (0.115) is now smaller than under the baseline approach. I re-estimate the structural model for both 1990 and 2010, keeping all targets the same as before except the wage complementarity moment that informs γ.[68] The newly estimated values of the elasticity of complementarity are 0.4571 (1990) and 0.5563 (2010). Repeating the counterfactual analyses suggests that under this more demanding specification, the increase in complementarities still explains nearly one quarter of the rise in the overall between-establishment share of wage inequality.

Summarizing the findings from Section 5, the rise in coworker complementarities has induced a distribution of workers and employers that is closer to the frictionless allocation; productivity is lower, but would have been even lower had no reallocation taken place. As a consequence, the economy is characterized by more marked coworker sorting, and thus wider gaps in the average wage paid by different firms. In particular, greater coworker complementarities, disciplined by the micro evidence, can account for around one quarter to one half of the rise in the between-establishment share of (residual) wage inequality observed in Germany from 1990 to 2010.

[68] For practical reasons relating to the analysis of the data at the FDZ-IAB, I use the regression-based estimates of wage complementarities.

6 Concluding remarks

This paper developed, characterized, and quantified a model of production with teams of heterogeneous workers who are hired on a search-frictional labor market. Conceptually, the aim is to open one small window into the black box of production, shedding light on the origins and implications of one cog in the machinery of firms – complementarities between coworkers' qualities. Constructing a firm-production function from the bottom up made it possible to consider how changes in the nature of work, happening at the micro level, shape the allocation of workers and firms, productivity, and wage inequality at the macro level.

In sum, this analysis yielded three take-aways. (1) When production involves division of labor among workers who are specialized in particular tasks, the quality of these workers are naturally complements; the more differentiated workers' task-specific skills are, the stronger the complementarities. (2) When labor markets are characterized by search frictions, coworker complementarities induce lower productivity, as coworker quality mismatch costs output, and bigger gaps between firms in workforce quality and pay. (3) Coworker complementarities have strengthened over time, paralleling an increase in the complexity of tasks, which helps rationalize the empirically observed rise in between-firm wage inequality.

The analysis also points toward several avenues for future research. One direction notes that the model here assumed the division of labor within firms to be perfect. But in practice, coordination frictions likely limit this division of labor (Becker and Murphy, 1992), and some firms are better than others in coordinating their workforce, as documented empirically by Coraggio *et al.* (2022) and Kuhn *et al.* (2022). Freund (2022) generalizes the team production model to incorporate such heterogeneity in coordination quality. The same trends in the nature of work considered in the present paper are predicted to have made this organizational capacity more important for productivity. This theoretical conjecture requires empirical and quantitative evaluation. A second avenue would be to carefully and quantitatively evaluate the implications of coworker complementarities for policies. For instance, labor market policies that impinge on labor market reallocation – whether through restrictions or blunted incentives to find the best match – are likely to worsen mismatch along this dimension. On the other hand, and when allowing for incomplete markets, policies that allow unemployed workers greater time to search for a good coworker have the potential to boost productivity (cf. Eeckhout and Sepahsalari, 2021; Huang and Qiu, 2022).

References

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, **67**(2), 251–333.

Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Acemoglu, D. and Restrepo, P. (2018). The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, **108**(6), 1488–1542.

Adenbaum, J. (2022). Endogenous Firm Structure and Worker Specialization. *mimeo*.

Adhvaryu, A., Bassi, V., Nyshadham, A., and Tamayo, J. A. (2020). No Line Left Behind: Assortative Matching Inside the Firm.

Ahmadpoor, M. and Jones, B. F. (2019). Decoding team and individual impact in science and invention. *Proceedings of the National Academy of Sciences*, **116**(28), 13885–13890.

Akcigit, U. and Ates, S. T. (2019). What Happened to U.S. Business Dynamism? Working Paper 25756, National Bureau of Economic Research.

Allen, T. (2019). Modern Spatial Economics.

Alon, T. (2018). Earning more by doing less: Human capital specialization and the college wage premium. *manuscript, University of California at San Diego*.

Alvarez, F. and Lucas, R. E. (2007). General equilibrium analysis of the Eaton–Kortum model of international trade. *Journal of Monetary Economics*, **54**(6), 1726–1768.

Alvarez, J., Benguria, F., Engbom, N., and Moser, C. (2018). Firms and the Decline in Earnings Inequality in Brazil. *American Economic Journal: Macroeconomics*, **10**(1), 149–189.

Andrews, D., Criscuolo, C., and Gal, P. N. (2019). The Best versus the Rest: Divergence across Firms During the Global Productivity Slowdown. *Working Paper*.

Andrews, M. J., Gill, L., Schank, T., and Upward, R. (2008). High wage workers and low wage firms: Negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **171**(3), 673–697.

Antonczyk, D., Fitzenberger, B., and Leuschner, U. (2009). Can a Task-Based Approach Explain the Recent Changes in the German Wage Structure? *Jahrbücher für Nationalökonomie und Statistik*, **229**(2-3), 214–238.

Arcidiacono, P., Foster, G., Goodpaster, N., and Kinsler, J. (2012). Estimating spillovers using panel data, with an application to the classroom. *Quantitative Economics*, **3**(3), 421–470.

Atakan, A. E. (2006). Assortative Matching with Explicit Search Costs. *Econometrica*, **74**(3), 667–680.

Atalay, E., Phongthiengtham, P., Sotelo, S., and Tannenbaum, D. (2020). The Evolution of Work in the United States. *American Economic Journal: Applied Economics*, **12**(2), 1–34.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The Fall of the Labor Share and the Rise of Superstar Firms. *The Quarterly Journal of Economics*, **135**(2), 645–709.

Autor, D., Chin, C., Salomons, A. M., and Seegmiller, B. (2022). New Frontiers: The Origins and Content of New Work, 1940–2018.

Autor, D. H. and Handel, M. J. (2013). Putting Tasks to the Test: Human Capital, Job Tasks, and Wages. *Journal of Labor Economics*, **31**(S1), S59–S96.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, **118**(4), 1279–1333.

Babet, D., Godechot, O., and Palladino, M. G. (2022). In the Land of AKM: Explaining the Dynamics of Wage Inequality in France. *mimeo*.

Bagger, J. and Lentz, R. (2019). An Empirical Model of Wage Dispersion with Sorting. *The Review of Economic Studies*, **86**(1), 153–190.

Barth, E., Bryson, A., Davis, J. C., and Freeman, R. (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.

Barth, E., Davis, J., and Freeman, R. B. (2018). Augmenting the Human Capital Earnings Equation with Measures of Where People Work. *Journal of Labor Economics*, **36**(S1), S71–S97.

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy*, **81**(4), 813–846.

Becker, G. S. and Murphy, K. M. (1992). The Division of Labor, Coordination Costs, and Knowledge. *The Quarterly Journal of Economics*, **107**(4), 1137–1160.

Berlingieri, G., Blanchenay, P., and Criscuolo, C. (2017). The great divergence(s). *OECD Discussion Paper*.

Bilal, A. and Lhuillier, H. (2021). Outsourcing, Inequality and Aggregate Output.

Bilal, A., Engbom, N., Mongey, S., and Violante, G. L. (2022). Firm and Worker Dynamics in a Frictional Labor Market. *Econometrica*, **90**(4), 1425–1462.

Bloesch, J., Larsen, B., and Taska, B. (2022). Which Workers Earn More at Productive Firms? Position Specific Skills and Individual Worker Hold-up Power. *mimeo*.

Bloom, N. and Reenen, J. V. (2011). Chapter 19 - Human Resource Management and Productivity. In D. Card and O. Ashenfelter, editors, *Handbook of Labor Economics*, volume 4, pages 1697–1767. Elsevier.

Bodker, J. E. (2019). Coworker Sorting: Wages, Productivity and Inequality. *mimeo*.

Boehm, J. and Oberfield, E. (2022). Growth and the Fragmentation of Production. *mimeo*.

Boerma, J., Tsyvinski, A., and Zimin, A. (2021). Sorting with Team Formation. Technical Report w29290, National Bureau of Economic Research, Cambridge, MA.

Bombardini, M., Gallipoli, G., and Pupato, G. (2012). Skill Dispersion and Trade Flows. *American Economic Review*, **102**(5), 2327–2348.

Bonhomme, S., Lamadon, T., and Manresa, E. (2019). A Distributional Framework for Matched Employer Employee Data. *Econometrica*, **87**(3), 699–739.

Boning, B., Ichniowski, C., and Shaw, K. (2007). Opportunity Counts: Teams and the Effectiveness of Production Incentives. *Journal of Labor Economics*, **25**(4), 613–650.

Borovičková, K. and Shimer, R. (2020). High Wage Workers Work for High Wage Firms. Technical Report w24074, National Bureau of Economic Research, Cambridge, MA.

Bresnahan, T., Brynjolfsson, E., and Hitt, L. M. (2002). Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence. *The Quarterly Journal of Economics*, **117**(1), 339–376.

Burdett, K. and Coles, M. G. (1999). Long-Term Partnership Formation: Marriage and Employment. *The Economic Journal*, **109**(456), F307–F334.

Caliendo, L., Mion, G., Opromolla, L. D., and Rossi-Hansberg, E. (2020). Productivity and Organization in Portuguese Firms. *Journal of Political Economy*.

Caplin, A., Lee, M., Leth-Petersen, S., Saeverud, J., and Shapiro, M. D. (2022). How Worker Productivity and Wages Grow with Tenure and Experience: The Firm Perspective.

Card, D., Heining, J., and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, **128**(3), 967–1015.

Card, D., Cardoso, A. R., and Kline, P. (2016). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *The Quarterly Journal of Economics*, **131**(2), 633–686.

Cardoso, A. R., Guimaraes, P., Portugal, P., and Reis, H. (2018). The Returns to Schooling Unveiled. Technical Report 11419, Institute of Labor Economics (IZA).

Chade, H. and Eeckhout, J. (2020). Competing Teams. *The Review of Economic Studies*, **87**(3), 1134–1173.

Chaney, T. and Ossa, R. (2013). Market size, division of labor, and firm productivity. *Journal of International Economics*, **90**(1), 177–180.

Coles, M. G. and Francesconi, M. (2019). Equilibrium Search with Multiple Attributes and the Impact of Equal Opportunities for Women. *Journal of Political Economy*, **127**(1), 138–162.

Coraggio, L., Pagano, M., Scognamiglio, A., and Tåg, J. (2022). JAQ of All Trades: Job Mismatch, Firm Productivity and Managerial Quality. *SSRN Electronic Journal*.

Cornelissen, T., Dustmann, C., and Schönberg, U. (2017). Peer Effects in the Workplace. *American Economic Review*, **107**(2), 425–456.

Correia, S. (2017). Linear models with high-dimensional fixed effects: An efficient and feasible estimator. *Unpublished manuscript, http://scorreia. com/research/hdfe. pdf (last accessed 25 October 2019).*

Costinot, A. (2009). On the origins of comparative advantage. *Journal of International Economics*, **77**(2), 255–264.

Criscuolo, C., Hijzen, A., Schwellnus, C., Barth, E., Chen, W.-H., Fabling, R., Fialho, P., Stadler, B., Upward, R., Zwysen, W., Grabska-Romagosa, K., Kambayashi, R., Leidecker, T., Skans, O. N., Riom, C., and Roth, D. (2021). Workforce composition, productivity and pay: The role of firms in wage inequality. *OECD Discussion Paper*.

Criscuolo, C., Freund, L. B., and Gal, P. N. (2023). Human Capital at Work: Skill Composition, Productivity, and Wages. *OECD Discussion Paper*.

Dauth, W. and Eppelsheimer, J. (2020). Preparing the sample of integrated labour market biographies (SIAB) for scientific analysis: A guide. *Journal for Labour Market Research*, **54**(1), 1–14.

De Loecker, J., Obermeier, T., and Van Reenen, J. (2022). Firms and inequalities. Technical report, Centre for Economic Performance, LSE.

De Ridder, M. (2021). Market Power and Innovation in the Intangible Economy. *mimeo*.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2020). Changing Business Dynamism and Productivity: Shocks versus Responsiveness. *American Economic Review*, **110**(12), 3952–3990.

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics*, **132**(4), 1593–1640.

Dessein, W. and Santos, T. (2006). Adaptive Organizations. *Journal of Political Economy*, **114**(5), 956–995.

Duranton, G. and Puga, D. (2004). Chapter 48 - Micro-Foundations of Urban Agglomeration Economies. In J. V. Henderson and J.-F. Thisse, editors, *Handbook of Regional and Urban Economics*, volume 4 of *Cities and Geography*, pages 2063–2117. Elsevier.

Dustmann, C., Ludsteck, J., and Schönberg, U. (2009). Revisiting the German Wage Structure. *The Quarterly Journal of Economics*, **124**(2), 843–881.

Dvorkin, M. and Monge-Naranjo, A. (2019). Occupation Mobility, Human Capital and the Aggregate Consequences of Task-Biased Innovations. *St. Louis Federal Reserve Bank Working Paper 2019-013C*.

Eaton, J. and Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica*, **70**(5), 1741–1779.

Eaton, J., Kortum, S., Neiman, B., and Romalis, J. (2016). Trade and the Global Recession. *American Economic Review*, **106**(11), 3401–3438.

Eberle, J. and Schmucker, A. (2017). Creating cross-sectional data and biographical variables with the Sample of Integrated Labour Market Biographies 1975-2014. Programming examples for Stata. page 18.

Edmond, C. and Mongey, S. (2021). Unbundling Labor. *mimeo*.

Eeckhout, J. and Kircher, P. (2011). Identifying Sorting—In Theory. *The Review of Economic Studies*, **78**(3), 872–906.

Eeckhout, J. and Sepahsalari, A. (2021). The Effect of Wealth on Worker Productivity. *mimeo*, (20/731).

Elsby, M. W. L. and Gottfries, A. (2022). Firm Dynamics, On-the-Job Search, and Labor Market Fluctuations. *The Review of Economic Studies*, **89**(3), 1370–1419.

Freund, L. B. (2022). Coordination is Key: Task Assignment and Firm Productivity. *mimeo*.

Garicano, L. (2000). Hierarchies and the Organization of Knowledge in Production. *Journal of Political Economy*, **108**(5), 874–904.

Garicano, L. and Hubbard, T. N. (2012). Learning about the nature of production from equilibrium assignment patterns. *Journal of Economic Behavior & Organization*, **84**(1), 136–153.

Garicano, L. and Rossi-Hansberg, E. (2006). Organization and Inequality in a Knowledge Economy. *The Quarterly Journal of Economics*, **121**(4), 1383–1435.

Garicano, L. and Rossi-Hansberg, E. (2015). Knowledge-Based Hierarchies: Using Organizations to Understand the Economy. *Annual Review of Economics*, **7**(1), 1–30.

Gartner, H. (2005). The imputation of wages above the contribution limit with the German IAB employment sample. *FDZ Methodenreport 2/2005*.

Goldschmidt, D. and Schmieder, J. F. (2017). The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure. *The Quarterly Journal of Economics*, **132**(3), 1165–1217.

Grigsby, J. (2022). Skill Heterogeneity and Aggregate Labor Market Dynamics. *mimeo*.

Grossman, G. M. and Helpman, E. (1990). Comparative Advantage and Long-Run Growth. *The American Economic Review*, **80**(4), 796–815.

Hagedorn, M., Law, T. H., and Manovskii, I. (2017). Identifying Equilibrium Models of Labor Market Sorting. *Econometrica*, **85**(1), 29–65.

Håkanson, C., Lindqvist, E., and Vlachos, J. (2021). Firms and Skills The Evolution of Worker Sorting. *Journal of Human Resources*, **56**(2), 512–538.

Hall, A. and Rohrbach-Schmidt (2020). BIBB/BAuA Employment Survey 2018. BIBB-FDZ Data and Methodological Report 1/2020. *BIBB Report*.

Haltiwanger, J. C. and Spletzer, J. R. (2020). Between Firm Changes in Earnings Inequality: The Dominant Role of Industry Effects. Working Paper 26786, National Bureau of Economic Research.

Handwerker, E. W. (2020). Outsourcing, Occupationally Homogeneous Employers, and Growing Wage Inequality in the United States. *BLS Economic Working Paper WP-522*.

Herkenhoff, K., Lise, J., Menzio, G., and Phillips, G. M. (2022). Production and Learning in Teams. Working Paper 25179.

Hobijn, B. and Sahin, A. (2009). Job-finding and separation rates in the OECD. *Economics Letters*, **104**(3), 107–111.

Hong, L. (2022). Coworker Sorting, Learning, and Wage Inequality. *mimeo*.

Hong, L. and Lattanzio, S. (2022). The Peer Effect on Future Wages in the Workplace. *Available at SSRN*.

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.

Huang, E. J. and Qiu, X. (2022). Precautionary mismatch. *mimeo*.

Jäger, S. and Heining, J. (2022). How Substitutable Are Workers? Evidence from Worker Deaths. page 52.

Jarosch, G., Oberfield, E., and Rossi-Hansberg, E. (2021). Learning From Coworkers. *Econometrica*, **89**(2), 647–676.

Jones, B. 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.

Jones, B. F. (2021). The Rise of Research Teams: Benefits and Costs in Economics. *Journal of Economic Perspectives*, **35**(2), 191–216.

Jones, C. I. (2011). Intermediate Goods and Weak Links in the Theory of Economic Development. *American Economic Journal: Macroeconomics*, **3**(2), 1–28.

Jung, P. and Kuhn, M. (2014). Labour Market Institutions and Worker Flows: Comparing Germany and the US. *The Economic Journal*, **124**(581), 1317–1342.

Kleinman, B. (2022). Wage Inequality and the Spatial Expansion of Firms. *mimeo*.

Kline, P., Saggio, R., and Sølvsten, M. (2020). Leave-Out Estimation of Variance Components. *Econometrica*, **88**(5), 1859–1898.

Kremer, M. (1993). The O-Ring Theory of Economic Development. *The Quarterly Journal of Economics*, **108**(3), 551–575.

Kremer, M. and Maskin, E. (1996). Wage Inequality and Segregation by Skill. Working Paper 5718, National Bureau of Economic Research.

Krusell, P., Ohanian, L. E., Ríos-Rull, J.-V., and Violante, G. L. (2000). Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica*, **68**(5), 1029–1053.

Kuhn, M., Luo, J., Manovskii, I., and Qiu, X. (2022). Coordinated Firm-Level Work Processes and Macroeconomic Resilience. *mimeo*.

Lazear, E. P. and Shaw, K. L. (2007). Personnel Economics: The Economist's View of Human Resources. page 37.

Lin, J. (2011). Technological Adaption, Cities, and New Work. *The Review of Economics and Statistics*, page 21.

Lind, N. and Ramondo, N. (2018). Trade with Correlation.

Lindenlaub, I. (2017). Sorting Multidimensional Types: Theory and Application. *The Review of Economic Studies*, **84**(2), 718–789.

Lopes de Melo, R. (2018). Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence. *Journal of Political Economy*, **126**(1), 313–346.

Lucas, R. (1978). On the Size Distribution of Business Firms. *Bell Journal of Economics*, **9**(2), 508–523.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, **60**(3), 531–542.

Marimon, R. and Zilibotti, F. (1999). Unemployment vs. Mismatch of Talents: Reconsidering Unemployment Benefits. *The Economic Journal*, **109**(455), 266–291.

Martellini, P. and Menzio, G. (2021). Jacks of All Trades and Masters of One: Declining Search Frictions and Unequal Growth. *American Economic Review: Insights*, **3**(3), 339–352.

Martinez, J. (2021). Putty-Clay Automation. *CEPR Discussion Paper DP16022*.

Mas, A. and Moretti, E. (2009). Peers at Work. *American Economic Review*, **99**(1), 112–145.

Mihaylov, E. and Tijdens, K. G. (2019). Measuring the Routine and Non-Routine Task Content of 427 Four-Digit ISCO-08 Occupations.

Minni, V. (2022). Making the Invisible Hand Visible: Managers and the Allocation of Workers to Jobs. *mimeo*.

Mion, G. and Opromolla, L. D. (2014). Managers' mobility, trade performance, and wages. *Journal of International Economics*, **94**(1), 85–101.

Mollick, E. R. (2011). People and Process, Suits and Innovators: The Role of Individuals in Firm Performance.

Neffke, F. M. H. (2019). The Value of Complementary Co-Workers. *Science Advances*, **5**(12), eaax3370.

Nix, E. (2020). Learning spillovers in the firm. Technical report, Working Paper.

Ocampo, S. (2022). A task-based theory of occupations with multidimensional heterogeneity. Technical report, Discussion Paper, retrieved from: https://sites. google. com/site/sergi ....

Palladino, M. G., Roulet, A., and Stabile, M. (2021). Understanding the Role of Firms in the Gender Wage Gap Over Time, Over the Life Cycle, and Across Worker Types.

Pearce, J. (2022). Idea Production and Team Structure. *Working paper*.

Porzio, T. (2017). Cross-Country Differences in the Optimal Allocation of Talent and Technology. *mimeo*.

Raveendran, M., Silvestri, L., and Gulati, R. (2020). The Role of Interdependence in the Micro-Foundations of Organization Design: Task, Goal, and Knowledge Interdependence. *Academy of Management Annals*, **14**(2), 828–868.

Rivkin, J. and Siggelkow, N. (2003). Balancing Search and Stability: Interdependencies Among Elements of Organizational Design. *Management Science*, **49**(3), 290–311.

Rohrbach-Schmidt, D. and Tiemann, M. (2013). Changes in workplace tasks in Germany—evaluating skill and task measures. *Journal for Labour Market Research*, **46**(3), 215–237.

Rosen, S. (1981). The Economics of Superstars. *The American Economic Review*, **71**(5), 845–858.

Rosen, S. (1982). Authority, Control, and the Distribution of Earnings. *The Bell Journal of Economics*, **13**(2), 311–323.

Rosen, S. (1983). Specialization and Human Capital. *Journal of Labor Economics*, **1**(1), 43–49.

Ruf, K., Schmidtlein, L., Seth, S., Stüber, H., Umkehrer, M., Griessemer, S., and Kaimer, S. (2021). Linked-Employer-Employee-Data of the IAB (LIAB): LIAB longitudinal model 1975-2019, version 1. *Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB)*.

Sandmann, C. and Bonneton, N. (2022). Existence of a non-stationary equilibrium in search-and-matching models: TU and NTU. *mimeo*.

Shimer, R. and Smith, L. (2000). Assortative Matching and Search. *Econometrica*, **68**(2), 343–369.

Silva, J., Leitao, M., and Montana, J. (2022). The Role of Firms in Wage Inequality Dynamics.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. W. Strahan and T. Cadell, London.

Song, J., Price, D. J., Guvenen, F., Bloom, N., and von Wachter, T. (2019). Firming Up Inequality. *The Quarterly Journal of Economics*, **134**(1), 1–50.

Sorkin, I. and Wallskog, M. (2021). The Slow Diffusion of Earnings Inequality. *Working Paper*.

Spitz-Oener, A. (2006). Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics*, **24**(2), 235–270.

Syverson, C. (2004). Product Substitutability and Productivity Dispersion. *The Review of Economics and Statistics*, **86**(2), 534–550.

Thompson, J. D. (1967). *Organizations in Action: Social Science Bases of Administrative Theory*. Organizations in Action: Social Science Bases of Administrative Theory. McGraw-Hill, New York, NY, US.

Tian, L. (2021). Division of Labor and Productivity Advantage of Cities: Theory and Evidence from Brazil. page 97.

Tomaskovic-Devey, D., Rainey, A., Avent-Holt, D., Bandelj, N., Boza, I., Cort, D., Godechot, O., Hajdu, G., Hällsten, M., Henriksen, L. F., Hermansen, A. S., Hou, F., Jung, J., Kanjuo-Mrčela, A., King, J., Kodama, N., Kristal, T., Křížková, A., Lippényi, Z., Melzer, S. M., Mun, E., Penner, A., Petersen, T., Poje, A., Safi, M., Thaning, M., and Tufail, Z. (2020). Rising between-workplace inequalities in high-income countries. *Proceedings of the National Academy of Sciences*, **117**(17), 9277–9283.

Weidmann, B. and Deming, D. J. (2021). Team Players: How Social Skills Improve Team Performance. *Econometrica*, **89**(6), 2637–2657.

Wilmers, N. and Aeppli, C. (2021). Consolidated Advantage: New Organizational Dynamics of Wage Inequality. *American Sociological Review*, **86**(6), 1100–1130.

Wood, S. and Bryson, A. (2009). High involvement management. pages 151–175. Cambridge University Press, Cambridge, UK.

Wuchty, S., Jones, B. F., and Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science*, **316**(5827), 1036–1039.

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

Lukas B. Freund

December 2022

Appendix

Table of Contents

A Empirical evidence

A.1 Data: German LIAB

This section provides further details on the processing and key characteristics of the longitudinal version of the Linked-Employer-Employee-Data (LIAB LM 7519). Access is provided by the Research Data Center of the German Federal Employment Agency at the Institute for Employment Research (IAB). A detailed description can be found in Ruf *et al.* (2021). To ensure best practices, I extensively rely on publicly available code by Eberle and Schmucker (2017) and Dauth and Eppelsheimer (2020).

The LIAB covers every worker at a random sample of establishments between the years 2008 and 2017 as well as, crucially, the complete employment biographies of each of these workers, even when not employed at the establishments in the sample. To maximize sample coverage, do not restrict myself to the establishments in the original sampling frame, but I do require a minimum number of persons in every establishment-year cell (see below).

The employment biographies come in spell format, where a spell can correspond either to employment or to a period of benefit receipt. I organize the resulting dataset as 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).

My sample selection criteria are similar to other studies using this dataset or studying similar topics. In a first step, I select employees aged 20-60 with workplaces in West German states who are liable to social security and are not in part-time or marginal employment (i.e., I limit the sample to full-time employeese). I also drop jobs with real daily earnings of less than 10 Euros. I drop observations in select industries and create a consistent industry classification at the 2-digit level of the OECD STAN-A38 nomenclature.[A.1]

A well-known and non-trivial limitation of the LIAB is that the earnings variable is top-coded at the so-called "contribution assessment limit" of the social security system ("Beitragsbemessungsgrenze"). To impute right-censored wages, I implement the recommended best practices. Specifically, I follow Card *et al.* (2013), who build on Gartner (2005) and Dustmann *et al.* (2009). This approach involves 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 (to include West German men only, in particular). I follow Card *et al.* (2013) in the specification of controls by including not only age, firm size, firm

[A.1] I drop Agriculture, forestry and fishing (1); Mining and quarrying (5); Utilities (35-36); as well as Activities of households as employers; undifferentiated activities of households for own use (97) and Activities of extraterritorial organizations and bodies (99). The selection partially aims to ensure consistency with analyses for the Portuguese case.

size squared and a dummy for firms with more than ten employees, but also the mean log wage of co-workers and fraction of co-workers with censored wages. Finally, following Dauth and Eppelsheimer (2020), whose publicly available code I heavily rely on, I limit imputed wages at 10 times the 99th percentile. I a second sample restriction step, I then drop establishment-year cells with fewer than ten full-time employees or worker-year observations and restrict attention to the largest connected set.

The final sample (1985-2017) includes 8,756,584 person-year observations for 686,761 unique persons, whose average age is 38.83. The median unweighted establishment size is 18. The average real daily wage in 2010 is 120 Euros (in 2015 units).

A.2 Data: Portuguese QP

The construction of the Portuguese dataset is described in detail in Criscuolo *et al.* (2023), from which the results shown in this paper are excerpted. In addition to Chiara Criscuolo and Peter Gal, who facilitated the analysis, I expressly thank the following individuals for their generous answers to my many questions: Ana Rute Cardoso, Priscilla Fialho, Paulo Guimarães, Pedro Portugal, Pedro Raposo and Marta Silva.

Here I provide a very brief summary. After an extensive data cleaning procedure to handle duplicate person and employer identifiers, we impose sample restrictions similar to those described in the preceding section. All observations retained relate to persons employed by third parties. Real hourly wages include base pay, regular benefits, and bonus pay. As wages are not top-coded, no imputation procedure is needed. I drop observations below the statutory minimum monthly pay, as in Cardoso *et al.* (2018). While primarily focussing on the 2010-2017 sample, where longer periods are considered I use a combination of official and unofficial but widely used harmonization crosswalks for occupations and industries.

The final sample (1986-2017) includes 21,200,256 person-year observations, of which 6,930,892 belong to the 2010-2017 sample. In 2017, there are 848,970 workers employed at 19,391 unique firms.

A.3 Measuring worker types

As stated in Section 3.2.3, I implement two alternative approaches to estimate each worker's time-invariant quality type in the data. One is to rank worker's by their fixed effect from Abowd *et al.* (1999) style log-linear wage regressions. An alternative that is less common by more consistent with the structure of the model is to adopt the non-parametric ranking algorithm proposed in Hagedorn *et al.* (2017). This section provides details on the implementation.

A.3.1 AKM models

This section describes the estimation of two-way fixed-effects regressions in the spirit of Abowd *et al.* (1999, AKM hereafter) to account for unobservable worker and firm effects. This approach is widely used in empirical studies of wage inequality and can, at a minimum, be viewed as a useful diagnostic tool.

My implementation approach follows the proposal of Bonhomme *et al.* (2019) to mitigate limited mobility bias, which is a form of incidental parameter bias arising from the fact that a large number of firm-specific parameters (i.e., fixed effects) that are solely identified from workers who move across employers. The idea is to reduce the dimensionality of the estimation problem by clustering similar firms.[A.2] To briefly review, clusters are found by solving a weighted k-means problem,

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

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

I use a baseline value of $K = 20$, but have experimented with $K = 10$ and $K = 100$ as well (the choice makes little practical difference, as reported also by Bonhomme *et al.* (2019)) and use firms' wage distributions over the entire sample period on a grid of 20 percentiles for clustering.

Different from Bonhomme *et al.* (2019), and in similarity to Palladino *et al.* (2021), I then stick to the two-way FE regression approach rather than estimating a correlated randon effects model. That is, after imputing a cluster to each worker-year observation, I then estimate two-way fixed effect (log) wage regression using cluster effects instead of firm effects:

$$\ln(w_{it}) = \alpha_i + \sum_{k=1}^{K} \psi_k \mathbf{1}(j(i,t) = k) + \epsilon_{it} \tag{A.2}$$

where $\mathbf{1}(J(i,t) = k)$ are dummies indicating which cluster k firm the employer of i in period t, $j(i,t)$ has been assigned to. I associate with each firm j the fixed effect of the cluster to which j belongs and denote it ψ_j. Just for the sake of exposition, I abstract from observables, but in practice I control for an index of time-varying characteristics that includes a cubic in age and a quadratic in

[A.2] For Portugal, I have experimented with a variety of bias-correction methods. In particular, the method discussed here as well as the approaches of Andrews *et al.* (2008) and Kline *et al.* (2020) yield very similar findings. Moreover, the uncorrected version appears significantly biased, overestimating the contributions of firm-specific pay premia and underestimating the degree of worker-firm and worker-worker segregation.

job tenure, exactly as in the construction of the residualized wages used to estimate the structural model.

Similar to Card *et al.* (2013), I estimate the AKM model separately for five overlapping sample periods. This estimation is implemented in Stata using the *reghdfe* package (Correia, 2017).

A.3.2 Non-parametric ranking algorithm

As is well known (Eeckhout and Kircher, 2011; Lopes de Melo, 2018; Bonhomme *et al.*, 2019), the restrictions imposed by the AKM model are inconsistent with common structural models, including the one set out in this paper. in particular, that correlations in output between workers at the same employer that are reflected in wages, are attributed to a common firm component, even if these correlations are due to coworker complementarities.

As an alternative way of estimating worker quality types, I therefore implement a version of the non-parametric ranking algorithm proposed in Hagedorn *et al.* (2017). The theoretical applicability of this method is proved in Appendix Section B.2.4.

The intuition is that since both the production function in equation (19) and the value of unemployment are increasing in worker productivity, wages within firms are also increasing in worker type. We can then infer a partial ranking of employees in a *given* establishment from their wages (Anna and Bob both work at Zara in Cambridge; if Anna's wage is higher than Bob's, the algorithm interprets that as Anna being ranked above Bob). If no firm matches with all workers, we have to aggregate the partial within-firm rankings to a global one by exploiting the logic of transitivity and worker mobility across employers (Bob changes job to work for H&M and has a new coworker, Carl; if Bob is ranked above Carl, then we infer Anna's rank to be greater than Carl's.)

To handle potential inconsistencies across partial rankings, measurement error being one relevant source, Hagedorn *et al.* (2017, Online Appendix C) adapt the Kemeny-Young method, which minimizes the sum of Kendall-tau distances between two rankings, and propose a computationally feasible approximation to the NP-hard problem of uncovering the exact solution. For convergence I require the correlation between the rankings obtained from consecutive iterations to be at least 0.995. The implementation in Octave, run on the FDZ-IAB servers, takes around a week when a ranking is created separately for each of five sample periods.

A.4 Additional empirical results

A.4.1 Cross-country evidence on the "firming up" of between-firm inequality

Complementing the evidence for Germany and Portugal constructed from micro-data, Figure A.1 illustrates cross-country trends for the between-firm share of wage inequality, as gathered by

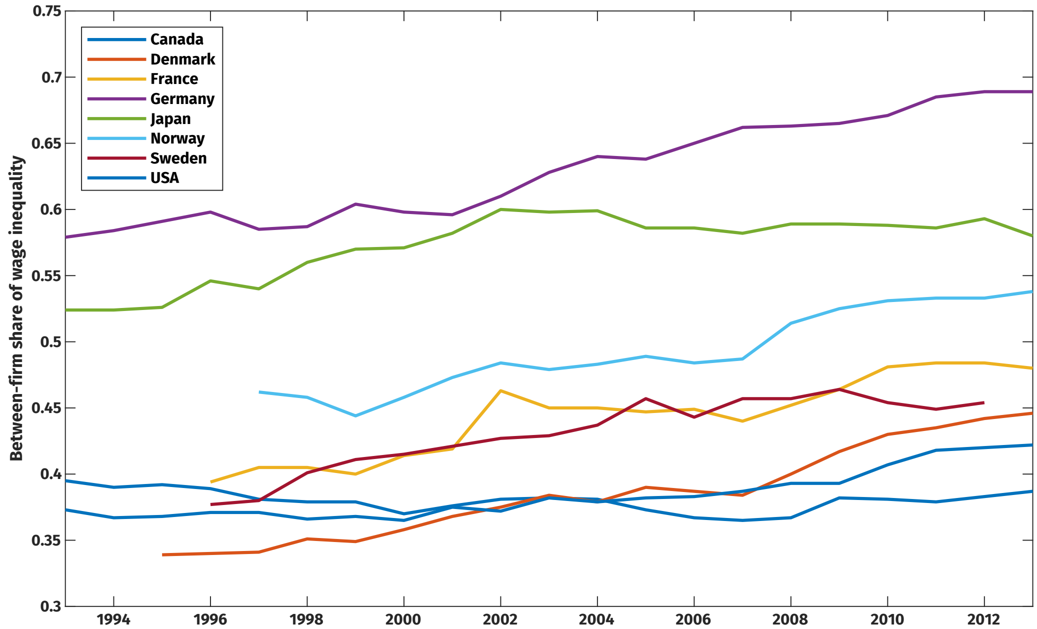


Figure A.1: Cross-country evidence on the between-firm share of wage inequality

Notes. This figure reports the evolution of the between-firm share of wage inequality for a set of OECD economies. The data are sourced from Tomaskovic-Devey *et al.* (2020).

Tomaskovic-Devey *et al.* (2020). While the levels are of limited comparability due to variation in the measure of earnings across different countries (e.g., hourly vs. daily vs. monthly earnings), one can observe a consistent upward trend for almost all countries.

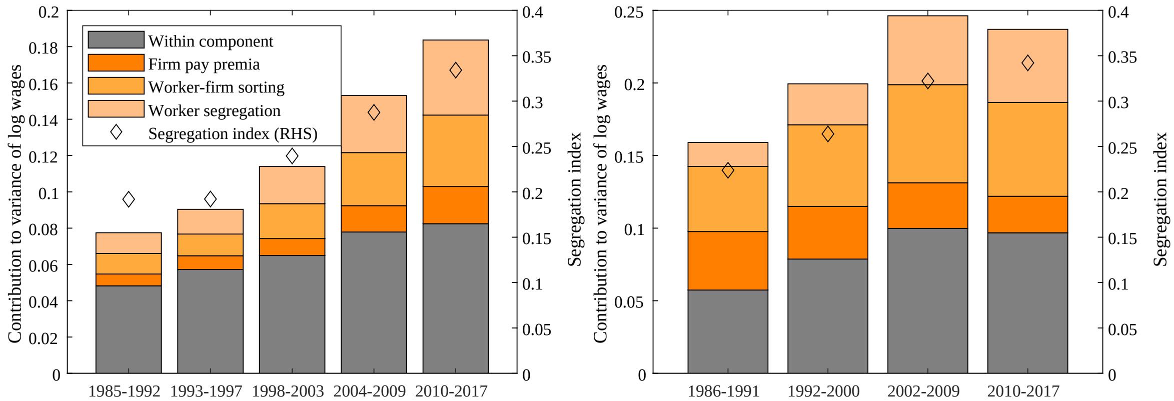
Interestingly, even in a country like France, where total wage inequality has broadly flatlined over the past few decades, the between-firm component tends to have increased due to rising sorting and segregation, whereas within-firm inequality has declined for a variety of reasons (Babet *et al.*, 2022).

A.4.2 AKM-based analyses

The log-linear structure of the AKM model facilitates simple diagnostics to decompose the variance of log wages into easily interpretable components. Given estimated worker and employer fixed effects, α_i and ψ_j, we can compute the following variance decomposition, as in Song *et al.* (2019).

$$\mathrm{Var}(w_{it}) = \underbrace{\mathrm{Var}(\alpha_i - \bar{\alpha}_j) + \mathrm{Var}(\epsilon_{it})}_{\text{within-component}} + \underbrace{\mathrm{Var}(\psi_j)) + 2\mathrm{Cov}(\bar{\alpha}_j, \psi_j) + \mathrm{Var}(\bar{\alpha}_j)}_{\text{between-components}}, \tag{A.3}$$

where $\bar{\alpha}_j$ is the average worker FE in firm j. A common interpretation then relates $\mathrm{Var}(\psi_k)$ to firm/cluster-specific pay premia, $\mathrm{Cov}(\bar{\alpha}_j, \psi_j)$ to worker-firm sorting, and $\mathrm{Var}(\bar{\alpha}_j)$ to worker segregation. It warrants renewed emphasis, though, that through the lens of a structural model with



(a) Germany (b) Portugal

Figure A.2: AKM-based wage variance decomposition

Notes. These panels report the components of the variance of log wages, following equation (A.3), based on the estimation of the AKM model described in equation (A.2). The left panel is for Germany, the right panel for Portugal (where I only use four sample periods to better span missing years in 1990 and 2001).

coworker complementarities these components are not neatly separable.

The left panel in Figure A.2 performs this decomposition for the German economy on a period-by-period basis. We see that all three between-components have risen over time. Moreover, the Kremer-Maskin segregation index (Kremer and Maskin, 1996), which indicates the share of the variance of worker FEs that occurs between rather than within establishments – and, thus, represents an alternative measure of coworker sorting – has risen over time.[A.3]

For comparison purposes, the right-hand panel reports the variance decomposition for the Portuguese economy. A key takeaway is that even though the between-firm share of the variance has remained roughly constant since the late 1990s, both $\mathrm{Cov}(\bar{\alpha}_j, \psi_j)$ and $\mathrm{Var}(\bar{\alpha}_j)$ have increased, as has the segregation index – similar to the German economy. In the absence of the firm-specific pay premia, therefore, the sorting and segregation components would have pushed up between-firm inequality, in line with what we observe for Germany and many other advanced economies. Silva *et al.*'s (2022) analysis of the Portuguese economy come to a similar conclusion.

A.4.3 Robustness: rising between-share of inequality in Germany

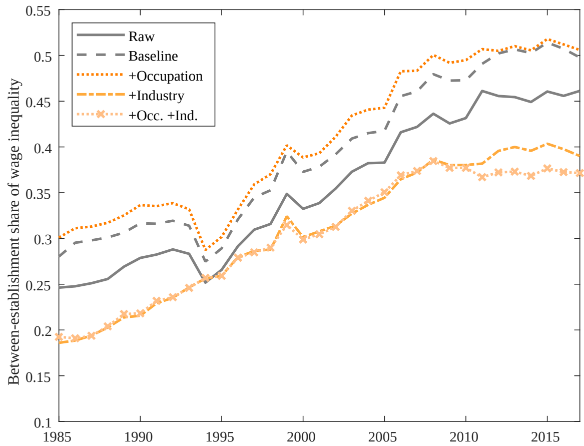
This section examines how robust the increase in the between-establishment share of the variance of log wages in Germany is to alternative measures of individual wages. Figure A.3 provides a graphical summary and indicates that controlling for different sets of covariates does not alter the main conclusion of a rising between-component.

[A.3] Interestingly, if one performs the same decomposition separately by major sector groups, the upward trend is most pronounced for knowledge-intensive services. In general, the between share of the variance of log wages tends to be higher for service sectors.

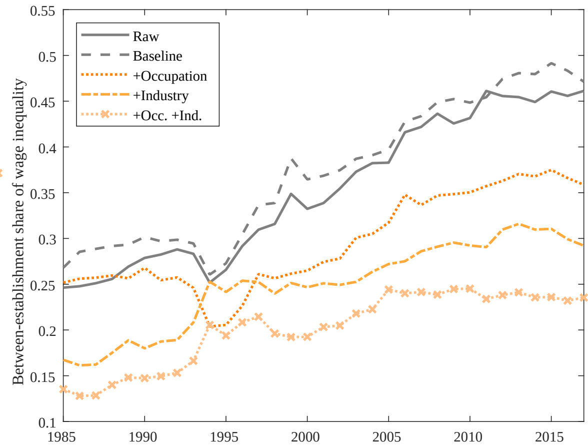
Starting with an explanation of the different wage measures considered, the solid line in Figure A.3, in either panel, depicts the standard between-within establishment decomposition of the variance of log raw wages, $\ln(\tilde{w})_{it}$. In the left panel, the dashed line in the same colors swaps in the baseline measure of wages fed into the model. This measure was constructed from a regression of log (raw) wages onto an index of time-varying characteristics, X_{it}, which includes age, tenure and year dummies, as well as person fixed effects (FEs). The residual measure subtracts from $\ln(\tilde{w}_{it})$ the component explained by X_{it}. Next, the dotted, dashed-dotted, and dotted-crossed lines also remove, respectively, (2-digit) occupation FEs, (2-digit) industry FEs, or both. Throughout, by including the worker FEs in the "residual" component, it is ensured that we do not omit variation in time-invariant individual earnings that is also present – indeed, central to –the structural model. For robustness, the right panel in the figure repeats the same exercise but without including person FEs in the regression in the first place.

Several observations stand out. First, and reassuringly, the rise in the between-establishment share of wage inequality is highly robust across all specifications. While the *level* of this share varies, the increase over time is quite uniform, ranging from 0.179 to 0.2175 (1985-2017). Second, and unsurprisingly, both the level and the percentage point increase in the between-establishment share are smaller in magnitude when worker FEs are not accounted for. Next, turn to the role of occupations and industries. Under the specification that minimizes the role of residual wage inequality, namely when controlling for both occupation and industry FEs, the between share rises from 0.1353 to 0.2354, with most of the increase occurring up until the Global Financial Crisis. That the increase in the between share is smaller, in percentage point terms, when taking out occupation FEs is consistent with two important ideas. The first is that a large fraction of between-firm inequality manifests across industries (Haltiwanger and Spletzer, 2020). The second is that some of the rise in between-establishment/firm inequality is due to occupational outsourcing. Goldschmidt and Schmieder (2017), in particular, document that outsourcing of cleaning, security, and logistics services accounts for around 9% of the increase in German wage inequality since the 1980s.[A.4] Moreover, through the lens of an AKM model – considered in the preceding Appendix section – outsourcing manifests in a reduction in all three of the between components. Interestingly, the *percentage* increase in the between-establishment share, after controlling for occupation and industry effects, is very similar to what is found in the baseline scenario, at around 75%. Overall, additional controls can account for some but not the majority of the rise in the between-establishment share of wage inequality.

[A.4] Bilal and Lhuillier (2021) argue, in the French context, that outsourcing leads to rising labor market sorting but relatively stable wage inequality, due to an offsetting general equilibrium mechanism associated with pro-competitive effects of contractors at the bottom of the job ladder.



(a) With worker fixed effects



(b) Without worker fixed effects

Figure A.3: Rising between-establishment share of wage inequality in Germany

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

A.4.4 Sorting patterns

In both Germany and Portugal, coworker sorting has risen over time – as reflected also in the worker segregation component just illustrated. Table A.1 indicates, for five sample periods, the correlation of a worker's type with the average type of their coworker. For robustness purposes, 2×2 specifications are indicated. One dimension refers to the method by which workers are ranked, based on AKM fixed effects or non-parametrically (see Appendix A.3.) The second dimension is whether the ranking is done on an economy-wide basis, which is denoted "definition 1" and represents the baseline, or whether workers are ranked within their respective two-digit occupation. The table shows that under any specification, coworker sorting has become more pronounced, though the magnitudes vary (the total rise from first to last period ranges from 0.057 in the second column to 0.119 in the first column).[A.5]

Similar to our analysis of the theoretical model, we can also construct a more disaggregated picture of coworker sorting in the data. Figure A.4 depicts for any worker percentile rank, the quality of their coworkers, in an early and a later year, for both Germany and Portugal. The worker type is based on the AKM method. Coworker quality is measured by the average percentile rank of a person's coworkers (i.e., in the same employer-year cell).[A.6]

For both countries, the patterns displayed conform to the theoretical predictions in Corollary 2,

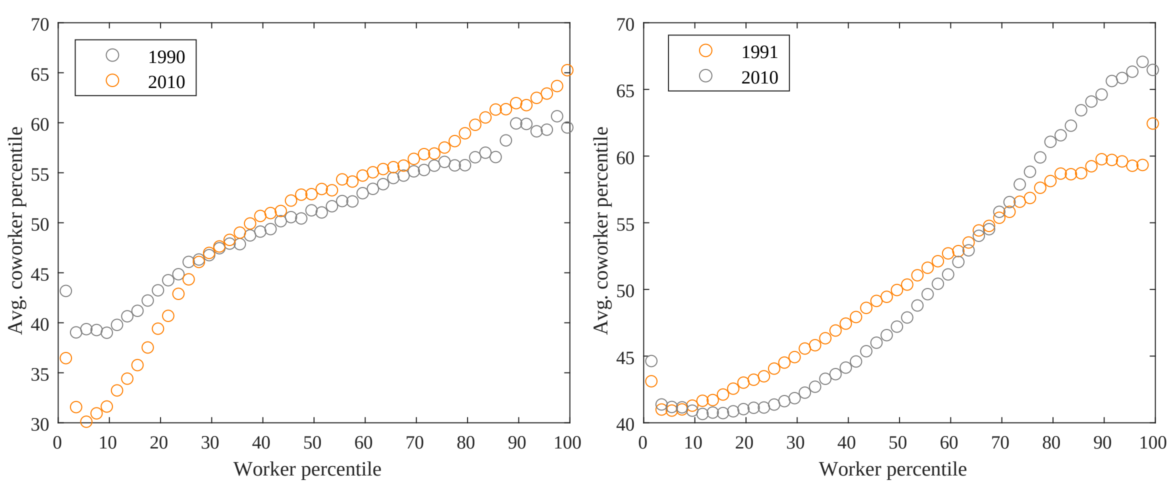
[A.5]Complementing the statistics reported in the main text, for Portugal the correlation of own rank and average coworker rank, using the AKM method and definition 1, has risen from 0.460 in 1991 to 0.534 in 2010.

[A.6]The results look very similar when using, instead, the percentile rank of their average coworker (i.e., when the "representative coworkers" is re-ranked to follow a uniform distribution. Results are available upon request.

	AKM-based		Non-parametric	
Sample period	Def. 1	Def. 2	Def. 1	Def. 2
1985-1992	0.458	0.393	0.590	0.487
1993-1997	0.435	0.357	0.580	0.489
1998-2003	0.489	0.385	0.632	0.519
2004-2009	0.542	0.433	0.663	0.546
2010-2017	0.577	0.450	0.695	0.561

Table A.1: Coworker sorting in Germany over time

Notes. This table reports the correlation of a worker's type with the average type of their coworkers, separately for five sample periods, using the German data. The moments are person-year weighted.



(a) Germany (b) Portugal

Figure A.4: Nonlinear sorting patterns in the data

Notes. This figure plots, for any percentile of the worker fixed-effect distribution, the average quality of coworkers. To enhance visual clarity, workers are first binned into 50 cells and then the coworker quality is computed for each cell. For practical reasons, this is done manually rather than using the popular "binscatter" module in Stata, but this makes no practical difference. Regarding the choice of years, for Portugal I use 1991 instead of 1990, as the latter is missing.

which was illustrated in Figures 3. In particular, we see an *increase* in the coworker quality for the highest-quality workers, and a *decrease* for those below (roughly) the median, while little changes for those in the center of the quality distribution. Reassuringly, the results are highly comparable across both countries, which is noteworthy also since the earnings data available for Portugal are uncensored.

A.4.5 Evidence on coworker wage complementarities

A.4.5.1 Complementarities: non-parametric vs. AKM-based worker rankings The main section reports estimation results for the regression in equation (35) based on worker and coworker quality types constructed based on AKM wage regressions. The first main column in Table A.2 repeats the point estimate of the coefficient on the interaction term, β_3. The second column shows the estimated coefficient when worker types are instead estiamted non-parametrically using the method proposed by Hagedorn *et al.* (2017) and described in Appendix A.3.2. Reassuringly, magnitudes and trend look very similar.

Sample period	AKM-based worker types	Non-parametrically estimated worker types
1985-1992	0.115	0.105
1993-1997	0.107	0.108
1998-2003	0.161	0.146
2004-2009	0.205	0.158
2010-2017	0.215	0.190

Table A.2: Empirical estimates of coworker wage complementarities: alternative worker

Notes. This table reports the point estimates for the coefficient β_3 when regression (35) is separately run for five sample periods and worker quality types are alternatively constructed based on AKM wage regressions or non-parametrically. Magnitudes are reported in terms of hourly wages (using a scaling factor of 7.5 working hours per day).

A.4.5.2 Evolution of complementarities in Portugal Figure A.5 shows that in Portugal, too, coworker wage complementarities have strengthened over time. Is is notable that levels are somewhat lower than in Germany – just as we would expect from the fact that PRT has a less “complex” economy. Moreover, the increase is more modest, which again is consistent with the observation that coworker sorting has increased only modestly in Portugal.

A.4.5.3 Complementarities across occupational groups in Germany Figure A.6 reports an analogue to Figure 6, likewise computed for the sample 2010-2017. In those years, the LIAB data include the KldB-2010 classification of occupations, which includes a grouping of occupations by the complexity of their tasks. When this grouping is used, the results are very similar to those obtained for Portugal, wherein complementarities are stronger when task requirements are more complex.

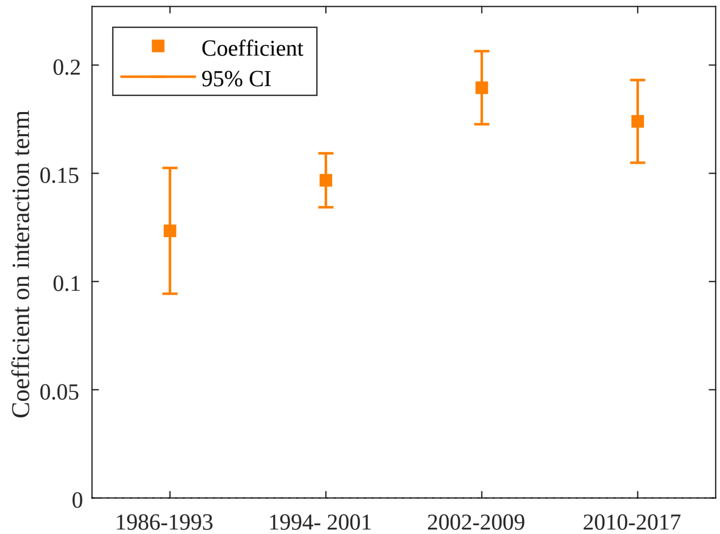


Figure A.5: Empirical estimates of coworker (wage) complementarities in PRT

Notes. This figure is the analogue of 4, but estimated using the Portuguese data. Standard errors are clustered at the firm level.

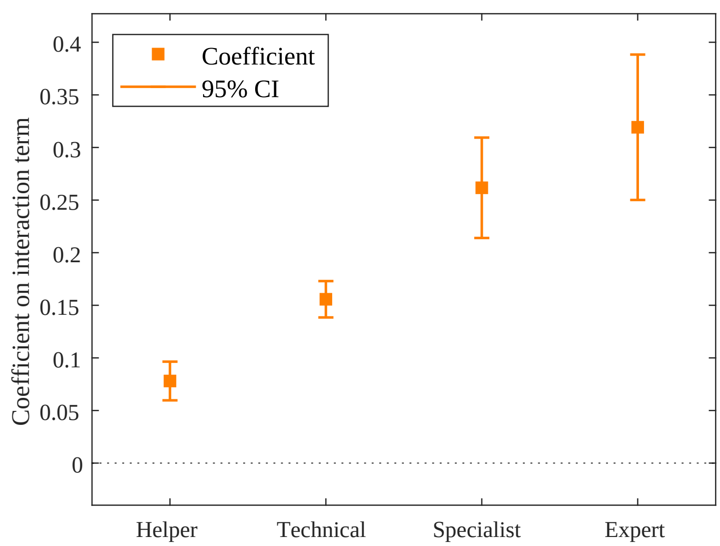
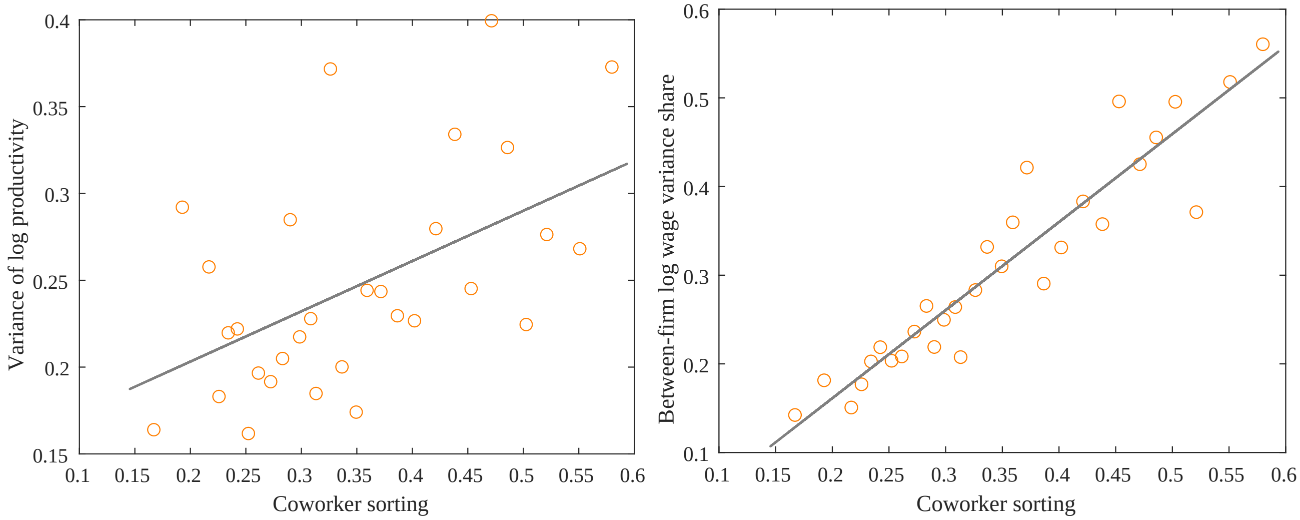


Figure A.6: Complementarities by task complexity: Germany

Notes. This figure is the analogue of 6, but estimated using German data. Standard errors are clustered at the employer level. Workers are grouped by the fifth digit of the KldB-2010 occupational classification. Requirement level 4 ("Experte") refers to highly complex activities; level 3 ("Spezialist") to complex specialist activities; level 2 ("Fachkraft") to technically oriented tasks; and level 1 ("Helfer") to "helping and training activities. For details, see the webpage of the Bundesagentur fuer Arbeit.



(a) Productivity dispersion

(b) Between-firm share of wage dispersion

Figure A.7: Coworker sorting is associated with greater between-firm inequality

Notes. This figure plots the variance of log wages (left panel) and the between-firm share of the variance of log wages (right panel) against the coworker sorting correlation. The unit of observation is a 4-digit industry, these are binned into 30 cells.

A.4.6 Cross-industry variation

Figure A.7 expands on the cross-industry findings from Section 3.3.2. It shows that industries characterized by relatively high degrees of coworker sorting also exhibit greater firm-level dispersion in productivity, and that a greater share of the variance of log wages occurs between firms.

B Theory

B.1 Team production

The derivations are lengthy but very standard in the trade literature. I particularly lean on the material in Eaton *et al.* (2016) and Allen (2019).

B.1.1 Derivation of equation (8)

Start with equation (1), repeated here for convenience:

$$Y = \left(\int_0^1 q(\tau)^{\frac{\eta-1}{\eta}} d\tau \right)^{\frac{\eta}{\eta-1}},$$

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 (7), rearranged as

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

and simplifying, we obtain

$$\left(\int_0^1 \left(q(\tau)\tilde{\lambda}\right)^{\frac{\eta-1}{\eta}} dQ(\tau)^{\frac{1}{\eta}} \left(\frac{1}{Y}\right)^{\frac{1}{\eta-1}} \left(\frac{1}{\lambda}\right)^{\frac{\eta-1}{\eta}} d\tau \right)^{\frac{\eta}{\eta-1}} = Y\lambda^{\frac{1}{\eta-1}}. \tag{B.1}$$

Now use $Q(\tau) = \tilde{\lambda}(\tau)q(\tau)$, bring the terms independent of τ outside the integral, and cancel exponents. Then

$$\int_{\mathcal{T}} Q(\tau) d\tau = \lambda Y. \tag{B.2}$$

B.1.2 Proof of Lemma 1

PART (I).

Step 1 is to derive the distribution of shadow costs of worker i providing task τ, $\lambda_i(\tau)$. Since the efficiency draws are independently and identically distributed, the probability of i producing that task at a shadow price less than p is the same for all $\tau \in \mathcal{T}$. Defining

$$G_i(p) := \Pr(\lambda_i(\tau) \leq p), \tag{B.3}$$

the properties of the Fréchet distribution together with equation (13) imply

$$
\begin{aligned}
G_{in}(p) &= \Pr\left\{\frac{\lambda_i^L}{a_1 z_i(\tau)} \leq p\right\} \\
&= \Pr\left\{\frac{\lambda_i^L}{a_1 p} \leq z_i(\tau)\right\} \\
&= 1 - \Pr\left\{z_i(\tau) \leq \frac{\lambda_i^L}{a_1 p}\right\} \\
&= 1 - \exp\left(-\left(\frac{\lambda_i^L}{\iota a_1 p x_i}\right)^{-\frac{1}{\chi}}\right).
\end{aligned}
$$

For step 2 we look at the probability that the firm can obtain a task $\tau \in \mathcal{T}$ for a shadow cost of less than p,

$$
G(p) := \Pr\{\tilde{\lambda}(\tau) \leq p\}. \tag{B.4}
$$

Using equation (11) and basic tools in probability:

$$
\begin{aligned}
G_n(p) &= \Pr\left\{\min_i \tilde{\lambda}_i(\tau) \leq p\right\} \\
&= 1 - \Pr\left\{\min_i \tilde{\lambda}_i(\tau) \geq p\right\} \\
&= 1 - \Pr\left\{\cap_{i\in\mathcal{S}}\left(\tilde{\lambda}_i(\tau) \geq p\right)\right\} \\
&= 1 - \Pi_{i\in\mathcal{S}}\left(1 - G_i(p)\right)
\end{aligned}
$$

The underlying intuition is that the lowest shadow price available is weakly lower than p unless the shadow price of producing that task is greater than p for each worker, so that the distribution $G(p)$ is the complement of the probability that for every $i \in \mathcal{S}$ the shadow cost of providing the task is greater than p.

In step 3, substitute in the last line above using the expression for $G_i(p)$ derived in the step 1:

$$\begin{aligned} G(p) &= 1 - \Pi_{i\in\mathcal{S}} \exp\left(\left(\frac{\lambda_i^L}{\iota a_1 p x_i}\right)^{-\frac{1}{\chi}}\right) \\ &= 1 - \exp\left(-(\iota p)^{\frac{1}{\chi}} \sum_{i\in\mathcal{S}} \left(\frac{\lambda_i^L}{a_1 x_i}\right)^{-\frac{1}{\chi}}\right) \\ &= 1 - \exp(-(\iota p)^{\frac{1}{\chi}}\Phi), \end{aligned}$$

where $\Phi = \sum_{i\in\mathcal{S}} \left(\frac{\lambda_i^L}{a_1 x_i}\right)^{-\frac{1}{\chi}}$

For the final step, consider equation (9), and substitute in the distribution of $\tilde{\lambda}(\tau)$ just derived. Then

$$\begin{aligned} \lambda^{1-\eta} &= \int_0^\infty p^{1-\eta} dG(p) \\ &= \int_0^\infty p^{1-\eta}\left(\frac{d}{dp}\left(1 - \exp(-(\iota p)^{\frac{1}{\chi}}\Phi)\right)\right)dp, \\ &= \frac{1}{\chi}\iota^{\frac{1}{\chi}}\Phi \int_0^\infty p^{\frac{1}{\chi}-\eta}\left(\exp(-(\iota p)^{\frac{1}{\chi}}\Phi)\right)dp. \end{aligned}$$

Now use a change of variables, with $m = (\iota p)^{1/\chi}\Phi$, so that $p = \iota^{-1}\left(\frac{m}{\Phi}\right)^{\chi}$ and $dp = \iota^{-1}\chi\left(\frac{m}{\Phi}\right)^{1-\chi}\frac{1}{\Phi}dm$. Performing the integration by substitution,

$$\begin{aligned} \lambda^{1-\eta} &= \frac{1}{\chi}\iota^{\frac{1}{\chi}}\Phi \int_0^\infty \left(\iota^{-1}\left(\frac{m}{\Phi}\right)^{\chi}\right)^{\frac{1}{\chi}-\eta} \exp(-m)\iota^{-1}\chi\left(\frac{m}{\Phi}\right)^{1-\chi}\frac{1}{\Phi}dm \\ &= \iota^{\eta-1}\Phi^{-\chi(1-\eta)} \int_0^\infty m^{\chi(\eta-1)} \exp(-m)dm \\ &= \iota^{\eta-1}\Phi^{-\chi(1-\eta)}\Gamma(1+\chi-\eta\chi) \end{aligned}$$

Finally, since $\iota = \Gamma(1+\chi-\eta\chi)^{\frac{1}{1-\eta}}$, we obtain

$$\begin{aligned} \lambda &= (\Phi_i)^{-\chi} \\ &= \left(\sum_{i\in\mathcal{S}}\left(\frac{a_1 x_i}{\lambda_i^L}\right)^{1/\chi}\right)^{-\chi}, \end{aligned}$$

as stated in the main text.

PART (II). What is the probability that i has the lowest shadow cost of providing a task $\tau \in \mathcal{T}$?

Because productivity draws are iid, and since tasks are on a continuum, by the law of large numbers this probability will be equal to the fraction of tasks that i produces according to the optimal plan. Define

$$\begin{aligned}
\pi_i &:= \Pr\left\{\tilde{\lambda}_i(\tau) \leq \min_{k\in\mathcal{S}\setminus i} \tilde{\lambda}_k(\tau)\right\} \\
&= \Pr\left\{\min_{k\in\mathcal{S}\setminus i} \geq p\right\} dG_i(p) \\
&= \int_0^\infty \Pr\left\{\cap_{k\in\mathcal{S}\setminus i}\, \tilde{\lambda}_k(\tau) \geq p\right\} dG_i(p) \\
&= \int_0^\infty \Pi_{k\in\mathcal{S}\setminus i}(1 - G_k(p)) dG_i(p)
\end{aligned}$$

Now substitute for the distribution of shadow costs to obtain:

$$\begin{aligned}
\pi_i &= \int_0^\infty \left[1 - \exp\left(-\left(\frac{\lambda_k^L}{\iota a_1 p x_k}\right)^{\frac{1}{\chi}}\right)\right] dp\left[\frac{d}{dp}\left(\left[1 - \exp\left(-\left(\frac{\lambda_i^L}{\iota a_1 p x_i}\right)^{-\frac{1}{\chi}}\right)\right]\right.\right] \\
&= \left(\frac{\lambda_i^L}{\iota a_1 x_i}\right)^{-\frac{1}{\chi}} \int_0^\infty \left(\frac{1}{\chi}\iota^{\frac{1}{\chi}} p^{\frac{1}{\chi}-1}\right)\left(\exp(-(p\iota)^{\frac{1}{\chi}}\Phi_n)\right) dp \\
&= \frac{(a_1 x_i)^{\frac{1}{\chi}} \left(\lambda_i^L\right)^{-\frac{1}{\chi}}}{\Phi_n}\left[-\exp(-(\iota p)^{\frac{1}{\chi}}\Phi_n)\right]_0^\infty \\
&= \frac{(a_1 x_i)^{\frac{1}{\chi}} \left(\lambda_i^L\right)^{-\frac{1}{\chi}}}{\Phi_n},
\end{aligned}$$

where the last equality follows given the expression for the cdf $G_n(p)$ derived above.

Lastly, we need to show that π_i is not only the fraction of *tasks* that i produces; it is also the fraction of the *value* of tasks. This result obtains because under the Fréchet distribution, the distribution of shadow prices of tasks that worker n actually uses from any team member will be the same. Here is a proof, adapted from Allen (2019).

The probability that the shadow cost of i providing a task τ is lower than $\tilde{p}$, conditional on that cost being the lowest for i among all team members, is

$$\Pr\left\{\tilde{\lambda}_i(\tau) \leq \tilde{p} | \tilde{\lambda}_i(\tau) \leq \min_{k\in\mathcal{S}\setminus i} \tilde{\lambda}_k(\tau)\right\} = \frac{1}{\pi_i}\int_0^{\tilde{p}} \Pr\left\{\min_{k\in\mathcal{S}\setminus i} p_k(\tau) \geq p\right\} dG_i(p),$$

were the first term on the right-hand side is the inverse probability that i has the lowest shadow cost of producing a given task, and the second term is the probability that the firm can be provided with a a task at a shadow cost lower than $\tilde{p}$ by a team member other than i.

Then using the same logic as in the derivation of π_i, we find that this is equal to

$$\begin{aligned}
\Pr\left\{\tilde{\lambda}_i(\tau) \leq \tilde{p} | \tilde{\lambda}_i(\tau) \leq \min_{k\in\mathcal{S}\setminus i} \tilde{\lambda}_k(\tau)\right\} &= \int_0^{\tilde{p}} \Pi_{k\in\mathcal{S}\setminus i}\Big(1 - G_k(p)\Big) dG_i(p), \\
&= \frac{1}{\pi_i}\frac{(a_1 x_i)^{\frac{1}{\chi}}\left(\lambda_i^L\right)^{-\frac{1}{\chi}}}{\Phi}\Big[-\exp(-(\iota p)^{\frac{1}{\chi}}\Phi)\Big]_0^{\tilde{p}} \\
&= \frac{1}{\pi_i}\frac{(a_1 x_i)^{\frac{1}{\chi}}\left(\lambda_i^L\right)^{-\frac{1}{\chi}}}{\Phi}\Big[[-\exp(-(\iota\tilde{p})^{\frac{1}{\chi}}\Phi) - (-\exp(0))\Big] \\
&= \frac{1}{\pi_i}\pi_i\Big(1 - \exp(-(\iota\tilde{p})^{\frac{1}{\chi}}\Phi)\Big) \\
&= G(\tilde{p}),
\end{aligned}$$

which is independent of i.

Intuitively, the planner makes team members with an absolute advantage provide a greater range of tasks exactly up to the point where the distribution of shadow costs associated with producing tasks is the same as the overall distribution of shadow costs.

PART (III). The final result immediately follows since

$$Q_i = \pi_i\bigg(\int_{\mathcal{T}} \tilde{\lambda}(\tau) q(\tau) d\tau\bigg) = \pi_i \int_{\mathcal{T}} Q(\tau) d\tau = \pi_i Q.$$

B.1.3 Proposition 1

When $\lambda = 1$, then part (i) of Lemma 1 implies that

$$1 = \sum_{i\in\mathcal{S}} \left(\frac{a_1 x_i}{\lambda_i^L}\right)^{\frac{1}{\chi}}, \tag{B.5}$$

Hence, from part (ii) of Lemma 1,

$$\pi_i = ((a_1 x_i)/\lambda_i^L)^{\frac{1}{\chi}}.$$

Now, integrating over equation (12) and using the time constraint (3) implies that

$$\int_{\mathcal{T}} \lambda_i(\tau) q_i(\tau) = \lambda_i^L.$$

Since $\tilde{\lambda}(\tau) = \lambda_i(\tau)$ when $y_i(\tau) > 0$, this also means that

$$Q_i = \lambda_i^L,$$

which says that the shadow value of worker i's time is equal to the shadow value of all tasks produced by that worker. Hence, from part (iii) of Lemma 1, and since $\lambda = 1$,

$$\begin{aligned} \lambda_i^L &= \pi_i Q, \\ \Leftrightarrow \lambda_i^L &= \pi_i Y, \\ \Leftrightarrow \left(\frac{a_1 x_i}{\lambda_i^L}\right)^{\frac{1}{\chi}} &= \lambda_i^L / Y. \end{aligned}$$

Substituting this expression into equation (B.5) and rearranging for Y yields

$$N^{1+\chi}\left(\frac{1}{N}\sum_{i=1}^{N}(a_1 x_i)^{\frac{1}{1+\chi}}\right)^{1+\chi}$$

B.2 Team hiring

B.2.1 Population dynamics

For any type x, the measure of unemployment satisfies

$$\begin{aligned} \delta\left(\int d_m(x,\tilde{y})d\tilde{y} + \int d_m(x,\tilde{y},\tilde{x}')d\tilde{x}'d\tilde{y}\right) = d_u(x)M_u\Bigg(\int \frac{d_v(\tilde{y})}{V}h(x,\tilde{y}) \\ + \int \frac{d_m(\tilde{x}',\tilde{y})}{V}h(x,(\tilde{y},\tilde{x}'))d\tilde{x}'d\tilde{y}\Bigg). \end{aligned} \tag{B.6}$$

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, y, the measure of one-worker matches follows

$$\begin{aligned} d_m(x,y)\left(\delta + M_f\int \frac{d_u(\tilde{x}')}{U}h(\tilde{x}',(y,x))d\tilde{x}'\right) = d_u(x)M_u\frac{d_v(y)}{V}h(x,y) \\ + \delta\int d_m(x,y,\tilde{x}')d\tilde{x}'. \end{aligned} \tag{B.7}$$

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

Finally, for all (x, y, x'),[B.1]

$$2\delta d_m(x, y, x') = d_u(x) M_u \frac{d_m(x', y)}{V} h(x, (y, x')) + d_u(x') M_u \frac{d_m(x, y)}{V} h(x', (y, x)). \tag{B.8}$$

The economic intuition parallels the aforementioned reasoning.

B.2.2 Derivation of surplus equations

B.2.2.1 $S(x, y)$ We start with the definition of $S(x, y)$, repeated here for convenience:

$$S(x, y) = \Omega(x, y) - V_u(x) - V_v(y). \tag{B.9}$$

Consider equation (31). The joint value of production with x and y. The term in $[\cdot]$ obviously corresponds to $S(x, y)$. Rearranging the surplus sharing equations furthermore implies that[B.2]

$$\big(- \Omega(x, y) + V_e(x, (y, x')) + V_p(x, y, x') \big) = (1 - \omega) S(x', (y, x))$$

In words, the joint value of y and x from being in a (x, y, x') coalition minus their joint outside option is equal to $(1 - \omega)$ of the marginal surplus of adding x'.

Hence, we can write

$$\begin{aligned} \rho \Omega(x, y) = {} & f(x, y) - \delta S(x, y) \\ & + M_f (1 - \omega) \int \frac{d_u(\tilde{x}')}{U} S(\tilde{x}', (y, x))^+ d\tilde{x}' \end{aligned}$$

Substituting this expression into equation (B.9) multiplied by ρ yields

$$\big(\rho + \delta\big) S(x, y) = f(x, y) - \rho \big(V_u(x) + V_v(y) \big) + M_f (1 - \omega) \int \frac{d_u(\tilde{x}')}{U} S(\tilde{x}', (y, x))^+ d\tilde{x}'. \tag{B.10}$$

[B.1] From an accounting perspective, one of the preceding three equations is redundant given because the adding-up constraint represented by equation (20) must hold and the distribution of worker types is exogenous. Similarly, $d_v(y)$ can be backed out from equation (21) given $d_f(y)$.

[B.2] To see this, start from the surplus sharing equation,

$$(1 - \omega) \big(V_e(x', (y, x)) - V_u(x') \big) = \omega \big(- \Omega(y, x) + V_p(y, x, x') + V_e(x, (y, x')) \big).$$

Adding on both sides $(1 - \omega)\big(- \Omega(x, y) + V_p(x, y, x') + V_e(x, (y, x')) \big)$ and using the definition of surpluses yields the above.

B.2.2.2 $S(x, y, z')$ The first equation we need is a relationship between the *marginal* surplus and the *total* join value. Start with the definition of $S(x, (y, x'))$, substitute using the sharing rule (23), and simplify:

$$\begin{aligned}
S(x,(y,x')) &= \Omega(x,y,x') - \underbrace{\Big(V_e(x',y) + V_p(x',y)\Big)}_{\Omega(x',y)} - V_u(x). \\
&= \Omega(x,y,x') - (V_u(x') + \omega S(x',y) + V_v(y) + (1-\omega)S(x',y)) - V_u(x) \\
&= \Omega(x,y,x') - S(x',y) - \Big(V_u(x) + V_u(x') + V_v(y)\Big)
\end{aligned}$$

Hence,

$$\Omega(x,y,x') = S(x,(y,x')) + S(x',y) + V_u(x) + V_u(x') + V_v(y). \tag{B.11}$$

and analogously

$$S(x',(y,x)) = \Omega(x,y,x') - \Big(S(x,y) + V_u(x) + V_u(x') + V_v(y)\Big). \tag{B.12}$$

Second, consider the HJB equation for $\Omega(x, y, x')$, equation (30). Substitute for the the terms in brackets using the the definitions of marginal surplus to get a second expression for the latter:

$$\begin{aligned}
\rho\Omega(x,y,x') &= f(x,y,x') + \delta\Big[-S(x,(y,x') - S(x',(y,x)\Big] \\
\Leftrightarrow \delta S(x,(y,x')) &= f(x,y,x') - \rho\Omega(x,y,x') - \delta S(x',(y,x))
\end{aligned}$$

Substituting for $S(x', (y, x))$ using equation (B.12) and for $\Omega(x, y, x')$ using equation (B.11), and collecting terms yields

$$\begin{aligned}
\delta S(x,(y,x')) &= f(x,y,x') - \rho\Omega(x,y,x') - \delta\Big[\Omega(x,y,x') - \Big(S(x,y) + V_u(x) + V_u(x') + V_v(y)\Big)\Big] \\
&= f(x,y,x') - (\rho+\delta)\Omega(x,y,x') + \delta\big[S(x,y) + V_u(x) + V_u(x') + V_v(y)\big] \\
&= f(x,y,x') - (\rho+\delta)\Big(S(x,(y,x')) + S(x',y) + V_u(x) + V_u(x') + V_v(y)\Big) \\
&\quad + \delta\big[S(x,y) + V_u(x) + V_u(x') + V_v(y)\big]
\end{aligned}$$

Simplifying yields

$$S(x,(y,x'))\big(\rho+2\delta\big) = f(x,y,x') - \rho\big(V_u(x)+V_u(x')+V_v(y)\big) + \delta S(x,y) - (\rho+\delta)S(x',y). \tag{B.13}$$

B.2.3 Derivation of the wage function

The value of employment for worker x at a firm y with coworker x' is

$$\rho V_e(x,(y,x')) = w(x,(y,x')) - 2\delta\omega S(x,(y,x')) + \delta\omega S(x,y).$$

Combining with the surplus sharing rule (27), repeated here for convenience,

$$V_e(x,(y,x')) = V_u(x) + \omega S(x,(y,x'))$$

yields

$$w(x,(y,x')) = \rho V_u(x) + (\rho+2\delta)\omega S(x,(y,x')) - \delta\omega S(x,y) \tag{B.14}$$

We can simplify this expression further. First, substitute for $(\rho+2\delta)S(x,(y,x'))$ from equation (B.13) to get

$$\begin{aligned} w(x,(y,x')) &= \rho V_u(x) + \omega\Big[f(x,y,x') - \rho\big(V_u(x)+V_u(x')+V_v(y)\big) \\ &\quad + \delta S(x,y) - (\rho+\delta)S(x',y)\Big] - \delta\omega S(x,y), \\ &= \omega f(x,y,x') + (1-\omega)\rho V_u(x) - \omega\rho\big(V_u(x')+V_v(y)\big) \\ &\quad - \omega(\rho+\delta)S(x',y), \end{aligned}$$

Next, substitute for $S(x',y)$ from equation (B.10) to obtain

$$w(x,(y,x')) = \omega f(x,y,x') + (1-\omega)\rho V_u(x) - \omega\rho\big(V_u(x')+V_v(y)\big) \tag{B.15}$$

$$- \omega\left[f(x,y) - \rho\big(V_u(x)+V_v(y)\big) + M_f(1-\omega)\int \frac{d_u(\tilde{x}')}{U} S(\tilde{x}',(y,x))^+ d\tilde{x}'\right] \tag{B.16}$$

$$= \omega\big(f(x,y,x') - f(x',y)\big) + (1-\omega)\rho V_u(x) \tag{B.17}$$

$$- \omega(1-\omega)M_f \int \frac{d_u(\tilde{x}'')}{U} S(\tilde{x}'',(y,x'))^+ d\tilde{x}''. \tag{B.18}$$

B.2.4 Monotonicity Lemma, with proof

The following lemma establishes that Hagedorn *et al.*'s (2017) non-parametric ranking algorithm extends to the present environment.

Lemma 2. *(i) The value of unemployment $V_u(x)$ is monotonically increasing in x, and (ii) so is the wage function $w(x,(y,x'))$.*

B.2.4.1 Proof of Part (i): $V_u(x)$ is monotonically increasing in x We start with the equilibrium equations. The value of unemployment for a worker type x can be written as[B.3]

$$\rho V_u(x) = \omega M_u \int_{M(x)} \left[\frac{d_v(\tilde{y})}{V} S(x,\tilde{y}) + \int_{M(x,\tilde{y})} \frac{d_m(\tilde{z},\tilde{y})}{V} S(x,(\tilde{y},\tilde{z}))d\tilde{z} \right] d\tilde{y}. \tag{B.19}$$

For the proof, it is convenient to denote $M(x) = \{y : S(x,y) \geq 0\}$ as the matching set which consists of unmatched firms of type y that are accepted by worker type x and vice versa[B.4]; and $M(x,y) = \{x' : S(x',(y,x)) \geq 0\}$ is the matching set consisting of worker types x' that are accepted by a coalition of firm y with worker type x.

The surplus functions obey

$$\begin{aligned} S(x,y)(\rho+\delta) = f(x,y) - \rho\big[V_u(x) + V_v(y)\big] \\ + (1-\omega)M_f \int_{M(x,y)} \frac{d_u(\tilde{x}')}{U} S(\tilde{x}',(y,x))d\tilde{x}' \end{aligned} \tag{B.20}$$

and

$$\begin{aligned} S(x,(y,x'))(\rho+2\delta) = f(x,y,x') - \rho\big[V_u(x) + V_u(x') + V_v(y)\big] \\ - \big(\rho+\delta\big)S(x',y) + \delta S(x,y) \end{aligned} \tag{B.21}$$

Notice that $S(x,y)$ – the surplus of a match between worker x and firm y – is symmetric, whereas $S(x',(y,x))$ – the marginal surplus of adding x' to a (y,x) coalition – is not.

On a preliminary note, the proof extensively relies on the fact that either $S(x,y) = 0$ at the interior boundaries of the matching set or the non-interior boundaries do not change with x; similar remarks apply to $S(x,(y,x'))$. Consider, for simplicity, $M(x) = [\underline{\varphi}(x), \overline{\varphi}(x)]$. If $\underline{\varphi}(x) \neq 0$, then

[B.3]For ease of exposition, I drop the home production term $b(x)$. This is without loss of generality as long as $b'(x) \geq 0$. Notice also that $M(x)$ takes the second argument of $S(x,y)$ but $M(x,y)$ takes the first argument of $S(x',(y,x))$.

[B.4]The matching sets are symmetric as under the bargaining protocol imposed matching decisions are decision is mutually efficient.

$S(x, \underline{\varphi}(x)) = 0$; if $\underline{\varphi}(x) = 0$, then $\frac{\partial \underline{\varphi}(x)}{\partial x} = 0$. Analogously, if $\overline{\varphi}(x) \neq 1$, then $S(x, \overline{\varphi}(x)) = 0$; if $\overline{\varphi}(x) = 1$, then $\frac{\partial \overline{\varphi}(x)}{\partial x} = 0$.

Step 1. Using the property just described, differentiating $V_u(x)$ yields:

$$\begin{aligned}\frac{\partial V_u(x)}{\partial x}\rho = \omega M_u \int_{M(x)} \Bigg[& \frac{\partial S(x, \tilde{y})}{\partial x} \frac{d_v(\tilde{y})}{V} \\ & + \int_{M(x,\tilde{y})} \frac{\partial S(x, (\tilde{y}, \tilde{x}'))}{\partial x} \frac{d_m(\tilde{x}', \tilde{y})}{V} d\tilde{x}' \Bigg] d\tilde{y}\end{aligned} \tag{B.22}$$

Differentiating $S(x, y)$:

$$\frac{\partial S(x, y)}{\partial x}\big(\rho + \delta\big) = \left[\frac{\partial f(x, y)}{\partial x} - \rho\frac{\partial V_u(x)}{\partial x} + (1 - \omega)M_f \int_{M(x,y)} \frac{d_u(\tilde{x}')}{U}\frac{\partial S(\tilde{x}', (y, x))}{\partial x} d\tilde{x}'\right] \tag{B.23}$$

where I again exploit the boundary property.

Differentiating $S(x, (y, x'))$:

$$\frac{\partial S(x, (y, x'))}{\partial x}(\rho + 2\delta) = \left[\frac{\partial f(x, y, x')}{\partial x} - \rho\frac{\partial V_u(x)}{\partial x} + \delta\frac{\partial S(x, y)}{\partial x}\right] \tag{B.24}$$

Step 2. Notice also that differentiating $S(x', (y, x))$ with respect to x gives:

$$\frac{\partial S(x', (y, x))}{\partial x}\big(\rho + 2\delta\big) = \left[\frac{\partial f(x, y, x')}{\partial x} - \rho\frac{\partial V_u(x)}{\partial x} - \big(\rho + \delta\big)\frac{\partial S(x, y)}{\partial x}\right] \tag{B.25}$$

Thus, we can rewrite equation (B.23) as follows by substituting for $\frac{\partial S(\tilde{x}', (y, x))}{\partial x}$:

$$\begin{aligned}\frac{\partial S(x, y)}{\partial x}\big(\rho + \delta\big) = \Bigg[& \frac{\partial f(x, y)}{\partial x} - \rho\frac{\partial V_u(x)}{\partial x} \\ & + \big(\rho + 2\delta\big)^{-1}(1 - \omega)M_f \\ & \int_{M(x,y)} \frac{d_u(\tilde{x}')}{U}\left(\frac{\partial f(x, y, \tilde{x}')}{\partial x} - \rho\frac{\partial V_u(x)}{\partial x} - \big(\rho + \delta\big)\frac{\partial S(x, y)}{\partial x}\right) d\tilde{x}' \Bigg]\end{aligned} \tag{B.26}$$

Define

$$\mu_1(x,y) := \left[\big(\rho+\delta\big) + \frac{(1-\omega)(\rho+\delta)}{\rho+2\delta} M_f \int_{M(x,y)} \frac{d_u(\tilde{x}')}{U} d\tilde{x}'\right]$$
$$\mu_2(x,y) := \rho\left[1 + \frac{1-\omega}{\rho+2\delta} M_f \int_{M(x,y)} \frac{d_u(\tilde{x}')}{U} d\tilde{x}'\right]$$

where $\mu_1(x,y) \geq 0$ and $\mu_2(x,y) \geq 0$ for any (x,y), as ρ and δ and ω all lie in the interval $[0,1]$.

Then we can write equation (B.26) as

$$\frac{\partial S(x,y)}{\partial x} = \mu_1(x,y)^{-1}\left(-\mu_2(x,y)\frac{\partial V_u(x)}{\partial x} + \frac{\partial f(x,y)}{\partial x} + \frac{(1-\omega)}{\rho+2\delta} M_f \int_{M(x,y)} \frac{d_u(\tilde{x}')}{U}\frac{\partial f(y,x,\tilde{x}')}{\partial x} d\tilde{x}'\right) \tag{B.27}$$

Step 3. Substitute equation (B.24) into (B.22):

$$\begin{aligned}\frac{\partial V_u(x)}{\partial x}\rho = \omega M_u \Bigg[\int_{M(x)} & \frac{\partial S(x,\tilde{y})}{\partial x}\frac{d_v(\tilde{y})}{V} \\ & + \int_{M(x,\tilde{y})} \left\{(\rho+2\delta)^{-1}\left[\frac{\partial f(x,\tilde{y},\tilde{x}')}{\partial x} - \rho\frac{\partial V_u(x)}{\partial x} + \delta\frac{\partial S(x,\tilde{y})}{\partial x}\right]\right\}\frac{d_m(\tilde{x}',\tilde{y})}{V} d\tilde{x}'\Bigg] d\tilde{y}\end{aligned}$$

Now collect terms relating to $S(x,\tilde{y})$:

$$\begin{aligned}\frac{\partial V_u(x)}{\partial x}\rho = {} & \omega M_u \int_{M(x)} \left[\frac{\partial S(x,\tilde{y})}{\partial x}\left(\frac{d_v(\tilde{y})}{V} + \delta\big(\rho+2\delta\big)^{-1}\int_{M(x,\tilde{y})}\frac{d_m(\tilde{x}',\tilde{y})}{V} d\tilde{x}'\right)\right] d\tilde{y} \\ & + \omega M_u \int_{M(x)}\int_{M(x,\tilde{y})} \left\{\big(\rho+2\delta\big)^{-1}\left[\frac{\partial f(x,\tilde{y},\tilde{x}')}{\partial x} - \rho\frac{\partial V_u(x)}{\partial x}\right]\right\}\frac{d_m(\tilde{x}',\tilde{y})}{V} d\tilde{x}' d\tilde{y}\end{aligned} \tag{B.28}$$

Step 4. Finally, substitute equation (B.27) into equation (B.28):

$$\begin{aligned}\frac{\partial V_u(x)}{\partial x}\rho = \omega M_u \int_{M(x)} & \Bigg[\Bigg(\mu_1(x,\tilde{y})^{-1}\Bigg(-\mu_2(x,\tilde{y})\frac{\partial V_u(x)}{\partial x} + \frac{\partial f(x,\tilde{y})}{\partial x} \\ & + \frac{(1-\omega)}{\rho+2\delta} M_f \int_{M(x,\tilde{y})} \frac{d_u(\tilde{x}')}{U}\frac{\partial f(x,\tilde{y},\tilde{x}')}{\partial x} d\tilde{x}'\Bigg)\Bigg)\left(\frac{d_v(\tilde{y})}{V} + \delta\big(\rho+2\delta\big)^{-1}\int_{M(x,\tilde{y})}\frac{d_m(\tilde{x}',\tilde{y})}{V} d\tilde{x}'\right)\Bigg] d\tilde{y} \\ & + \omega M_u \int_{M(x)}\int_{M(x,\tilde{y})} \left\{\big(\rho+2\delta\big)^{-1}\left[\frac{\partial f(x,\tilde{y},\tilde{x}')}{\partial x} - \rho\frac{\partial V_u(x)}{\partial x}\right]\right\}\frac{d_m(\tilde{x}',\tilde{y})}{V} d\tilde{x}' d\tilde{y}\end{aligned}$$

Collecting terms yields:

$$\mu_3 \frac{\partial V_u(x)}{\partial x} = \\ \omega M_u \int_{M(x)} \left[\left(\mu_1(x,\tilde{y})^{-1} \left(\frac{\partial f(x,\tilde{y})}{\partial x} + \frac{(1-\omega)}{\rho+2\delta} M_f \int_{M(x,\tilde{y})} \frac{d_u(\tilde{x}')}{U} \frac{\partial f(\tilde{y},x,\tilde{x}')}{\partial x} d\tilde{x}' \right) \right) \right. \\ \left. \left(\frac{d_v(\tilde{y})}{V} + \delta\big(\rho+2\delta\big)^{-1} \int_{M(x,\tilde{y})} \frac{d_m(\tilde{x}',\tilde{y})}{V} d\tilde{x}' \right) \right] d\tilde{y} \\ + \frac{\omega}{\rho+2\delta} M_u \int_{M(x)} \int_{M(x,\tilde{y})} \frac{\partial f(x,\tilde{y},\tilde{x}')}{\partial x} \frac{d_m(\tilde{x}',\tilde{y})}{V} d\tilde{x}' d\tilde{y},$$

where

$$\mu_3 := \rho + \omega M_u \int_{M(x)} \mu_1(x,\tilde{y})^{-1} \mu_2(x,\tilde{y}) \left(\frac{d_v(\tilde{y})}{V} + \delta(\rho+2\delta)^{-1} \int_{M(x,\tilde{y})} \frac{d_m(\tilde{x}',\tilde{y})}{V} d\tilde{x}' \right) d\tilde{y} \\ + \frac{\omega\rho}{\rho+2\delta} M_u \int_{M(x)} \int_{M(x,\tilde{y})} \frac{d_m(\tilde{x}',\tilde{y})}{V} d\tilde{x}' d\tilde{y}.$$

Since $\frac{f(x,y)}{\partial x} > 0$, $\frac{f(x,y,x')}{\partial x} > 0$, we have that $\mu_3 \geq 0$. Hence, $\frac{\partial V_u(x)}{\partial x} \geq 0$. □

B.2.4.2 Proof of Part (ii): $w(x,y,x')$ is monotonically increasing in x Differentiating equation (B.15) with respect to x yields

$$\frac{\partial w(x,(y,x'))}{\partial x} = \omega \frac{\partial f(x,y,x')}{\partial x} + (1-\omega)\rho \frac{\partial V_u(x)}{\partial x}.$$

As long as $\omega > 0$ this is strictly increasing in x, since $\frac{\partial f(x,y,x')}{\partial x} > 0$ by assumption and we previously showed that $\frac{\partial V_u(x)}{\partial x} \geq 0$. □

B.3 Closed-form results for a stylized version

The full, quantitative model does not admit closed-form characterization of matching decisions and the distribution of wages. The fault lies with the value of either sides' outside option (waiting) which at once informs the matching decision and at the same time is endogenous to the matching decisions as taken by all agents in the economy. After all, it is those decisions pin down who you are likely to meet if you do wait.[B.5] To solve for matching decisions and distributions by hand, then, we need to sever their mutual interdependence.

[B.5]This general-equilibrium interplay between individual decisions and distributional dynamics is characteristic of the class of mean-field games to which the present model belongs.

To do so, I write down a version of the model of Eeckhout and Kircher (2011). I focus on matching of pairs of workers into a team, each managed by a firm, rather than matching into pairs consisting of one worker and one firm. I show how one can obtain not only a tight characterization of the key matching decision but also closed-form expressions for the degree of coworker sorting and between-firm inequality.

B.3.1 Environment

As in the main model, there is unit mass of workers, with types uniformly distributed over $\mathcal{X} = [0, 1]$. There is a mass $d_f = \frac{1}{2}$ of firms. As before, we denote their type $y \in \mathcal{Y}$, where $\mathcal{Y}$ is the set of potential firm types. But as in the baseline quantification I will assume that firms are ex-ante homogeneous with productivity normalized to unity.

Production with a single worker is normalized to be zero and $f(x, x') = x + x' - \gamma(x - x')^2$, where $\gamma > 0$ again indicates the strength of complementarities. This specification of team production is motivated by a second-order Taylor approximation to the micro-founded CES function in Proposition 1 around the mean worker type, i.e., around $\frac{x+x'}{2}$. Then $(x - x')^2$ is proportional to the variance of types and γ measures the production loss due to coworker mismatch.

We consider a finite-horizon setup in which there are three stages $s \in \{0, 1, 2\}$. All the focus will be on the outcomes in $s = 1$; the other two stages merely serve to frame the decision problem in $s = 1$. Everyone starts out unmatched. In $s = 0$, each firm meets and matches with one worker. Our focus is on the following matching decision, to be taken in $s = 1$. Each worker-firm pair, of which there is a mass $\frac{1}{2}$, is randomly paired with one of the remaining unmatched workers. *Either* they match, produce, and share the output. The agents are then idle in $s = 2$. *Or* they decide to *wait*, in which case all agents separate, each worker pays a *fixed search cost* c, and they can be active in stage $s = 2$. Specifically, in return for paying the search cost, each worker in the final stage is paired with their optimal match.

Relative to the full model, we have complete destruction of matches after the production stage ($\delta = 1$); every firm is guaranteed to meet one worker in every period; and there is zero discounting. As in Atakan (2006), the search cost is *explicit* instead and, as such, *type-independent*. This contrasts with the full model, where search costs are *implicit*. In that case, more productive agents have greater opportunity costs of time due to discounting, hence search frictions disproportionately erode their value of search (Sandmann and Bonneton, 2022).

Finally, payoffs are determined according to the following, simplified protocol. Firms have no bargaining power and each worker receives their respective outside option plus half the surplus generated by the match.

B.3.2 Solving the model by backward induction

FRICTIONLESS MATCHING IN $s = 2$. Matching in the second stage is frictionless. As described also in Boerma *et al.* (2021), the problem of a firm y is to choose workers x and x' to maximize profits, taking the wage schedule for workers, $w : [0,1] \longrightarrow \mathbb{R}_+$, as given:

$$v(y) = \max_{x,x'} \Big(f(x,x') - w(x) - w(x') \Big).$$

Any worker x chooses an employer y and coworker x' to maximize their wage income, taking as given the firm value, $v(y)$, and the coworker's wage schedule:

$$w(x) = \max_{y,x'} \Big(f(x,x') - w(x') - v(y) \Big).$$

Formally, an assignment is a probability measure π over workers, coworkers and firms. Let $\tilde{\Phi}_j$ denote the marginal distribution of workers ($j = x$) and firms ($j = y$), respectively; the tilde notation indicates that this will refer to the agents who arrive in the final stage. The set of feasible assignments is $\Pi := \Pi(\tilde{\Phi}_x, \tilde{\Phi}_x, \tilde{\Phi}_y)$, the set of probability measures π on the product space $\mathcal{X} \times \mathcal{X} \times \mathcal{Y}$, such that the marginal distributions of π onto $\mathcal{X}$ and $\mathcal{Y}$ are equal to $\tilde{\Phi}_x$ and $\tilde{\Phi}_y$, respectively. An equilibrium is a tuple (π, w, v) such that firms solve their profit maximization problem, each worker solves their respective problem, alongside a feasibility constraint:

$$\int f(x,x')d\pi = \int w(x)d\tilde{\Phi}_x(x) + \int w(x')d\tilde{\Phi}_x(x') + \int v(y)d\tilde{\Phi}_y.$$

The firm's first-order condition for a worker is

$$f_1(x,x') - \frac{dw(x)}{dx} = 0.$$

From the second-order conditions, per Becker (1973), we know that when the production function is supermodular in worker types, i.e., when $\gamma > 0$, then the optimal assignment among workers features positive assortative matching between coworkers. That is, we have a deterministic coupling, or matching function, $\mu : [0,1] \longrightarrow [0,1]$ and, specifically, $\mu^*(x) = x$.[B.6]

[B.6] Formally, the pure assignment corresponds to the case where π vanishes outside Graph(μ). Notice that from the viewpoint of optimal assignment, minimizing the quadratic loss, $\frac{1}{2}(x - x')^2$ is equivalent to maximizing the bilinear product, $x \cdot y$.

The wage schedule is then derived by integrating over the first-order condition,

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

where the integration constant is 0 since $f(0,0) = 0$. Under our parametric assumptions on the production function and the bargaining process, this yields $w^*(x) = x$ and $v^*(y) = v^* = 0$ for any $y \in \mathcal{Y}$.

STAGE-1 MATCHING DECISION. Each coalition of a firm with one worker, the latter being denoted x', is randomly matched with an unmatched worker whose type is x. *Either* they match, the team produces and the output value is shares. Then they are idle in the final stage. Else, the partners are all unmatched, each worker pays a fixed search cost c, and all agents actively participate in the stage-2 matching process.

A firm with worker x' that is randomly matched a worker x decides to hire them if the joint value of production is (weakly) greater than the sum of the respective outside options, taking into account the cost of search.[B.7] Using the same notation as in the full model, this means that $h(x,(y,x')) = 1 \Leftrightarrow S(x,(y,x')) > 0$, where[B.8]

$$S(x,(y,x')) = f(x,y,x') - \Big[w^*(x) + w^*(x') + v^*(y) - 2c\Big].$$

Substituting for the last-period payoffs, i.e., $w^*(x) = x$ and $v^*(y) = 0$, and simplifying shows that a match is formed whenever $|x' - x| > s^*$, where the equilibrium threshold s^* satisfies

$$s^* = \sqrt{2c/\gamma}. \tag{B.29}$$

Equivalently, the (symmetric) matching set is $M(x') = \{x \in \mathcal{X} : x' - s^* > x < x' + s^*\}$.[B.9]

It immediately follows that greater complementarities, corresponding to a higher value of γ, render the matching set narrower, whereas greater search costs, c, render the matching set wider.

STAGE-0 MATCHING DECISION. At the very beginning, each firm is randomly paired with one worker. It is optimal to match, since the opportunity cost of doing so is zero.[B.10]

[B.7] I assume that when the parties are indifferent between producing now and waiting they opt for the former. This assumption does not, of course, affect the substantive results.

[B.8] As discussed in Footnote 7 of Eeckhout and Kircher (2011), for low types the surplus in the next period may not exceed the total waiting cost. In order to avoid keeping track of endogenous entry, I assume that people will search even if that is the case.

[B.9] This simple formulation also emerges as a special case in Eeckhout and Kircher (2011). The result also resembles circular production models à la Marimon and Zilibotti (1999).

[B.10] To resolve the indifference case, we could of course assume an infinitesimally small positive production value.

B.3.3 Characterization

Next, we can characterize the sorting patterns and wage-distributional outcomes implied by this matching rule in closed-form.[B.11] To ensure maximum comparability with the full model, I take the following approach: I focus on the outcomes for workers that are part of matched formed in stage 1; and I re-weight these workers according to the (uniform) population distribution. That is, the marginal distribution of worker types participating in stage-1 matches is taken to be uniform. In this way, the results are not biased by the fact that stage-2 outcomes exhibit no mismatch by construction; nor is the workforce composition mechanically biased towards workers with intermediate skill levels who accept a wider range of partners than those with relatively lower or higher skills.[B.12]

B.3.3.1 Matching patterns The key step affording analytical tractability is to notice that, for any threshold level s, the distribution of coworker types conditional on type x, denoted $\Phi(x'|x)$, takes the form of a piecewise uniform distribution. Lemma 3 summarizes.

Lemma 3 (Conditional type distribution)**.** *Given a threshold distance s, the conditional distribution of coworkers for $x \in \mathcal{X}$ is*

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

A simple summary statistic capturing the degree of sorting between coworkers is the correlation coefficient.[B.13] The following Proposition provides a very simple closed-form expression.

Proposition B.1 (Coworker sorting)**.** *For a given threshold s,*

(i) *the coworker correlation is given by*

$$\rho_{xx} = (2s+1)(s^2-1)^2;$$

[B.11] All proofs and are collected in Section B.3.4.

[B.12] Concretely, since the matching set is $M(x) = \{x' \in \mathcal{X} : x - s^* < x > x + s^*\}$, for any $s^* < 1$, workers outside the interval $[s^*, 1-s^*]$ would be under-represented relative to those inside that interval.

[B.13] This coefficient is defined as

$$\rho_{xx} = \frac{\int\int (x-\bar{x})(x'-\bar{x})d\Phi(x'|x)d\Phi(x)}{\int (x-\bar{x})^2 d\Phi(x)},$$

where the numerator is the covariance; $\bar{x}$ is the average worker type, equal to $\frac{1}{2}$ given the uniform weighting; and the denominator is the variance.

(ii) *the average coworker type is*

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

Part (i) of Proposition B.1, together with equation (B.29) summarizing the equilibrium matching decision demonstrates that the coworker correlation coefficient, ρ_{xx} is *increasing* in γ – stronger complementarities give rise to more pronounced coworker sorting. In particular, $\rho_{xx} \to 1$ for $\gamma \to \infty$. Conversely, greater search costs dilute the incentive to wait for the best match and accordingly lower the degree of sorting observed in the economy. This result therefore sharply summarizes the tradeoff between complementarities and search costs in determining coworker sorting patterns in the economy.[B.14]

Part (ii) of Proposition B.1 paints a more disaggregated picture of matching patterns that reveals interesting *nonlinearities* arising from *asymmetries* in the matching set. It defines for every worker type the *average* coworker type for a given threshold level s, denoted $\hat{\mu}(x)$. Figure 3 in the main text graphically illustrates. The dashed-dotted and dotted lines describe matching patterns under, respectively, the deterministic coupling $\mu(x) = x$ that captures matching decisions under PAM in the frictionless economy (dashed 45-degree line); and the independent coupling that intuitively corresponds to a random matching process (dotted, horizontal line). The solid line describes $\hat{\mu}(x)$ for a low value of γ and the dashed line illustrates $\hat{\mu}(x)$ for a higher value of γ.

Three observations are warranted. First, for "middle types" – with $x = \frac{1}{2}$ being the paradigmatic case – the *average* coworker type is invariant to changes in γ. Intuitively, stronger complementarities mean that such an agent is less likely to be in a team with a much better agents but the likelihood of being teamed up with a much worse agent shrinks in symmetric fashion. Second, the presence of search costs means that low types are, on average, paired up with agents better than them, whereas high types are typically paired up with coworkers *worse* than them. Finally, and consequently, a strengthening of complementarities *increases* the average coworker type for the *best*; and it *lowers* it for the worst. This force has thus the potential to engender polarizing dynamics wherein "superstar teams" pull away while "laggards" fall behind (cf. Andrews *et al.*, 2019; Autor *et al.*, 2020)

[B.14] In practice, the worker types x are not observable, of course. Section B.3.3.3 below defines an observable measure of types following Borovičková and Shimer (2020), denoted $\lambda(x)$. The corresponding coworker sorting correlation coefficient is virtually indistinguishable from ρ_{xx} (in population, that is). It is also characterizable in closed form but less intuitively, being equal to $\rho_{\lambda\lambda} = \frac{10\gamma^2 s^6 - 3\gamma^2 s^5 + 3240 s^3 - 4860 s^2 + 1620}{-60\gamma^2 s^6 + 36\gamma^2 s^5 + 1620}$.

B.3.3.2 Wage distribution The production function, matching patterns, and wage sharing rule jointly determine the distribution of wages in the economy. Here we are particularly interested in the fraction of wages that occurs *between* teams. A tedious but otherwise straightforward sequence of integration and algebra steps – extensively leveraging Lemma 3 – affords the following characterization.

Proposition B.2. *Given a threshold distance s and a value of γ, the between-firm share of the variance of wages is equal to*

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

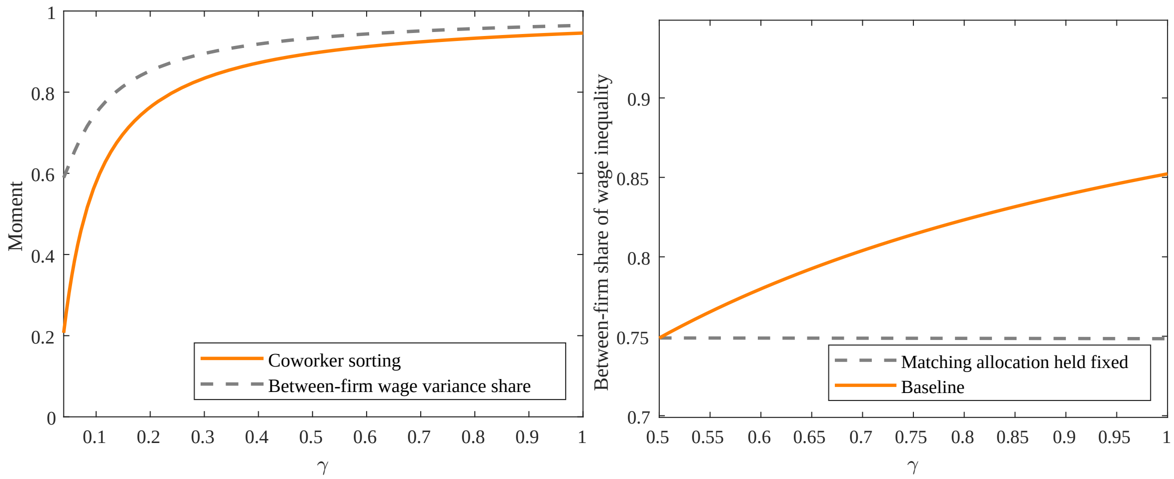
Despite this expression being slightly unwieldy, it makes precise the following two points. First, the between-firm variance share unambiguously increases with the strength of coworker complementarities, captured by γ. Figure B.1 illustrates this point by plotting this moment alongside the coworker correlation, ρ_{xx}, confirming their close co-movement. Second, for $s = 0$, which in particular obtains in equilibrium when search costs are absent, between-firm inequality accounts for *all* of the dispersion in wages. Third, since we have deliberately not yet substituted in for s using equation (B.29), we can connect the between-firm share even more tightly with workforce reallocation. To avoid mathematical expressions that are hard on the eye, the right panel provides a graphical illustration of the following exercise. We initially fix s at $s_0^* = \sqrt{2c/\gamma_0}$ and increase γ. The dashed line shows that the direct effect on the between-firm inequality share is almost nil, the reason being that the degree to which a greater cost of mismatch proportionally erodes the value of production does not systematically differ between high-paying and low-paying teams. By contrast, if we allow matching decisions to change, i.e., s^* moves alongside γ, then the associated reallocation toward greater coworker sorting pushes up between-firm inequality.[B.15]

B.3.3.3 Measurement of types & fixed effects Thus far, we derived moments of interest on the basis of worker's true types, respectively, their rank in the productivity distribution. Those are, of course, unobserved. To bridge the gap to empirical work, this section supposes that for every worker we observe many repeated (stage-1) outcomes and arrange these observations as a panel.

Borovičková and Shimer (2020) propose measuring a worker's (cardinal) type as the expected log wage received by a worker. The corresponding measure in the present context (in levels) is

$$\lambda(x) = \int_0^1 w(x,x')d\Phi(x'|x).$$

[B.15] I use the term "reallocation" in a loose sense, as we are merely conducting a comparison of steady states.



(a) Between-share tracks coworker sorting

(b) Workforce (re-)allocation is key

Figure B.1: Coworker sorting and between-firm inequality in the stylized model

Notes. The left panel plots ρ_{xx} and the between-firm share of wage variance against γ. The experiment illustrated in the right panel is described in the main text.

Performing the relevant integration steps, utilizing Lemma 3, and simplifying yields

$$\lambda(x) = \begin{cases} \lambda^b(x) := x - \frac{\gamma}{6}\left[(x-s)^2 + xs\right] & \text{for } x \in [0,s) \\ \lambda^m(x) := x - \frac{\gamma}{6}s^2 & \text{for } x \in [s, 1-s] \\ \lambda^h(x) := x - \frac{\gamma}{6}\left[1 + s - 2x - s^2 - x^2 - sx\right] & \text{for } x \in (1-s, 1]. \end{cases}$$

In population, and given a wage rule that is very sensitive to the worker's true type, $\lambda(x)$ and x are highly correlated in population and even in short samples.[B.16]

In practice, and given a matched employer-employee dataset (indexing each worker by i and employers by j), alternative approach to measuring worker types, following AKM, considers a reduced-form, two-way fixed effect (FE) wage equation of the form

$$w(x_i, y_{j(i)}, \epsilon) = \alpha_i + \psi_{j(i)} + \epsilon_{i,j(i))}$$

where $\alpha_i = \alpha(x_i)$ is worker i's FE, $\psi_{j(i)} = \psi(y_{j(i)})$ is the employer's FE, and ϵ_{ij} is an error term such that the distribution has mean-zero for all (i, j) pairs.

Following the reasoning of Eeckhout and Kircher (2011), even if this equation is misspecified, s.t. it doesn't hold for all types, we can still define the FEs as the solutions to the following moments

[B.16] It is straightforward to verify this in simulations.

conditions:

$$\alpha(x) = \int_{\mathcal{X}} \Big(w(x,y) - \psi(y) \Big) d\Phi(y|x)$$
$$\psi(y) = \int_{\mathcal{Y}} \Big(w(x,y) - \alpha(x) \Big) d\Phi(x|y)$$

Performing the appropriate integration steps shows that ϵ_{ij} does indeed drop out when integrating over the conditional match distribution. In the absence of ex-ante firm heterogeneity, two points immediately follow. First, the true firm FE is zero. Second, for any type x, the worker FE $\alpha(x)$ coincides with the expected wage, $\lambda(x)$.[B.17]

It bears highlighting that the worker FE – which is often interpreted as a proxy for a worker's ability – is the solution to a system of differential equations wherein we integrate over the *conditional* distribution of coworkers. One interpretation thereof is that this represents a mathematical interpretation of the following remark by Neffke (2019): "[T]here is a social dimension to human capital, rooted in a deepening division of labor that progressively distributes collective know-how [...] as knowledge becomes distributed, human capital acquires a distinct relational dimension: Each worker packages a specific set of skills, and, as a consequence, a worker's productivity will depend on the skills of the people he or she works with."

B.3.4 Proofs and derivations

B.3.4.1 Lemma 3 The distribution of coworkers for a worker of type x is

$$\Phi(x'|x) = \frac{\int_{\mathcal{X}} \mathbf{1}\{\tilde{x} \in M(x)\} d\Phi(\tilde{x})}{\Phi(\overline{m}(x)) - \Phi(\underline{m}(x))}$$

where $\overline{m}(x)) = \sup M(x)$ and $\underline{m}(x)) = \inf M(x)$; given the assumption of uniform weights on types among the matched $\Phi(x)$ is the uniform cdf.

Assuming that potential coworkers are uniformly distributed, the equilibrium matching rule then implies that

$$x' \sim \begin{cases} U[0, x+s] & \text{for } x \in [0,s) \\ U[x-s, x+s] & \text{for } x \in [s, 1-s] \\ U[x-s, 1] & \text{for } x \in [1-s, 1] \end{cases}$$

[B.17] Observe, however, that the error term ϵ_{ij}, which in this case captures match-specific effects, is not uncorrelated with the regressors; e.g., for $x \in [s, 1-s]$, $\lambda(x) = x - \frac{\gamma}{6}(s^*)^2$ and $\epsilon(x, x') = \frac{\gamma}{6}(s^*)^2 - \frac{\gamma}{2}(x - x')^2$.

Writing this out more carefully for three different segments, for $x \in [0, s)$:

$$\Phi(x'|x) = \begin{cases} 0 \quad \text{for} \qquad x' = 0 \\ \frac{x'}{x+s} \quad \text{for} \quad x' \in [0, x+s) \\ 1 \quad \text{for} \qquad x' > x+s. \end{cases}$$

Next, for $x \in [s, 1-s]$:

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

Finally, for $x \in (1-s, 1]$:

$$\Phi(x'|x) = \begin{cases} 0 \quad \text{for } x' < x-s \\ \frac{x'-(x-s)}{1-(x-s)} \quad \text{for } x' \in (x-s, 1] \\ 1 \quad \text{for } x' = 1. \end{cases}$$

□

B.3.4.2 Proposition B.1 Part (ii) immediately follows from 3. For part (ii), notice first that under the uniform distribution, the mean type is $\bar{x} = \frac{1}{2}$ and the variance is $\frac{1}{12}$. We need to derive the covariance term, which is defined as

$$\bar{c}_{xx} = \int_X \int_X (x-\bar{x})(x'-\bar{x}) d\Phi(x'|x) d\Phi(x).$$

Step 1. Substituting for the piece-wise conditional match distribution and the uniform population weights, for a threshold s,

$$\begin{aligned} \bar{c}_{xx} &= \int_0^s \int_0^{x+s} (x-\frac{1}{2})(x'-\frac{1}{2}) \frac{1}{x+s} dx' dx \\ &+ \int_s^{1-s} \int_{x-s}^{x+s} (x-\frac{1}{2})(x'-\frac{1}{2}) \frac{1}{2s} dx' dx \\ &+ \int_{1-s}^{1} \int_{x-s}^{1} (x-\frac{1}{2})(x'-\frac{1}{2}) \frac{1}{1-(x-s)} dx' dx \end{aligned}$$

Step 2. Integrate over x'.

The different components are:

$$\int_0^{x+s}(x-\frac{1}{2})(x'-\frac{1}{2})\frac{1}{x+s}dx' = \frac{1}{2}(x-\bar{x})(s+x-2\bar{x})$$

$$\int_{x-s}^{x+s}(x-\frac{1}{2})(x'-\frac{1}{2})\frac{1}{2s}dx' = (x-\bar{x})^2$$

$$\int_{x-s}^{1}(x-\frac{1}{2})(x'-\frac{1}{2})\frac{1}{1-(x-s)}dx' = \frac{1}{2}(x-\bar{x})(1-s+x-\bar{x})$$

Step 3. Integrate over x.

$$\begin{aligned}
\bar{c}_{xx} &= \int_0^s \frac{1}{2}(x-\bar{x})(s+x-2\bar{x})dx + \int_s^{1-s}(x-\bar{x})^2 dx + \int_{1-s}^1 \frac{1}{2}(x-\bar{x})(1-s+x-\bar{x})dx \\
&= \left[\frac{5}{12}s^3 - \frac{5}{8}s^2 + \frac{1}{4}s\right] + \frac{(2s-1)^3}{12} + \left[\frac{5}{12}s^3 - \frac{5}{8}s^2 + \frac{1}{4}s\right] \\
&= \frac{1}{12}\left[(2s+1)(s-1)^2\right]
\end{aligned}$$

Hence, dividing by the variance, the *correlation* is

$$\bar{\rho}_{xx} = (2s+1)(s^2-1)^2.$$

□

B.3.4.3 Proposition B.2 To find the between-team share of the variance of wages, we first compute the total wage dispersion as the same of the variance of wages between and within *types*; then we find the variance of the average wage by team; the between-share is the ratio of the latter over the former.

Many of the steps are straightforward but tedious and lengthy, following similar steps as in the preceding section, hence I only sketch the procedures and intermediate results.

Total wage dispersion. The average wage is equal to the average type $\lambda(x)$ and can be computed as

$$\bar{\lambda} = \int_0^s \lambda^b(x)dx + \int_s^{1-s}\lambda^m(x)dx + \int_{1-s}^1 \lambda^h(x)dx$$

After some further integration and simplification, this simplifies to read

$$\bar{\lambda} = \frac{1}{2} - \frac{1}{6}\gamma s^2(1-\frac{s}{3})$$

The between-type variance is equal to

$$\bar{\sigma}^2_{\text{between-type}} = \int_0^s \left(\lambda^b(x) - \bar{\lambda}\right)^2 dx + \int_s^{1-s} \left(\lambda^m(x) - \bar{\lambda}\right)^2 dx + \int_{1-s}^1 \left(\bar{\lambda}^h(x) - \bar{\lambda}\right)^2 dx,$$

which corresponds to the contribution of worker heterogeneity to total wage inequality.

After more integration and algebra, we find that this is equal to

$$\bar{\sigma}^2_{\lambda} = \frac{1}{12} - \frac{\gamma^2 s^5}{12}\left(\frac{s}{27} - \frac{1}{45}\right)$$

Conditional on a type x, the variance of wages is equal to

$$\sigma^2_{\lambda(x)} = \int_0^1 \left(w(x,x') - \lambda(x)\right)^2 d\Phi(x'|x).$$

Calculating this expression for the three segments of worker types – bottom, middle, and high – we then obtain the average within-type wage variance as

$$\sigma^2_{\text{within}\,\lambda} = \int_0^s \sigma^2_{\lambda^b(x)} dx + \int_s^{1-s} \sigma^2_{\lambda^m(x)} dx + \int_{1-s}^1 \sigma^2_{\lambda^h(x)} dx.$$

Performing the usual piece-wise integration yields

$$\sigma^2_{\text{within-type}} = \frac{\gamma^2\, s^4\, \left(2280\, s\, \ln(2) - 1639\, s + 80\right)}{3600},$$

and, hence, the total wage variance is equal to

$$\sigma^2_w = \frac{1}{12} + \frac{\gamma^2\, s^4}{45} - \frac{4897\,\gamma^2\, s^5}{10800} - \frac{\gamma^2\, s^6}{324} + \frac{19\,\gamma^2\, s^5\, \ln(2)}{30}.$$

Between-share of wage inequality. Finally, we compute the variance of the average wage in a firm, which given our assumptions on the firm earning zero return just corresponds to output per worker (i.e., productivity). That is,

$$
\begin{aligned}
\sigma^2_{w,\text{between-firm}} &= \int_0^s \left[\int_0^{x+s} \left(\frac{f(x,x')}{2} - \bar{\lambda} \right)^2 \frac{1}{x+s} dx' \right] dx \\
&+ \int_s^{1-s} \left[\int_{x-s}^{x+s} \left(\frac{f(x,x')}{2} - \bar{\lambda} \right)^2 \frac{1}{2s} dx' \right] dx \\
&+ \int_{1-s}^{1} \left[\int_{x-s}^{1} \left(\frac{f(x,x')}{2} - \bar{\lambda} \right)^2 \frac{1}{1+s-x} dx' \right] dx \\
&= \frac{1}{12} - \frac{13\,\gamma^2\,s^5}{2400} + \frac{\gamma^2\,s^4}{80} + \frac{5\,s^3}{36} - \frac{s^2}{6},
\end{aligned}
$$

where the last equality follows after another sequence of integration and algebra. The result in Proposition B.2 then follows by taking the ratio $\frac{\sigma^2_{w,\text{between-firm}}}{\sigma^2_w}$. □

C Estimation & quantitative analysis

C.1 Validation of identification approach

Following Bilal *et al.* (2022), I conduct two exercises to validate the identification of the vector of jointly estimated parameters, ψ. First, to support the argument in the main text that each element of ψ is closely linked to a particular moment, Figure C.1 plots each moment against the respective parameter. As required for local identification, the relationships are monotonic and exhibit significant amount of variation. For the second exercise, let a given parameter ψ_i vary around the estimated value ψ_i^* and plot the distance function $\mathcal{G}(\psi_i, \psi_{-i}^*)$, where the remaining parameters are allowed to adjust to minimize $\mathcal{G}$. Reassuringly, Figure C.2 indicates that $\mathcal{G}(\psi_i, \psi_{-i}^*)$ has a steep U-shape, suggesting that ψ is indeed well-identified.

C.2 Computation of the cross-partial wage derivative

To non-parametrically approximate the average cross-partial derivative of the empirical wage function, I proceed as follows.

(i) Construct the non-parametric wage function, which given an approximation of the worker type space on n_x grid points is a 10×10 matrix. Denote as w_{ij} the average wage of a worker in decile i of the type distribution whose average coworker is in decile j of the coworker type distribution, $i = 1, ..., n_x$ and $j = 1, \dots, n_x$.

(ii) To prevent local nonlinearities in the conditional wage function from biasing the estimate, I fit a cubic polynomial separately for each worker type x. Denote the predicted values $\tilde{w}_{ij}$.(In that sense, the approach here is not *completely* non-parametric.)

(iii) Given a matrix of these predicted values, I then numerically compute the cross-partial derivative based on the matrix $[\tilde{w}_{ij}]_{i=1,\dots,n_x}^{j=1,\dots,n_x}$ using finite difference methods: approximate the derivative using forward differences for for $i, j = 1$; using backward differences for $i, j = n_x$; and using central differences everywhere else.

(iv) Finally, I compute a weighted average of these based on the observed match density; $\widehat{\frac{\partial w^2(x,x')}{\partial x \partial x'}} = \sum_{i=1}^{n_x} \sum_{j=1}^{n_x} \frac{d_m(x_i,x_j)}{\sum_{i=1}^{n_x} \sum_{j=1}^{n_x} d_m(x_i,x_j)} \times \left[\widehat{\frac{\partial w^2(x,x')}{\partial x \partial x'}} \right]_{ij}$.

C.3 Adjustment procedure

Since in the structural model the number of workers in each production unit is normalized to two, the between-unit share of the variance of wages — or, for that matter, that of types — will be greater than zero even under random matching. More generally, for any degree of coworker sorting less

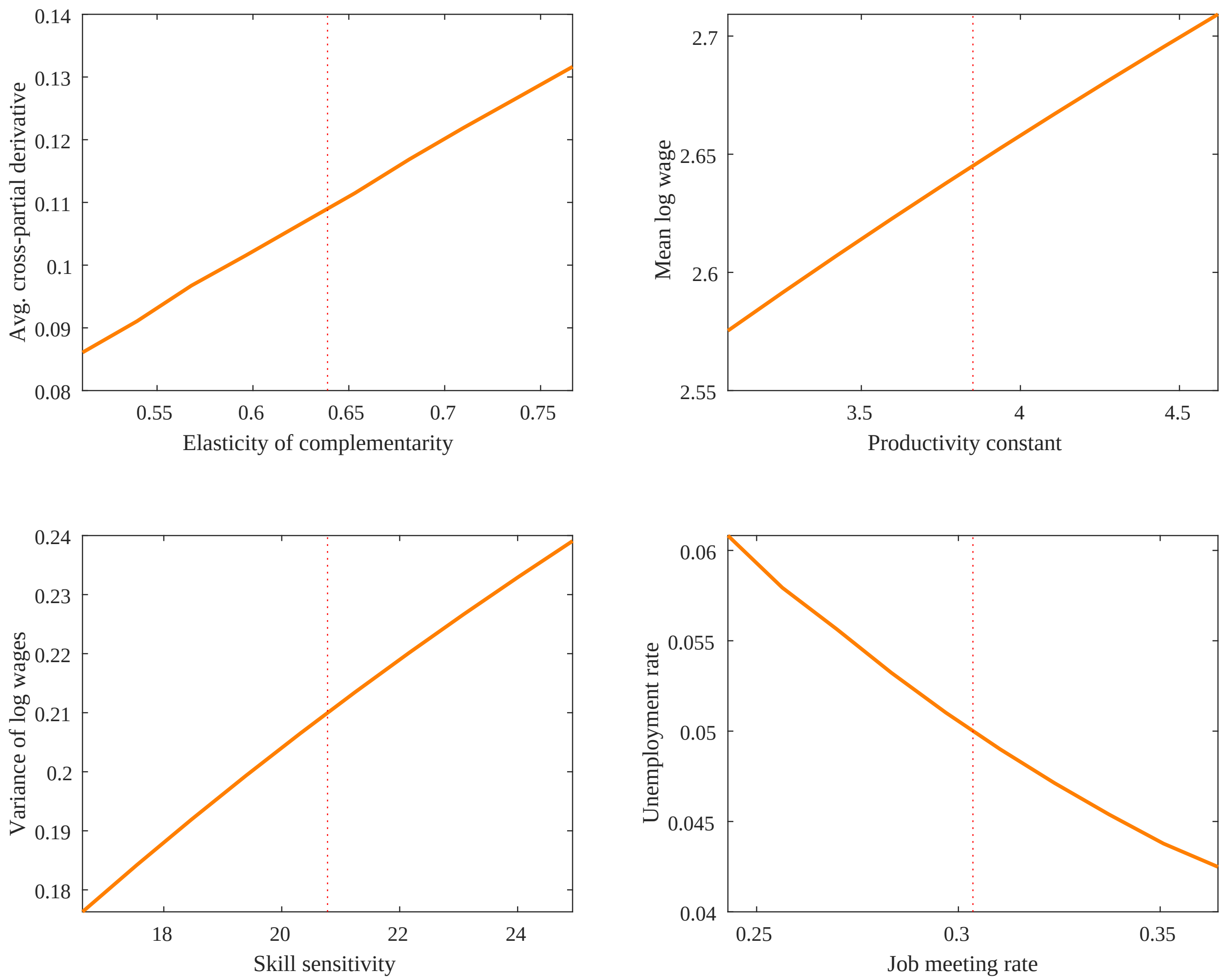


Figure C.1: Validation of identification method: moment against parameter

Notes. This figure plots the targeted moment against the relevant parameter, holding constant all other parameters. The parameters considered are, respectively, γ, a_0, a_1, and M_u.

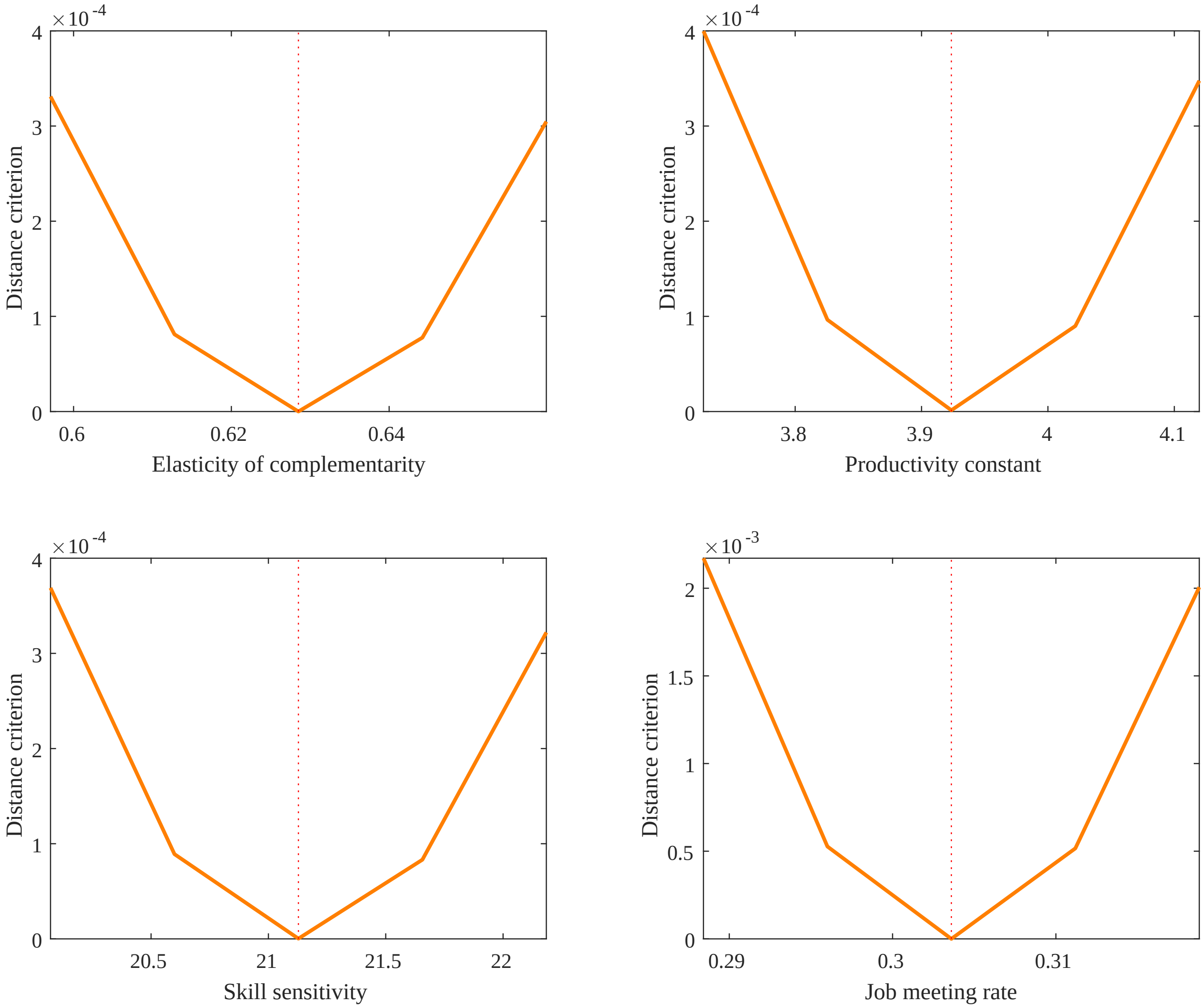


Figure C.2: Validation of identification method: distance criterion

Notes. This figure plots the distance function $\mathcal{G}(\psi_i, \psi^*_{-i})$ when varying a given parameter ψ_i around the estimated value ψ^*_i. The remaining parameters are allowed to adjust to minimize $\mathcal{G}$.

than unity, i.e. $\rho_{xx} < 1$, the *level* of the between-share in a model with team size $m = 2$ will be biased upward relative to the case of $m > 2$ and, in particular, $m \to \infty$. These observations are simply an outgrowth of the law of large numbers not applying within production units. As such, it is a statistical phenomenon rather than an economically interesting mechanism. Furthermore, the upward bias is *greater* when the coworker correlation is *lower*. One can immediately verify this result intuitively by noting that for $\rho_{xx} = 1$, *all* dispersion is *across* and *none* within units, regardless of the value of m. Figure C.3 illustrates these ideas graphically (its construction is described below).

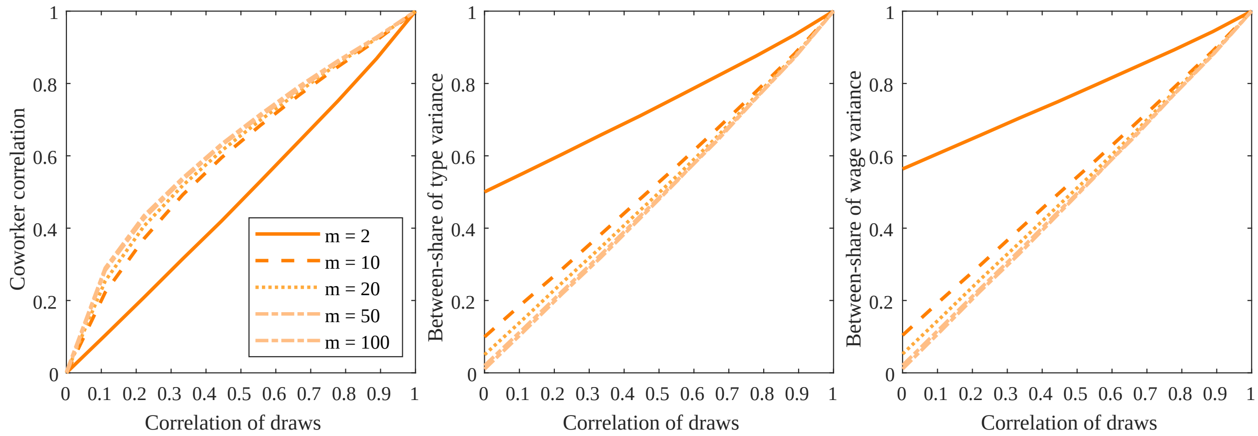
This statistical bias has two implications. First, without any further adjustment, the *level* of the between-share predicted by the estimated model, which assumes $m = 2$, will be excessively high relative to the real world, in which $m > 2$. Second, insofar as the estimated model predicts greater stronger coworker sorting than the earlier period, the predicted increase in the between-share of wage inequality is a *lower bound*, because the statistical upward bias in the later period will be smaller than in the later period. This section proposes an approach to correct the level of the between-share, but all results for *changes over time* that are reported in the main text are a conservative estimate, insofar as I apply the same correction factor to both periods.

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 $X = (X_1, X_2, \dots, X_m)'$ whose distribution is described by a Gaussian copula over the unit hypercube $[0,1]^m$, with an $m \times m$ dimensional correlation matrix $\Sigma(\rho^c)$, which contains ones on the diagonal and the off-diagonal elements are all equal to a parameter ρ_c. Formally, the Gaussian copula with parameter matrix $\Sigma(\rho_c)$ is $C_{\Sigma}^{\text{Gauss}}(x) = \Phi_R(\Phi^{-1}(x_1), \dots, \Phi^{-1}(x_m))$, 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 $\Sigma(\rho_c)$. To map this onto our empirical context, m may be interpreted as the average team size. Each vector of observations drawn from the distributions of X, $x_j = (x_{1j}, x_{2j}, \dots, x_{mj})'$, describes the types of workers in that team, indexed by j.

This setup affords us with a closed-form description of the population between-team share of the variance of types as a function of m and ρ_c. Since the marginals of the Gaussian copula are simply continuous uniforms defined over the unit interval, the variance of the union of all draws is just $\frac{1}{12}$. Furthermore, the mean of the elements of X is itself a random variable, $\bar{X}$. That is, for some realization x_j, we can define $\bar{x}_j = \frac{1}{m}\sum_{i=1}^{m} x_{ij}$. Since the elements of X all have the same variance and we specified their correlation profile, the variance of $\bar{X}$ will be $\frac{1}{m^2}\left(\frac{m}{12} + m(m-1)(\frac{\rho_c}{12})\right)$. Taking the ratio, we find that the between share, as a function of ρ_c and m, is equal to

$$\sigma^2_{x,\text{between-share}}(\rho_c, m) = \frac{1}{m}\left(1 + (m-1)\rho_c\right).$$

In the main text, and letting the empirical average size be $\hat{m}$, the adjusted results for the between-



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

Figure C.3: Graphical illustration of the adjustment method

Notes. See text for a description of the method used. The result are based on binning workers into ten groups and one million draws.

share are therefore obtained by subtracting the following correction factor:

$$\text{correction-factor} = \tfrac{1}{2}\Big(1+\rho_c\Big) - \tfrac{1}{\hat{m}}\Big(1+(\hat{m}-1)\rho_c\Big),$$

The value of ρ_c I feed into this formula is 0.42, which is the average coworker correlation value in the earlier sample.[C.1] The implied value of the correction factor is ≈ 0.25.

There are two potential concerns with this approach. First, ρ_c is not the same measure as the coworker correlation, ρ_{xx} that we typically consider in both the empirical analysis and the structural model. To compare the two measures, suppose we draw n samples (i.e., distinct teams) from X, so that the total number of observations is $m \times n$. Each individual observation is indexed by $i = 1, \dots, m \times n$, and the sample to which i belongs is $j(i)$. Then we can define the leave-out-mean $\bar{x}_{-i,j} = \frac{1}{m-1}\sum_{k \neq i} x_{k,j(i)}$. As in our main analysis, the coworker correlation is $\rho_{xx} = corr(x_i, \bar{x}_{-i})$, where the j indexed is suppressed to emphasize that we are considering a worker-weighted statistic. Figure C.3 confirms that ρ_{xx} and ρ_{cc} track each other quite closely, even though for larger values of m and intermediate values of ρ_c we find that $\rho_{xx} > \rho_{cc}$.[C.2]

The second concern is that the adjustment approach pertains, strictly speaking, to the between-unit share of the variance of *types*, as opposed to that of *wages*. However, given a distribution of workers

[C.1] I choose to base the correction on the earlier sample, yielding a bigger downward adjustment, to avoid over-stating the degree of between-firm inequality that the model can generate without assuming ex-ante firm heterogeneity. In addition, I use $\hat{m} = 15$. (The exact value of $\hat{m}$ does not matter much, since for reasonable values of $\hat{m}$ the implied correction factors are very close to each other. The magnitude of the bias rapidly diminishes as $\hat{m}$ grows, as is evident from the above formula.)

[C.2] Notice that in order to match the structural model as closely as possible, I binned the draws in the same way as I did in the structural model and for empirical analyses. Of course, the statistical environment makes it possible to examine the implications of such binning and, reassuringly, it makes little difference to the results.

across production units derived from the statistical model, we can easily impute wages based on the structural wage function derived from the structural model, and then repeat the variance decomposition for wages. The figure confirms that the two adjustment factors obtained when looking at types and wages, respectively, are very similar to one another. Overall, the proposed adjustment approach therefore seems to accomplish the desired goal with some reliability.

D The task content of production in Germany

To document patterns in the task content of production I draw on the Employment Surveys (ES) carried out by the German Federal Institute for Vocational Training (Bundesinstitut fuer Berufsbildung, BIBB).[D.1] Following the influential methodology introduced in Autor *et al.* (2003) and first applied to the ES by Spitz-Oener (2006), I use the tasks that employees report to have performed at the workplace to measure variations in the nature of work.

These surveys have several attractive features: they provide detailed information on tasks performed at work; the survey has been run, in repeated waves, since 1985 (namely: 1985/86, 1991/92, 1998/99, 2006, 2012, and 2018), facilitating time series analyses; each wave has a large sample size between 20,000 to over 30,000 respondents per wave, facilitating between-group comparisons; responses are at the *worker-level* and consistent occupation codes can be used across multiple waves, making it possible to capture changes in nature of work not only associated with employment shifts across occupations but also within-occupation (on the importance of which see, e.g., Spitz-Oener (2006); Atalay *et al.* (2020)); and a supplemental survey in 2012 allows enriching binary task indicators with information on the actual shares of time spent by employees in different occupations on various tasks.

The analysis uncovers several key findings.

(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 is robust and not merely the result of composition effects, and that there is substantial cross-sectional variation that can be exploited in regression analysis, as done in the main text.

[D.1] I thank the Research Data Center of the Federal Institute for Vocational Education and Training for providing access to scientific use files as well as guidance. All remaining errors are my own.

D.1 Methodology

SAMPLE RESTRICTIONS. As detailed in Rohrbach-Schmidt and Tiemann (2013) and Hall and Rohrbach-Schmidt (2020), time comparisons with the BIBB/IAB surveys require a standardization of the sample basis. To that end, I follow the steps detailed in those reports and focus on employees from West Germany, aged 15 to 65, who belonged to the labor force (defined as having a paid employment situation) with a regular working time of at least ten hours per week.[D.2] The final sample comprises 91,152 worker-year observations.

In addition, I entirely omit the 1998/99 wave from my analysis, because the number of activities queried in that wave is substantially lower than in the other surveys. While doing so reduces the overall sample size, this choice avoids bias to the results due to the limited comparability in tasks. For example, none of the activities "accommodating", "caring", "storing", "protecting", "programming" and "cleaning" were queried in the 1998/99 survey.[D.3]

TASK CLASSIFICATION. As the detailed discussion in Rohrbach-Schmidt and Tiemann (2013) makes clear, comparisons over task intensities using the BIBB ES over time need to be implemented carefully and account for variation over time in what tasks are queried and whether their content has changed in meaning. Writing in the context of typical studies that compare task items in the categories non-routine analytical, non-routine interactive, non-routine manual, routine-cognitive and routine-manual, the authors highlight in particular that routine cognitive tasks are difficult to classify (e.g. "measuring" may be routine cognitive or routine manual; also see the findings of Antonczyk *et al.* (2009) in comparison to those by Spitz-Oener (2006)).[D.4]

Given my focus on non-routine or complex tasks, these classification problems are less severe though, as "these items are regularly observable throughout the cross-sections, their content did not change significantly from year to year, and measurement validity is comparatively strong," as Rohrbach-Schmidt and Tiemann (2013)) note when suggesting to researchers to foucs on the increase in these tasks.

As summarized in Table D.1, I therefore include several task items in the index of abstract tasks — guided by the classification of non-routine tasks in Spitz-Oener (2006), Rohrbach-Schmidt and Tiemann (2013) and Atalay *et al.* (2020) – and compare them with all other tasks, i.e., those that are

[D.2] In addition, I drop observations who report having performed none of the activities queried in at least two waves and, given the extensive use of occupational codes, any occupations with fewer than thirty observations across all waves.

[D.3] I thank Daniela Rohrbach-Schmidt for her generous and patient advice on how to handle the older waves as well as for sharing programs illustrating how the scientific use files available via GESIS can be rendered maximally consistent with the original data.

[D.4] Autor and Handel (2013) also treat the "physical" dimension of tasks as a combined measure of physical and routine tasks. Meanwhile, Acemoglu and Autor (2011) subsume non-routine analytical and non-routine interactive into "abstract", while routine-cognitive and routine-manual tasks are subsumed into "routine".

Task classification	**Task name**	**Description**
Complex/abstract	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 D.1: Classification of tasks in the BIBB Employment Surveys

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

broadly categorized as routine or manual.[D.5]

TASK INDEX. Given this classification, I then define an index capturing the usage of abstract/complex tasks for worker i in period t, following Antonczyk *et al.* (2009):

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

To illustrate, if worker i performs five distinct activities in sample period t and two of those belong to the category of abstract/complex tasks, then the complexity index for her work is 0.4.

OCCUPATIONAL CLASSIFICATION. To ensure a consistent classification of occupations when using information from multiple waves, I use the German Classification of Occupations 1988 (KldB88). As the oldest classification available in the two most recent waves (2012 and 2018) is the KldB92 classification (cf. Table 9 in Hall and Rohrbach-Schmidt (2020))), in processing these two waves I rely on a conversion table KldB92→KldB92; the conversion quality is high as the two classifications

[D.5] I do not use the task items "managing", "applying law" and "negotiating", because they are only measured in the early waves. Moreover, I associate buying/selling with "other", since even though they may be hard to automate (even that seems questionable in light of self-checkouts and e-commerce), they are arguably not among the most complex activities. This decision makes no practical difference, though.

are very similar.[D.6]

D.2 Results

PATTERNS OVER TIME. The first column in Table D.2 indicates that the aggregate usage share of complex tasks in workers' activities has monotonically increased from 1986 to 2018, with the increase being particularly pronounced in the first period.

The second to fourth columns decompose the period-by-period change in the importance of abstract/complex tasks into two components: a "between" component that captures shifts in occupational employment shares and a "within" component that measures changes in the task content within occupations. Formally, as in Atalay *et al.* (2020), I decompose changes in the usage of abstract tasks between periods t and $t-1$ according to the equation

$$\Delta \bar{T}_t^{\text{complex}} = \sum_o \omega_{o,t-1}(\bar{T}_{t,o}^{\text{abstract}} - \bar{T}_{t-1,o}^{\text{complex}}) + \sum_o (\omega_{o,t} - \omega_{o,t-1})\bar{T}_{t,o}^{\text{abstract}}$$

where $\bar{T}_{t,o}^{\text{complex}}$ measures the average usage of complex tasks by members of occupation o in period t and $\omega_{o,t}$ is the period- t employment share of occupation o.

Consistent with the findings of Atalay *et al.* (2020) for the US, this decomposition reveals that about three quarters of the increase in complex tasks over the sample period have occurred *within* occupations.

Education offers an alternative lens through which to view the changing task content. As shown in Figure 1a in the main text, the share of complex tasks in the portfolio of university-educated individuals is substantially greater than that of persons with less formal education. The *increase* over time takes place across the board, however.

CROSS-SECTIONAL PATTERNS. The share of complex tasks also varies greatly by occupation. To provide a comparison to the cross-sectional analysis of ISCO-08 2-digit occupations in Section 3.3.2 I compute the average task shares by occupation in the waves 2012 and 2018 according to that classification. Table D.3 lists the bottom 5 and top 5 occupations and shows both the complexity measure based on the ES and the non-routine abstract score from Mihaylov and Tijdens (2019). The comparison reveals large variation in task shares according to either measure.

TIME USAGE. One concern with the main analysis might be that I only considered whether a given task represents an important activity in the respondent's job as opposed to measuring how

[D.6] This crosswalk is based on the Klassifikationsserver der Statistischen Ämter des Bundes und der Länder, current occupations coded in the 2006 wave in which both KldB88 and KldB92 are available and personal decisions. I thank Anett Friedrich for creating and sharing the crosswalk.

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

Table D.2: The evolving task content of production in Germany

Notes. This table reports the within-between occupation decomposition of the change in the share of complex tasks over time. The "Total" column aggregates across all individuals. The decomposition is performed at the level of KldB-1988 2-digit occupations.

ISCO-08 2-digit occupation	$\bar{T}_0^{\text{complex}}$	MT-NRA
Business and administration professionals	0.84	0.47
Legal, social and cultural professionals	0.83	0.67
Business and administration associate professionals	0.82	0.29
Teaching professionals	0.81	0.57
Administrative and commercial managers	0.81	0.58
...	...	...
Drivers and mobile plant operators	0.2	0
Agricultural, forestry and fishery labourers	0.14	0
Food preparation assistants	0.14	0
Market-oriented skilled forestry, fishery and hunting workers	0.12	0
Cleaners and helpers	0.12	0

Table D.3: Top- and bottom-5 occupations in terms of task complexity

Notes. This table reports the top-5 and bottom-5 ISCO-08 2-digit occupations when ranked by $\bar{T}_0^{\text{complex}}$ in pooled 2012 and 2018 waves. The column "MT-NRA" shows the non-routine abstract score taken from Mihaylov and Tijdens (2019) after collapsing to the ISCO-08-2d level using occupational employment shares (computed using the Portuguese data, for lack of representative German data).

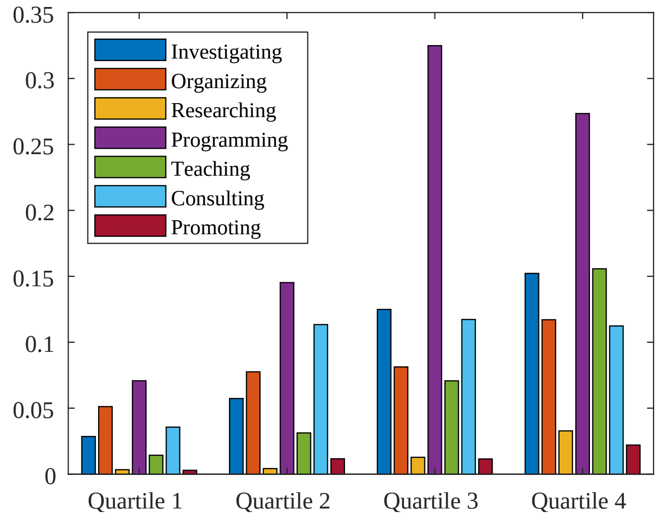


Figure D.1: Allocation of time to complex tasks by occupational groups

Notes. This table reports the share of time spent on complex tasks by different occupational groups. Occupations are first ranked according to their task complexity index and grouped into four groups of approximately equal size. Then the average share of time members of these occupational groups spend on the various tasks labelled as "complex" in Table D.1 is computed.

important that activity is relative to others. To address this concern, I draw on a Supplemental Survey from 2012 that precisely details the amount of time a subset of workers spent on the different tasks on a given day. Figure D.1 charts the shares of time spent on the seven abstract/complex activities for different occupational groups. Specifically, I rank occupations according to their task index and group them into 4 equally sized groups. Drawing on the supplemental survey, I then compute the average share of time members of these occupational groups spent on the various tasks. The results show, firstly, that at least in more recent periods each occupational group spends some time on such tasks as organizing or using the computer ("programming"). However, secondly, the fraction of time spent on each of these tasks is multiples greater for the top quartile than for the bottom quartile, without one single task driving the overall increase in the complexity index.