Human Capital at Work: Five Facts about the Role of Skills for Firm Productivity, Growth, and Wage Inequality*

Chiara Criscuolo
IFC

Lukas B. Freund
Columbia University

Peter Gal
OECD

October 22, 2024

Abstract

Firms vary in their production processes, leading to different occupational skill requirements, and they employ workers with varying skill levels. The sorting of workers with heterogeneous skills into firms differing in productivity, size and age matters for both economic efficiency and distributional outcomes. This paper applies a unified measurement approach to comprehensive administrative micro data from Portugal to establish five facts about the relationship between workforce skills, firm productivity and dynamics, and wage differentials: (1) Firms at the productivity frontier do not only rely more on high-skill occupations, they also tend to hire the most skilled workers within each occupation. (2) Such differences in workforce composition statistically explain close to a fifth of firm-level productivity dispersion. (3) Young firms with a high-quality workforce are more likely to experience rapid employment growth. (4) More than half of the large-firm wage premium can be attributed to large firms employing more skilled workers. (5) Working alongside highly skilled colleagues raises wages, and the clustering of talented workers in the same firms contributes about as much to the variance of log wages as worker-firm sorting. Together, these results highlight the significant interaction between human capital factors and firm dynamics.

Keywords: productivity, human capital, labor market sorting, wage inequality
JEL Classification: D24, J24, J31, O47.

*For helpful conversations and feedback we thank Ana Rute Cardoso, Vasco Carvalho, Priscilla Fialho, Paulo Guimarães, Pedro Portugal, Pedro Raposo, Nathalie Scholl and Marta Silva.

Corresponding authors are Chiara Criscuolo (ccriscuolo@ifc.org), Lukas Freund (lukas.beat.freund@gmail.com), and Peter Gal (Peter.Gal@oecd.org).

First version: July 2023.

1 Introduction

Who works for – and with – whom? The answer to this question potentially matters for aggregate productivity and distributional outcomes alike, as individuals vary in skills, while workplaces differ in requirements and advantages – characteristics that correlate with both firms' productivity and workers' wages. Indeed, the sorting of heterogeneous workers into heterogeneous firms receives substantial attention across various fields of economics, from labor to macro and trade.

In practice, workers and firms differ along a variety of dimensions, and data limitations mean that it is challenging to disentangle which characteristics on both sides shape any particular economic outcome. Typically, either firm-level or employee-level datasets have been used to examine the links between skills and productivity, firm size and wages, the role of workforce quality for the fate of startups, or the importance of peer effects for workers' performance and remuneration. This leaves many important questions unanswered. Are the most productive firms distinctive because their production processes rely disproportionately on high-skill occupations or because they recruit the most capable workers within each occupation? Do large employers pay higher average wages partly because they employ more workers in high-paid occupations? What is the role of employee skills in determining which start-ups grow quickly and which remain small? And regarding the impact of one's workplace for wages, is it the employer's characteristics or the attributes of your coworkers that count? Recently, increased data availability has enabled a growing body of papers to address these questions using linked employer-employee data. However, these papers employ a variety of measurement strategies, complicating a unified interpretation.[1]

To inform the debate about these questions, we perform a comprehensive measurement exercise, yielding five facts about the role of human capital at the workplace. These facts describe the sorting of workers, characterized along multiple dimensions of skills, into firms and provide new insights about the implications of such sorting for firm dynamics and worker wages.

Methodologically, we leverage high-quality matched employer-employee panel data from Portugal, enriched with information from income statements. Quite uniquely, these data allow measuring firm-level outcomes and their occupational employment structure as well as workers' job and wage histories over time.[2] We exploit this richness to account for the fact that firms may

[1]To mention just a few examples: Criscuolo *et al.* (2021) focus on the occupational composition of firms and implications for productivity, while studies decomposing the sources of wage inequality often employ a wage fixed-effects approach (Abowd *et al.*, 1999), and the literature on coworker interactions itself uses a plurality of variables to measure worker skills, from wages (Jarosch *et al.*, 2021) through fixed effects (Cornelissen *et al.*, 2017; Hong and Lattanzio, 2022) to years of schooling (Barth *et al.*, 2018; Cardoso *et al.*, 2018; Nix, 2020).

[2]Studying the Portugese economy may also illustrate a more general point for other Southern European countries with similar labor market institutions, and a similar macroeconomic environment (financial and then sovereign debt crises, coupled with high cost of dismissals which could impact worker reallocation).

vary, firstly, in the production processes they rely on and hence different tasks and occupational requirements, while also, secondly, employing workers of different skill levels conditional on their occupation. Concretely, each worker is characterized in three main dimensions: a person wage fixed-effect (FE) that proxies for overall worker human capital; an occupational measure of the cognitive requirements of the tasks performed; and a "relative" FE measure that captures the worker's human capital relative to peers in the same occupation. Further, we characterize firms based on their position in the industry-year specific distribution of labor productivity (value-added per worker) as well as their employment size and age. The two-sided panel structure of the data then allows us to compare the employment composition of firms in terms of the three worker skill measures and to track outcomes in both cross-section slices of data and over time.

Using these rich data, we establish the following five facts.

Fact 1. *More productive firms employ not only a greater share of high-skill occupations but they also disproportionately hire the best workers within each occupation – especially at the frontier and among workers in cognitively intensive occupations.*

Fact 2. *Differences in workforce skill can statistically explain a non-trivial share of variation in firm productivity, and within-occupation quality matters similarly as occupational composition. Yet, a large unexplained share of firm-level productivity dispersion remains.*

Fact 3. *Young firms with a high-quality workforce are more likely to experience fast employment growth.*

Fact 4. *Large firms pay higher wages than smaller firms, but more than half of this premium can be accounted for by worker characteristics, especially within-occupation quality.*

Fact 5. *Having high-quality coworkers is associated with substantial wage gains, and positive assortative matching of workers into teams – more productive workers tends to have better colleagues – emerges as an important contributor to wage inequality. Quantitatively, coworker effects account for around 15% of the total wage variance, on par with the contribution made by worker-firm sorting.*

We next discuss these facts in the context of the relevant literatures. First, we contribute to an extensive literature that empirically describes labor market sorting between heterogeneous workers and firms.[3] Such allocative patterns in the labor market carry significance for efficiency, insofar as complementarities across attributes mean that positive assortative matching raises total output. In addition, these patterns shape the level and firm-level structure of earnings inequality and productivity dispersion. While most studies consider *either* occupation – as an observable characteristic — *or* wage-based worker fixed-effects – proxying for all (observed and

[3]For instance, Doms *et al.* (1997); Abowd *et al.* (1999, 2018); Lindenlaub (2017); Bonhomme *et al.* (2019); Borovičková and Shimer (2020); Lentz *et al.* (2022); Lindenlaub and Postel-Vinay (2023); Freund (2023). For an excellent survey, see Eeckhout (2018).

unobserved) time-invariant worker characteristics – our analysis highlights that a comprehensive account of worker-firm sorting needs to explain both why the occupational composition of more productive firms diverges from less productive firms and why, conditional on occupation, better workers tend to sort into better firms. The quantitative moments collected in this study can, furthermore, aid in disciplining structural, quantitative models that incorporate multiple sources of labor market sorting.

Second, our study helps disentangle the determinants of wages and wage inequality. Regarding the literature on the large-firm wage premium (LFWP),[4] we highlight in particular the importance of differences in within-occupation workforce quality. This finding clarifies that the skill component of the LFWP cannot be traced back solely to non-homotheticities in occupational structure such as an increasing role for non-production workers in large firms. Our results could be consistent with a model in which size is not by itself a source of wage premia but, instead, just a proxy for firm productivity, which in turn is positively correlated with workforce quality due to production complementarities. In addition, we document variation across sectors in the magnitude, curvature and sources of the large-firm wage premium; this variation can inform attempts to build structural models of the relationship between firm size, workforce quality and wages.

Furthermore, our study highlights the quantitative importance of coworkers in determining wages. Specifically, we estimate two-way fixed effects wage regression (Abowd *et al.*, 1999) augmented for coworker interactions (Cornelissen *et al.*, 2017) and use the resulting estimate to perform a decomposition of the variance of log wages. We find that coworker sorting accounts for a similar share of the total variance as worker-firm sorting. While the latter effect has received considerable attention in the literature,[5] coworker effects were typically abstracted from. Our findings thus lend support to recent efforts to empirically estimate and structurally model the sources and economic implications of such effects.[6] In short, wage premia derive not only from who one works *for* (Mortensen, 2003) but also relate to whom one works *with*.

Third, we contribute empirical facts that inform and motivate – or challenge – models of the joint dynamics of workers and firms. First, we find that the sign of the correlation between firm size and workforce quality flips depending on whether we consider cross-firm variation (positive) or within-firm variation (negative).[7] This sign flipping highlights a complex interplay between heterogeneous-firm dynamics and workforce quality composition that is absent from most

[4] See, for instance, Brown and Medoff (1989); Berlingieri *et al.* (2018); Bloom *et al.* (2018).

[5] See, for instance, Abowd *et al.* (1999); Card *et al.* (2013); Alvarez *et al.* (2018); Song *et al.* (2019); Lachowska *et al.* (2020); Lochner and Schulz (2022).

[6] See, for instance, Herkenhoff *et al.* (2018); Nix (2020); Boerma *et al.* (2021); Jarosch *et al.* (2021); Hong and Lattanzio (2022); Freund (2023).

[7] This result echoes Gulyas (2020).

canonical models that either focus on endogenous firm dynamics with decreasing returns to scale (Hopenhayn, 1992; Bilal *et al.*, 2022) or on assortative matching under two-sided heterogeneity (Shimer and Smith, 2000; Hagedorn *et al.*, 2017).[8] Second, we document that the workforce quality of firms is a predictor of future employment growth. This relationship is stronger for young than for old firms, and it is not driven by industry or occupation factors alone. Alongside Mueller and Murmann (2016), Babina *et al.* (2019) and Choi *et al.* (2023), this evidence suggests that differences in employee talent in the very early stages of a company are, at least, candidate element of the ex-ante characteristics determining which young firms succeed, turning into "gazelles," and which fail (cf. Sterk *et al.*, 2021).

OUTLINE. The remainder of the paper is organized as follows. Section 2 describes the micro data and our measurement strategy. We then present, in Section 3, a succession of five key facts. Finally, Section 4 concludes with a discussion of implications for future research.

2 Data and measurement

Our analysis relies on comprehensive, matched employer-employee panel data from Portugal. Using these data we can measure worker and firm characteristics along multiple dimensions and compare outcomes in the cross-section and over time. This section describes the data and the measurement approach we adopt.

2.1 Data description

Our primary data source is the Quadros de Pessoal/Relatório Único (QP, henceforth), which is an annual mandatory census of all employers in Portugal. It provides detailed information on workers' employment status, hourly wage (including bonus payments), gender, education, age, and tenure, among other things, as well as firms' industry, employment size and birth year. By merging in income statement variables from the Informação Empresarial Simplificad, we can, furthermore, compute value-added per worker as a measure of firm-level productivity. These data are comprehensive and of high quality; they comprise the universe of Portuguese employers and employees; and wages are derived from administrative sources, rather than surveys, and

[8]The recent literature has made considerable progress in this direction, though. Eeckhout and Kircher (2018) integrate the classic theory of firm boundaries with a model of sorting between workers and firms, but the model abstracts from firm dynamics or within-firm workforce heterogeneity. The models in Bilal *et al.* (2022) and Elsby and Gottfries (2022), on the other hand, do feature rich firm dynamics in a frictional labor market with on-the-job search but abstracts from worker heterogeneity. Finally, Gulyas (2020) studies firm dynamics with heterogeneous workers, but assumes that firms produce with a linear production technology, thus assuming away sources of decreasing returns such as span of control or taste for variety that are traditionally viewed to be the key force giving rise to a non-degenerate firm size distribution. Also see Elsby and Michaels (2013) and Acemoglu and Hawkins (2014).

they are not top-coded.

After standard data cleaning steps, we impose sample restrictions similar to those commonly employed in the literature (e.g., Card *et al.*, 2018). Specifically, we select persons aged 20-60 living in continental Portugal who are full-time employed as third parties. We also require that they earn at least the mandatory, year-specific monthly minimum wage and have non-missing entries for wage and 2-digit ISCO-08 occupation codes, as we use these two variables extensively. On the employer side, we restrict attention to firms in the manufacturing and non-financial services sectors with at least ten employees and non-missing entries for STAN-A38 industry codes and value-added per worker. We furthermore restrict attention to the largest connected set, as we rely on job-to-job mobility to identify worker and firm fixed effects (as described in detail below). Nominal values such as wages or sales are deflated using the Portuguese consumer price index (2012 = 100). Additional steps and summary statistics for workers are listed in Appendix A.1. Our baseline sample spans the years 2010-2017.[9]

Next we describe how these data allow measuring worker and firm characteristics along multiple dimensions, and then describe howe we construct the firm-level dataset that serves as the main input into our analysis.

2.2 Measuring worker characteristics

Considering worker characteristics first, our measurement approach is informed by the following, simple observation. Our objective is ultimately to understand the determinants of wages, productivity and firm dynamism. These outcomes may vary, across individuals or firms, with the type of tasks performed and with the efficiency with which a given set of tasks is performed. To take Google as a prominent example of a successful company, most employees are primarily tasked with solving cognitively demanding problems like product design or software engineering and, in addition, the company may be able to selectively recruit the most capable workers from a vast pool of applicants for any given role. Employees at less successful peers, such as former competitor AltaVista, may have performed similar tasks but its employees may have proved less expert, on average, at designing, engineering, and marketing a commercially successful product. Finally, employees working for a local corner-shop perform completely different tasks, which on average are likely to be less cognitively demanding, such as operating the till or stocking shelves. These workers can command a lower market wage rate, even if they are exceptionally adept at performing these tasks. In this example, the former two companies thus differ from the latter in terms of task composition, while the employees of those two companies differ from each other in how well they execute a given set of tasks.

[9] Selecting this sample period allows us to use consistent occupation codes without any need for cross-walking.

To empirically operationalize these different dimensions of workforce skill, we use three measures. They capture, respectively, overall time-invariant worker human capital, the cognitive skill requirements of the tasks bundled in a worker's occupation, and the human capital relative to other individuals who are in the same occupation.

Our first measure of worker human capital is the person fixed effect from a two-way fixed effect (FE) log-linear wage regressions (Abowd *et al.*, 1999, henceforth "AKM"). This measure is a commonly used proxy for a worker's time-invariant and partly unobservable human capital. Our specific implementation of the AKM model ensures that our estimates are not affected by limited mobility bias (Andrews *et al.*, 2008). This is accomplished by grouping employers using the pre-clustering approach of Bonhomme *et al.* (2019).[10] After imputing a cluster to each worker-year observation, we estimate the following wage equation[11]

$$w_{it} = \alpha_i + \sum_{k=1}^{K} \psi_k \mathbf{1}(j(i,t) = k) + X_{it}'\beta + \epsilon_{it} \tag{1}$$

where α_i is an individual fixed effect capturing the (time-invariant) component of earnings ability that is transferable across jobs, $\mathbf{1}(j(i,t) = k)$ are dummies indicating which cluster k the employer of worker i in period t, denoted $j(i,t)$, has been assigned to. We associate with each firm j the fixed effect of the cluster to which j belongs. Lastly, X_{it} is a vector of time-varying controls that contains year dummies, a cubic in age and a quadratic in job tenure.

Our second measure uses a worker's occupation as an empirically feasible proxy for the type of tasks performed.[12] [13] Our occupational classification is the ISCO-08 2-digit system, and our data

[10]The idea is to reduce the dimensionality of the estimation problem by clustering similar firms. Clusters are found by solving a weighted k-means problem,

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

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

[11]This estimation is implemented in Stata using the *reghdfe* package (Correia, 2017).

[12]Consistent with this idea, the International Standard Classification of Occupations 2008 (ISCO-08) includes, for each of 427 four-digit occupations, a description of the tasks and duties associated with that group. For example, the description of the 4-digit occupation 2441 "Economists" lists the following (ILO, 2008): "(a) studying, advising on, or dealing with various economic aspects [...]; (b) compiling, analysing and interpreting economic data using economic theory and a variety of statistical and other techniques; (c) advising on economic policy and course of action to be followed in the light of past, present and projected economic factors and trends; d) preparing scholarly papers and reports; (e) performing related tasks; (f) supervising other workers.

[13]This is not to say that the occupational proxy for tasks is perfect by any means. For instance, the task boundaries

contain 39 unique occupations. We follow Criscuolo *et al.* (2021) and partition these occupations into three groups on a measure of the cognitive skills typically required to perform the tasks associated with a given occupation. This measure is computed from the OECD Survey of Adult Skills (PIAAC), which elicits the cognitive abilities of individuals through test scores that capture numeracy, literacy, and problem solving in technology-rich environments. In their multi-country study, Criscuolo *et al.* (2021) average these score results by country and occupation to arrive at a measure of general skill intensity, which is used to rank occupations. We use the same ranking. For our sample specifically, 34% of observations belong to the low-skill occupations, while 56% and 10% pursue medium-skill and high-skill occupations, respectively. The latter group includes, for instance, "Information and communications technology professionals" (ISCO-08 code 25), an example of a medium-skill occupation are "Metal, machinery and related trades workers" (code 72), and "Personal service worker" (code 51) would fall under the low-skill category.

The third and final classification bridges the occupation-based and AKM-based analyses. We take an individual's FE, α_i, and compute the within-occupation FE as $\alpha_i^{-o} = \alpha_i - \bar{\alpha}_{o(i)}$ where $\bar{\alpha}_{o(i)}$ is the average FE among individuals in the occupation o of individual i.[14] Simply put, this third classification allows measuring whether a given worker is a good engineer or a mediocre engineer, a top-ranked chef or a bad chef.

2.3 Measuring firm characteristics

Turning to firms, we focus on three characteristics that are widely studied in the literature on firm dynamics: productivity, size, and age.

Considering productivity first, we follow Criscuolo *et al.* (2021) and group firms into five productivity segments. We proceed as follows. Our measure of productivity is log value-added per employee. Firms are then grouped into deciles of the annual productivity distribution within STAN-A38 industries. Groupings are blanked if less than 10 firm-level observations are available within STAN A38 x year cells. The "frontier" then comprises firms in the top decile, "medium-frontier" includes the 7th-9th decile, "medium" the 5th and 6th decile, low-medium the 2nd-4th decile, and lastly, "laggards" are firms in the first decile.

Firm size is measured based on the total employment count indicated in the administrative records. As such, firm sizes also accounts for the number of employees working part-time or below minimum-wage (that is, employees that we dropped from the worker panel on which we

of an occupation may be redrawn in response to technological innovations, and two workers of same occupation may perform different tasks at different firms, e.g., they may be more specialized in larger firms (Chaney and Ossa, 2013; Adenbaum, 2022).

[14] To be precise, since an individual may switch occupation, α_i^o is time-specific even if α_i is not. The time subscript is suppressed for ease of exposition.

estimate the AKM models). We consider three main groups: small (10-49 employees), medium (50-249), and large (≥250).

The age of a firm is likewise taken straight from the administrative data. Following Haltiwanger *et al.* (2013), we distinguish between young firms, that are less than 10 years old, and mature ones, which are 10 years and above.

2.4 Constructing the firm-level dataset

The majority of analyses are based on an annual firm-level panel, which is constructed from the worker-level records and the above skill measurements as follows. At the firm-year level we construct the share of employment accounted for by workers of different types. This is straightforward for the occupational classification. To obtain a comparable employment share representation of both FE based measures, we bin workers into three tertiles and then consider for each firm what share of workers belong to the bottom, medium, or top tertile of the overall and occupation-specific FE distributions.

In addition, we also consider a continuous measure based on the worker FEs (WFE). First, we calculate for each firm-year the average WFE, denoted $\bar{\alpha}_{jt}$. Second, we decompose $\bar{\alpha}_{jt}$ into two components that capture occupational composition and within-occupation quality as follows,

$$\bar{\alpha}_{jt} = \underbrace{\sum_{o \in O} \frac{N_{jot}}{N_{jt}} \bar{\alpha}_o}_{\text{occupational quality}\,:\bar{\alpha}^{o}_{jt}} + \underbrace{\frac{1}{N_{jt}} \sum_{i} \alpha_i^{-o} \mathbf{1}\{j(i,t) = j\}}_{\text{within-occupation quality}\,:\bar{\alpha}^{-o}_{jt}}, \tag{2}$$

where N_{jt} is the number of worker observations associated to firm j in year t, $N_{j,o,t}$ the nuber of workers in occupation o (where O is the set of occupations), $\mathbf{1}\{j(i,t) = j\}$ is an indicator for whether worker i is employed at firm j in period t, and $\bar{\alpha}_o$ is the average FE of workers in occupation o.

Finally, we restrict ourselves to firms with non-missing observations for every one of the following variables: labor productivity, employment shares, average FE(s), average log wage, and STAN-A38 industry.[15] The final sample comprises 39,648 unique firms and 202,938 firm-years. Table 1 provides a set of summary statistics. Appendix table A.1 describes the worker panel underlying the AKM estimation.

[15]As a result of these restrictions, the (person-weighted) average of $\bar{\alpha}_j$ is no longer exactly equal to zero.

	Mean	Std.
# firm-year obs.	202,938	-
Log value-added per worker.	10.00	0.64
Log employment	3.22	0.87
Firm age	20.25	15.17
Avg. age (years)	40.92	4.71
Avg. log real hourly wage	1.73	0.39
Female share (%)	0.37	0.30
Share in low-skill occupations (%)	0.34	0.35
Share in medium-skill occupations (%)	0.56	0.35
Share in high-skill occupations (%)	0.10	0.22
Share movers (%)	0.03	0.07

Table 1: Summary statistics for the firm-level panel

Notes. This table provides summary statistics for the firm-level annual panel (2010-2017). Workforce characteristics and outcomes (bottom eight rows) are weighted by underlying person-years. Real values are in 2012 Euros.

3 Five facts about skills, firm dynamics, and wages

Using the data and measurement strategy just set out, in this section we establish five facts about the economic interactions between workforce skills and firm performance, as well as other firm characteristics.

3.1 The anatomy of worker-firm sorting

We start by describing how workers with different skills systematically sort into firms that vary in terms of productivity. A particular goal is to understand in how far firms at the productivity frontier are distinctive in terms of their occupational composition as compared to the quality of workers conditional on occupation.

Beginning with a simple overview, the top three rows of Table 2 indicates the employment shares at firms belonging, respectively, to the laggard, medium, and frontier within-industry productivity groups. The employment shares are defined in terms of our three alternative definitions of worker groups and weighted by firm-level person-year observations. Thus, within each firm group (rows) and skill definition (three main columns) the shares sum to one.

The headline finding emerging from these summary statistics is that the workforce of frontier firms is more highly skilled in two dimensions. These firms employ a relatively greater share of high-skill occupations, and within each individual occupation they tend to hire the best workers.

Going into greater detail, and considering first the comprehensive classification of workers

		FE groups			Occ. groups			Within-occ. FE groups		
		L	M	H	L	M	H	L	M	H
Prod.	Laggard	0.52	0.35	0.13	0.45	0.51	0.04	0.41	0.38	0.22
	Medium	0.38	0.39	0.24	0.33	0.60	0.08	0.39	0.34	0.27
	Frontier	0.19	0.24	0.57	0.32	0.51	0.17	0.23	0.28	0.50
Size	10-49	0.39	0.37	0.24	0.39	0.54	0.08	0.37	0.36	0.27
	50-249	0.35	0.31	0.34	0.36	0.53	0.11	0.35	0.32	0.33
	>250	0.28	0.33	0.40	0.27	0.61	0.12	0.29	0.32	0.39
Age	Young	0.34	0.35	0.31	0.37	0.51	0.12	0.31	0.37	0.32
	Mature	0.33	0.33	0.34	0.33	0.57	0.10	0.34	0.33	0.34

Table 2: Employment composition across 3 skill levels by firm characteristics

Notes. This table describes how firms' employment composition across skill levels (L = low; M = medium; H = high) varies with firm productivity ("Prod."), size, and age. The three columns indicate, respectively, the shares of workers in the bottom, medium and top tertile of the economy-year wide worker fixed effect distribution ("FE groups"), the shares of low-, medium, and high-skill occupations ("Occ. groups"), and share of workers in the three tertiles of the occupation-year specific worker fixed effect distribution ("Within-occ. FE groups"). Firm-level observations are weighted by the underlying number of person-year observations, and te row "Observations" indicates the effective number of underlying person-years. Shares may not sum to one due to rounding.

in terms of economy-wide FEs ("classification 1"), we see that the economy is characterized by substantial positive assortative matching between firms and workers. On average, 57% of the workforce of frontier firms belongs to the top tertile of workers. The analogous figure is merely 24% for medium firms and 19% for laggards. Conversely, the share of workers in the lowest tertile ranges from 52% for the least productive firms to 13% for frontier firms. Moreover, much of the substitution of worker types that occurs as we move along the firm productivity distribution appears to involve the low- and high-types; the employment share of the middle group of workers is relatively more similar across firm productivity groups (ranging from 24% to 39%).

Next, consider how worker-firm sorting operates along occupational lines ("classification 2"). In line with Criscuolo *et al.* (2021), we find that firms at the productivity frontier employ relatively more workers in high-skill occupations than their less productive peers, with correspondingly fewer employees in low-skill occupations. The share of high-skill occupations in frontier firms is roughly twice that at the typical median performer (17% vs. 8%, a gap of 9 percentage points), with the gap rising to 13 percentage points for the "laggard" firms in the bottom decile of the within-industry productivity distribution. Conversely, 45% of employees at laggards are in low-skill occupations compared to 32% in frontier firms.

Beyond occupational composition, the second dimension of worker-firm sorting is that more productive firms also tend to hire more of the best workers within their respective occupation

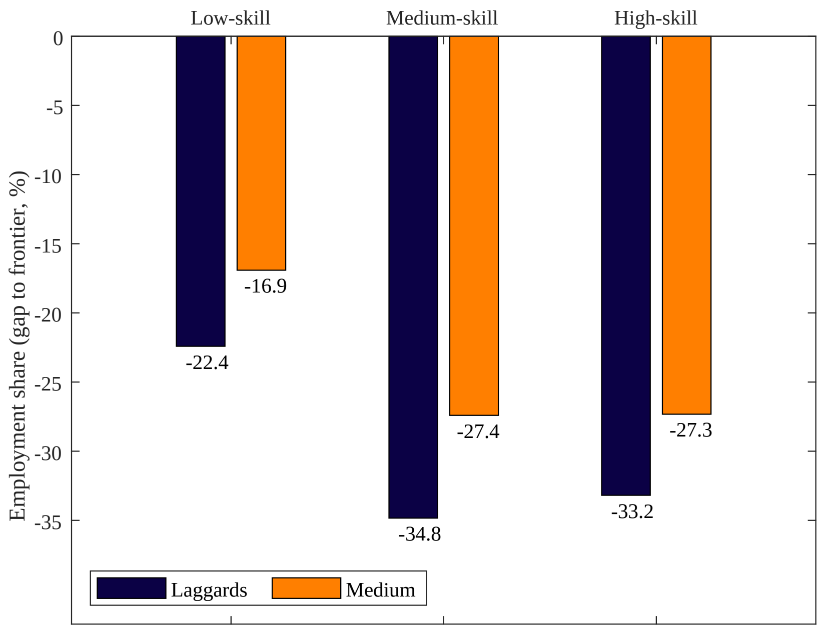

Figure 1: Within-occupation quality gaps to frontier, by occupational groups

Notes. This figure depicts the gap in the employment share of top tertile workers relative to the group of frontier firms, defined separately for each occupation and then taking an (unweighted) average within the three groups of low-, medium- and high-skill occupations. The gap is shown for laggard and medium firms.

("classification 3"). The third main column of Table 2 reveals that frontier firms employ a greater share of workers in the top tertile of the occupation-specific worker FE distribution. Compared to classification 1, the between-firm gaps shrink by less than half. Whereas the high-type gap between frontier and medium-firms using classification 1 is 33 percentage points (44 points between frontier and laggards), this shrinks to 23 (respectively 28) when considering within-occupation differences, as per classification 3.

A natural follow-up question is whether the extent to which higher-ability members of a given occupation tend to work in more productive firms varies with the skill intensity of tasks performed. To answer this question, we consider medium and laggard firms and compute for each individual occupation the average gap to the frontier in terms of the share of workers in the top tertile of their occupation-specific FE distribution. We then take the average across these occupations within low-, medium- and high-skill occupation groups. The results are plotted in Figure 1.

Two main results stand out from Figure 1. First, frontier firms consistently employ a greater share of top-tertile workers within each group of occupations. This means that the aggregate result is not driven by a small subset of occupations. Second, though, the magnitude of this gap tends to be larger for high- and medium-skill occupations than for low-skill occupations. For example, the frontier-laggard gap is 22% among low-skill occupations, rising to 33% among

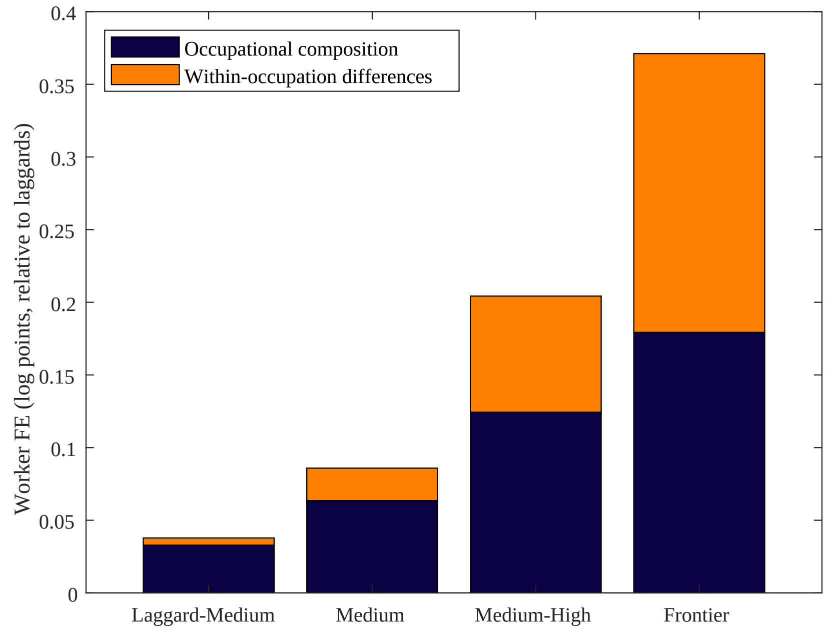

Figure 2: Firm-level variation in average worker FE: occupations vs. within-occupation

Notes. This figure decomposes the gap in average worker FE across firm productivity groups into two components, using equation (2): differences in occupational employment shares and variation in FEs conditional on a worker's occupation. The firm-level gaps are defined in log points relative to firms in the "laggard" groups of the industry-specific productivity distribution.

high-skill occupations.

In light of these observations, we next disentangle how much of the gap across firms in the average worker FE is driven by, respectively, differences in occupational composition and the within-occupation, worker-specific component of skills. To this end, we decompose the gap in the average WFE between laggard and the four other firm productivity groups using equation (2). This approach also enables us to study if and how the relative weight of the two components shifts as we move farther away from the bottom toward the top of the firm productivity distribution.

Figure 2 plots the resulting decomposition and illustrates that both factors play an important role. For example, of the 17 log point difference between medium-high and laggard firms, 9 are due to occupational composition and 8 arise from within-occupation differences. As the two bars to the right indicate, occupational composition matters somewhat less when considering the gap between frontier and medium-high firms as compared to the gaps among firms in the bottom half of the distribution. While the occupational composition of medium-high and frontier firms is similar (as indicated by their average occupation-specific FEs), the typical employee of frontier firms has a significantly higher FE conditional on their occupation. This pattern is consistent, for example, with the idea that the most productive firms assemble "superstar teams" of the most productive workers within each occupation (Freund, 2023). We summarize as follows:

Fact 1. *Firms at the productivity frontier disproportionately employ workers in high-skill* *occupations, relative to other firms, and they tend to hire the best workers for each job. Both occupational composition and within-occupation quality differences contribute in roughly equal measures to between-firm gaps in workforce quality; within-occupation differences are particularly important in differentiating the frontier from follower-up firms, and more generally within high-skill occupations.*

FIRM SIZE AND AGE. To close this section, we discuss how the workforce composition varies along the lines of two other firm characteristics that the literature has commonly considered, namely size and age. Regarding the former, worker selection is, for instance, one candidate explanation for the observation that larger employers pay higher wages. Indeed, as Table 2 shows, in our data we see that larger firms employ a higher share of high-skill occupations, and a notably greater share of the most skilled within each occupation. In terms of magnitudes, the gap between large and small firms is more modest than the gap between frontier and laggard firms, though. This holds true under any of the skill definitions.

Interestingly, the sign of the relationship between firm employment size and workforce quality flips depending on whether we exploit variation *within* a firm over time or compare *across* firms. Table 3 reports the result from estimating regressions of the average worker FE – considering our three firm-level measures $\bar{\alpha}_{jt}$, $\bar{\alpha}^{o}_{jt}$, and $\bar{\alpha}^{-o}_{jt}$ – under alternative specifications. Whereas firm size positively predicts worker quality within industry-year across firms, the two variables are negatively related when including a firm fixed effect in the regression. The estimated elasticity of the average FE – by construction, the FEs are in logs – to firm size is 0.0397 *across* firms (column 1), but -0.0186 when estimated on *within*-firm variation (column 2). These opposite signs are robust across the different workforce quality measures (columns (2)-(3) and (5)-(6)). This means that, even though large firms employ better workers, as firms grow, they "downgrade" the quality of their workforce (and "upgrade" as they shrink).

Finally, we turn to firm age. Brown and Medoff (2003) called attention to the possibility that worker sorting across firms of different ages may help explain why pay in younger firms tends to be lower, on average. In Table 2, though, we do not observe major differences in workforce composition across firms of different ages.[16] In Section 3.3 we study the role of skill differences *among* young firms in predicting their future employment dynamics, and in Section 3.4, we consider the contribution of different workforce characteristics in accounting for the large-firm pay premium. Before doing so, we briefly go deeper into the relationship between worker skills and firm productivity.

[16]This finding differs from what Babina *et al.* (2019) document for the U.S., for instance, indicating cross-country differences.

	(1) $\bar{\alpha}_{jt}$	(2) $\bar{\alpha}^{o}_{jt}$	(3) $\bar{\alpha}^{-o}_{jt}$	(4) $\bar{\alpha}_{jt}$	(5) $\bar{\alpha}^{o}_{jt}$	(6) $\bar{\alpha}^{-o}_{jt}$
Log employment	0.0397***	0.0136***	0.0260***	-0.0186***	-0.00589***	-0.0127***
Industry x Year FEs	Yes	Yes	Yes			
Firm FEs				Yes	Yes	Yes
Year FEs				Yes	Yes	Yes
Observations	202,938	202,938	202,938	199,830	199,830	199,830
R-squared	0.259	0.358	0.074	0.886	0.873	0.803

Table 3: Regression of workforce quality on log employment, across and within firms

Notes. This table describes the estimation results for a regression of workforce quality on firm size. Specifically, it lists the estimated coefficient for log employment size in separate regressions using three different firm-level measures of workforce quality as dependent variables: the average worker FE, $\bar{\alpha}_{jt}$ (columns 1 and 4); the average occupation FE, $\bar{\alpha}^{o}_{jt}$ (columns 2 and 5); and the average within-occupation FE, $\bar{\alpha}^{-o}_{jt}$ (columns 3 and 6). The first three columns reflects estimates off of cross-firm variation, the final three columns are estimated on within-firm variation. Standard-errors clustered at the firm-level are indicated in parentheses. *** p<0.001, ** p<0.01, * p<0.05.

3.2 Workforce characteristics and productivity dispersion

Our data reflect the well-known fact that firm productivity is widely dispersed, even within narrowly defined industries. Specifically, the average annual standard deviation of log labor productivity is 0.64; and the typical frontier firm is more than three times as productive as the typical median performer, with the gap more than doubling relative to laggard firms.[17] If frontier firms systematically employ better workers, a natural question is to what extent these differences can account for productivity dispersion. While the *causal* relationship between worker skills and firm productivity is well beyond the scope of this study, this section statistically evaluates the share of firm-level productivity dispersion that is accounted for by variation in workforce skills (cf. Fox and Smeets (2011), Criscuolo *et al.* (2021) and the review in Syverson (2011)).

A simple approach to evaluate the explanatory power of workforce skills for productivity dispersion is to compare the standard deviation of labor productivity before and after controlling for different skill measures in a linear regression. Table 4 indicates the results obtained from such an exercise. Starting from the top, each row indicates the firm-level dispersion in productivity under different specifications. The first two rows tell us that the standard deviation of log labor productivity drops from 0.64 to 0.57 when taking out industry-year fixed effects. Next, we control

[17] We report the unweighted averages across industries. For instance, if y_{jst} denotes the average log labor productivity of firm j in industry s in year t, then the standard deviation reported refers to $\frac{1}{S}\frac{1}{T}\sum_{s=1}^{S}\sum_{t=1}^{T} sd_j(y_{jst})$. Similarly, if y_{gjt} denotes the (unweighted) average log labor productivity of firms in productivity group $g \in \{1,2,3,4,5\}$, where e.g. $s = 5$ refers to the frontier, in industry s in year t, then the leader-laggard gap reported refer to $\exp\left(\frac{1}{S}\frac{1}{T}\sum_{s=1}^{S}\sum_{t=1}^{T}\left(y_{5st} - y_{1st}\right)\right)$. Note that we use log differences and the exponentiation is taken after the averaging.

	Std. Dev.	Relative to Raw	Relative to Baseline
Raw	0.638	-	-
Baseline	0.568	0.89	1.00
Control for occ. shares	0.549	0.86	0.97
Control for avg. occ WFE	0.533	0.83	0.94
Control for avg. within-occ. WFE	0.542	0.85	0.95
Control for avg. WFE	0.480	0.75	0.84
Control for avg. occ WFE + avg. within-occ. WFE	0.474	0.74	0.83

Table 4: Dispersion of labor productivity before and after controlling for workforce skills

Notes. The three main columns indicate the standard deviation of productivity (residuals). The "baseline" corresponds to the dispersion of labor productivity after taking out industry-year fixed effects. Subsequent rows add, separately, the indicated controls in addition to industry-year fixed effects.

for shares in high-, medium- and low-skill occupations; the average occupation WFE and its square; the average within-occupation WFE and its square; their combination; and the average WFE and its square.[18]

In brief, this exercise suggests that relative to a baseline without workforce skills, the skill measures generate a meaningful reduction of up to 17%.[19] Considering occupational employment shares only significantly understates the statistical importance of workforce skills in explaining firm-level productivity dispersion. It is key to also consider within-occupation skill differences and, crucially, the interaction between occupational composition and within-occupation differences. In summary:

Fact 2. *Differences in workforce skill can statistically explain a non-trivial share of variation in firm productivity, and within-occupation quality matters similarly as occupational composition. Yet, a large unexplained share of firm-level productivity dispersion remains.*

3.3 Workforce characteristics and firm dynamics

In addition to gaps in productivity, another key dimension in which there are well-documented and pronounced differences among firms is employment growth and job creation. A large literature on business and labor market dynamism highlights how these differences are tied to observable characteristics, notably firm age and productivity. Specifically, young firms are key

[18]A related approach discussed by Syverson (2011) considers how the (adjusted) R-squared value changes as we include different, or more, regressors. On this approach, too, we see that the share of variation explained by the regressors is greatest when skills are measured by the rank in the economy-wide distribution of fixed effects; and that the within-occupation measure of workforce quality explains at least as much as the occupational employment shares.

[19]Similarly, Fox and Smeets (2011) find that in Danish data, the 90/10 ratio of TFP quantiles declines by 18% from adding a detailed set of human capital and wage bill controls.

drivers of net job creation, far in excess of their share of total employment (e.g., Haltiwanger *et al.*, 2013). At the same time, there is pronounced dispersion among these young firms in terms of survival rates as well as employment growth rates conditional on survival.[20] Most of these differences appear to result from ex-ante heterogeneity rather than persistent shocks post-entry (Sterk *et al.*, 2021). Yet relatively little is known about the sources of such heterogeneity. One candidate source emerging from the literature is variation in the quality of young firms' workforce. This hypothesis is typically examined with a focus on company founders (Lazear, 2004; Karmakar *et al.*, 2021) or founding teams (Choi *et al.*, 2023). In this section, we instead use administrative evidence to provide complementary, reduced-form evidence relating young firms' workforce composition to future employment dynamics.

To this end, we extend the approach of Babina *et al.* (2019) and estimate the following regression:

$$l_{j,t+s} = \beta_0 + \beta_1 \hat{x}_{j,t} + \beta_2 l_{j,t} + \eta_{\text{industry}(j)} + \eta_{\text{birth-year}(j)} + \epsilon_{j,t} \tag{3}$$

where the dependent variable, $l_{j,t+s}$, is log employment of firm j in period $t+s$ for $s>0$, the regressor of interest is a measure of j's workforce quality in t, and we control for period-t log employment as well as industry and birth-year fixed effects. Firm employment is measured by the number of worker-year observations per firm and, thus, includes only full-time employees satisfying our sample restrictions. By construction, this regression is run on the subsample of firms that we observe as surviving until $t+s$.

We estimate this regression separately for six firm-age groups and three alternative measures of workforce quality, $\hat{x}_{j,t}$: the average WFE ($\bar{\alpha}_{j,t}$), the average occupation-WFE ($\bar{\alpha}^o_{j,t}$), and the average within-occupation WFE ($\tilde{\alpha}_{j,t}$). We consider a five-year horizon, $s=5$. Table 5 presents the coefficient estimates for β_1 by firm age category (columns) and workforce quality measure (rows).

The regression estimates in Table 5 indicate that workforce composition, as measured by the average WFE, positively predicts future firm employment growth in a statistically significant way. Three additional observations stand out. First, across workforce quality measures, the point estimate for the relationship between current workforce quality and future employment growth is greater for younger firms than for old firms. According to the first row, it is almost twice as large for firms of age 0-4 than for firms aged 30-49. One potential explanation is that young firms are more likely to be farther away from their optimal size. An alternative possibility is that the quality of early-stage employees plays an outsized role, because they create persistent organizational capital (Choi *et al.*, 2023) or influence the quality of subsequent hires (Freund, 2023).

[20] Table B.1 in the appendix quantifies how the rates of job creation, job destruction and net job creation vary across firm age groups.

$\ln$(Employment in 5 years)	(1) 0-4y	(2) 5-9y	(3) 10-19y	(4) 20-29y	(5) 30-49y	(6) >50y
Avg. WFE	0.261***	0.246***	0.189***	0.196***	0.134**	0.0973
	(0.0694)	(0.0533)	(0.0308)	(0.0347)	(0.0445)	(0.0764)
Avg. occupation-WFE	0.0285	0.0955	0.0360	0.0108	0.0268	-0.0211
	(0.0793)	(0.0576)	(0.0380)	(0.0432)	(0.0601)	(0.103)
Avg. within-occ. WFE	0.292***	0.229***	0.215***	0.279***	0.184***	0.183
	(0.0804)	(0.0543)	(0.0336)	(0.0412)	(0.0511)	(0.0999)
Observations	4669	7349	17160	12256	8410	2316

Table 5: Regressions of 5-year ahead log employment on workforce quality, by firm age

Notes. This table indicates the estimated values for β_1 in a sequence of regressions of five-year ahead employment on current log employment as well as alternative measures of workforce quality. The regression is run separately for 3 different measures of worker quality – represented by rows – and 6 different firm age groups – indicated by columns. All regressions include industry and firm birth-year fixed effects, in addition to controlling for current log employment. Standard errors are clustered at the firm-level and indicated in parentheses. *** p<0.001, ** p<0.01, * p<0.05

The second observation is that the magnitude of the estimated coefficient for young firms is economically meaningful. Among firms in the age-group 0-4y, a one standard-deviation increase in the average WFE predicts a rise in employment of 5.5 log points. The final observation concerns the difference in relevance between occupational composition and within-occupation worker quality. According to the baseline estimates presented in Table 5, it is the latter but not the former that is positively associated with future firm size. In Appendix B.2, though, we extend the sample under consideration by an additional six years, covering 2004-2017, and repeat the exercise. In addition to the magnitude of the estimated coefficient β_1 being considerably greater, the average occupation WFE is now also positively associated with five-year ahead employment growth in a statistically significant manner. In summary:

Fact 3. *Young firms with a high-quality workforce are more likely to experience fast growth in the future.*

While decidedly correlational in nature, Fact 3 is consistent with a range of explanations (cf. Babina *et al.*, 2019). One possibility is that some firms are endowed with inherently higher growth potential – perhaps reflecting the quality of an initial product idea — or technical efficiency. These high-potential firms are able to attract more resources and they hire a higher-quality workforce due to production complementarities. If that is the case, it is possible that frictions which limit access to the relevant skills may cause the high potential of some start-ups not to be realized. Another possibility is that ex-ante heterogeneity in intrinsic potential is actually of limited importance and, instead, it is the quality of a firm's initial workforce that by itself determines the future performance of the business. In such a scenario, assortative matching between entrants and workers would play a lesser role in terms of efficiency (though complementarities across coworkers or co-founders may still matter). Differentiating between these explanations – with the help of explicitly causal empirical designs or structural, quantitative models – would be important to understand which policies are conducive to fostering business dynamism and growth.

3.4 Firm size and wages

While young firms – the focus of the preceding section – are typically small, a different literature with long tradition instead concentrates on the well-documented observation that larger companies pay higher wages, especially in manufacturing. In the search for explanations, the interplay between firm size and workforce characteristics plays a prominent role, as the positive relationship between firm size and pay may be partly due the fact that the employees of large firms tend to be more skilled (cf. Section 3.1). On the other hand, it may also reflect that high-productivity firms will be larger, earn higher rents, and share some of them with their workers (Lucas, 1978; Burdett and Mortensen, 1998; Card *et al.*, 2018). Of course, these explanations are not mutually exclusive and may interact, as high-productivity firms may be larger and hire skilled workers due to production complementarities as opposed to an inherent interaction with firm size. In this section, we revisit the role of workforce quality differences in explaining the large-firm wage premium (LFWP), asking in particular which dimensions of worker skills are pertinent.

In a first step, we perform a set of simple firm-level regressions of the average log wage on log employment, introducing alternative controls step by step. Table 6 summarizes the estimation results. Columns (1) and (2) indicate that firms with twice as many employees pay wages that are on average 3.8% higher, shrinking to 2.6% when controlling for industry. When controlling for the average WFE, however, this premium is estimated to be close to zero. This mediating effect of workforce skills appears to be driven by within-occupation quality differences, as controlling for the average occupation WFE leaves the LFWP intact.

However, this first finding does not tell us that workforce quality explains the LFWP, as worker FEs may be correlated with other pay relevant firm characteristics.[21] In particular, we saw in Section 3.1 that high-type workers tend to sort into high-productivity firms.

To disentangle the contributions of worker quality and firm pay premia we next rely on the fact that, in equation (1), log wages are additively separable into worker and firm components, α and ψ. Consequently, the coefficients in regressions of AKM components on firm size (groups) mechanically add up to the total coefficient of log wage on firm size (Bloom *et al.*, 2018). Moreover,

[21] Consistent with this notion, the regression coefficient on worker FEs exceed unity.

$\ln$(wage)	(1)	(2)	(3)	(4)	(5)
Log employment	0.0431***	0.0325***	0.00216*	0.0289***	0.00782
	(0.0127)	(0.00904)	(0.00123)	(0.00529)	(0.00596)
Avg. worker FE			1.635***		
			(0.0129)		
Avg. occupation worker FE				1.509***	
				(0.0449)	
Avg. within-occupation FE					1.522***
					(0.0319)
Year FEs	Yes				
Industry x Year FEs	No	Yes	Yes	Yes	Yes
Observations	4,947,499	4,947,499	4,947,499	4,947,499	4,947,499
R-squared	0.050	0.373	0.952	0.606	0.677

Table 6: The large-firm wage premium and workforce characteristics

Notes. This table presents estimates from a firm-level regression of the average log wage on log employment as well as alternative measures of workforce quality. Firm-level observations are weighted by associated person-years. Standard errors are clustered at the firm-level and indicated in parentheses. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

per equation (2), we can split the average worker FE into occupational and within-occupation components.

Figure 3 displays the resulting decomposition of the firm pay premium, considering the difference in log wage and AKM components relative to the group of small firms with 10-19 workers.[22]

Three findings stand out. First, unconditionally the average wage is almost 30 log points higher in a firm with 500+ employees compared to a small firm with 10-19 employees. Interestingly, average pay peaks at firms with 250-499 employees, declining again when moving to mega firms.[23] Second, the firm-pay AKM component explains around 40% of the pay premium at each point in the size distribution. To understand the origins of the LFWP, it is thus crucial to account for the positive correlation between worker and firm components. Third, the within-occupation WFE component accounts for a larger share of the LFWP than differences in occupational composition. Indeed, the former is monotonically increasing across firm size bins, whereas it is the decline in the latter in the very largest firms that drives the non-monotonicity at the top end of the size distribution. We summarize:

[22]The decomposition is shown for the entire economy but looks similar when evaluated at and then aggregated from the industry level.

[23]This patterns matches evidence for the U.S. over the period 2007-2013 presented in Bloom *et al.* (2018) and for the services segmented in Berlingieri *et al.* (2018).

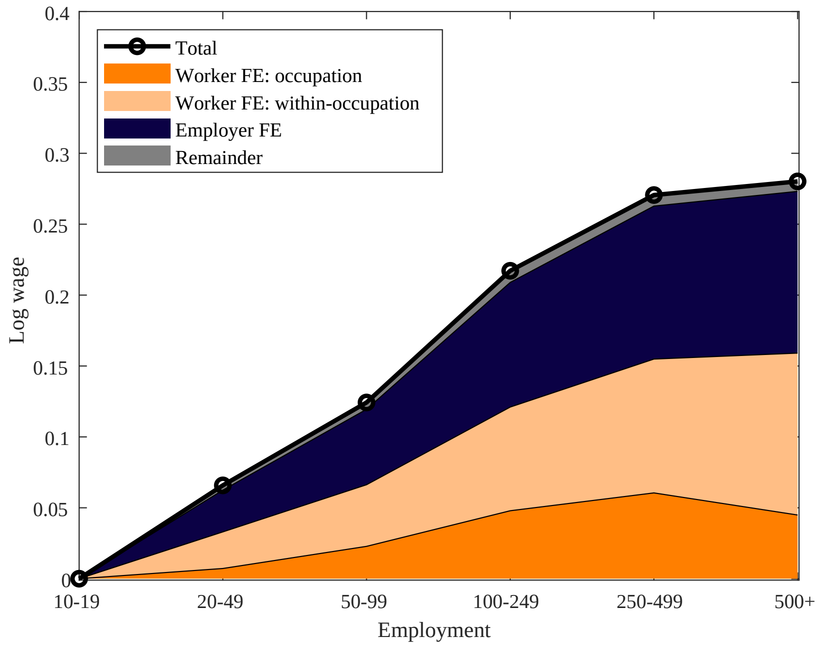

Figure 3: Decomposition of the large-firm wage premium into AKM components

Notes. This figure decomposes the wage premium relative to firms of size 10-19 for 5 other size groups into its AKM components, estimated based on equation (1).

Fact 4. *Large firms pay higher wages than smaller firms, but this premium shrinks by more than half when controlling for time-invariant worker characteristics, especially within-occupation quality.*

Appendix B.3 examines the role of sectoral heterogeneity in accounting for the aforementioned non-monotonicity. We quantify and decompose the LFWP in four different sectors, distinguishing between manufacturing and non-financial market services as well as knowledge-intensive and less knowledge-intensive industries. The picture emerging from this analysis is that the monotone relationship between wages and firm size is intact in all sectors except low-knowledge intensity services, a sector comprising about half of Portuguese employment, including in industries such as wholesale and retail trade. Only in that sector do the very largest firms pay relatively less, reflected in a combination of lower firm pay premia and worker characteristics. Interestingly, these changes mirror a decrease in labor productivity for 500+ firms relative to the size category below. Jointly, these results underline the challenge of creating “good jobs” for all workers in economies that have shifted from the traditional paradigm of a manufacturing economy towards a model dominated by (low knowledge-intensity) services.

3.5 Coworker quality and wages

Thus far, we focussed on the interaction between heterogeneous firms and heterogeneous workers. But viewed from the individual worker's perspective, an important reason to care about the workforce composition of their employer, over and beyond their own contribution, is that the quality of their coworkers may influence their own productivity. This perspective is indeed natural — most people work in teams and interactions with coworkers are an essential part of their job – and has been at the heart of a series of recent papers that revisit the sources of wage inequality and human capital accumulation (Herkenhoff *et al.*, 2018; Nix, 2020; Jarosch *et al.*, 2021; Freund, 2023). It is, however, absent from the log-linear wage specification underlying the standard AKM model. In this section, we therefore study the influence of coworker quality on wages using reduced-form tools. Throughout, the average quality of worker i's coworker group in their firm at time t is proxied by the time-t average coworker fixed effect, $\bar{\alpha}_{-i,t}$.[24]

In the spirit of Cornelissen *et al.* (2017), we estimate a coworker-augmented version of the AKM model:[25]

$$w_{it} = \alpha_i + \gamma \bar{\alpha}_{-i,t} + \sum_{k=1}^{K} \psi_k \mathbf{1}(j(i,t) = k) + X'_{it}\beta + \epsilon_{it} \tag{4}$$

where γ is the coefficient of interest. It quantifies the elasticity of the wage with respect to average coworker quality. We experiment with three alternative definitions for the coworker group ("team"): every employee in the same firm, in the same firm and occupation, or in the same firm and one of seven hierarchical layers (see, e.g., Caliendo *et al.*, 2020). We estimate this model using the iterative method proposed by Arcidiacono *et al.* (2012) and implemented efficiently by Cardoso *et al.* (2018).

Panel A of Table 7 summarizes the regression estimates for γ alongside the sample standard deviation of the average coworker quality. Depending on the team definition, the estimated value of γ is between 0.283 and 0.389, which means that a 10% increase in the quality of coworkers, as measured by $\bar{\alpha}_{-it}$, is associated with an increase in the wage by between 2.8 and 3.9%. Taking into account the sample variability of $\bar{\alpha}_{-it}$, a one standard-deviation increase in coworker quality is predicted to raise wages by an economically meaningful amount of 8.3%, 8.0% or 8.6%, respectively.[26]

Next, we examine the implications of such coworker wage spillovers for the overall dispersion

[24] Note that while the individual fixed effects are, by construction, time invariant, the average coworker FE is time-varying due to changes in the coworker group's composition.

[25] Relative to these papers, and to deal with limited mobility bias, we proceed as described in Section 2 by estimating cluster fixed effects as opposed to individual-firm fixed effects.

[26] These estimates are in a similar ballpark as those presented by Cardoso *et al.* (2018) and Hong (2022) and somewhat larger than those provided by Battisti (2017) and Cornelissen *et al.* (2017).

		Baseline	Coworker (1)	Coworker (2)	Coworker (3)
Panel A.	$\hat{\gamma}$	-	0.389	0.281	0.271
Estimates	Std. $\bar{\alpha}_{-i}$	-	0.215	0.281	0.271
Panel B. Wage variance decomposition (%)	Worker effects	58.4	53.6	44.9	43.4
	Firm effects	9.0	3.5	6.3	6.1
	Worker-firm sorting	25.9	13.8	18.3	17.8
	Coworker spillovers		2.7	2.5	2.9
	Coworker sorting		12.9	15.1	16.7

Table 7: Regression estimates and variance decomposition in an augmented AKM model

Notes. This table summarizes the estimation results for the coworker-augmented AKM model alongside a decomposition of the variance of wages into the resulting fixed effect variance and covariance terms. The model is estimated for 3 different specifications of the coworker group: firm-year (1), firm-occupation-year (2), firm-layer-year (3). The variance decomposition is also performed for the baseline AKM model.

in wages. To what extent these spillovers contribute to overall wage inequality critically depends on the degree of positive assortative matching across coworkers, i.e., do high-FE workers tend to be matched together with other high-FE into the same workplaces.

We use the estimates of the fixed effects and γ to decompose the variance of log wages into five components: in addition to the "conventional" terms – the variance of worker FEs and firm/cluster FEs, respectively, and twice the covariance between worker and firm/cluster FEs (e.g., Song *et al.*, 2019) – allowing for coworker interactions introduces two additional terms. These are the variance of coworker effects, $\mathrm{Var}(\gamma\bar{\alpha}_{-it})$, and (twice) the covariance of workers and coworker effects, $2\mathrm{Cov}(\alpha_i, \gamma\bar{\alpha}_{-it})$.[27]

As a reference point, the first column of Panel B shows the variance decomposition arising from the estimation of the standard AKM model in equation (1). Worker effects explain more than half of total wage dispersion, firm effects close to ten percent, and worker-firm sorting more than a quarter.

In the augmented model, on the other hand, the variance of worker effects accounts for between 43%-54%, less than in the baseline model, and the contribution of firms – through both variance and covariance terms – likewise declines. Instead, the positively assortative sorting of high-type workers into the same employers and teams, emerges as another key contributor to the total variance of log wages. Quantitatively, the covariance of worker fixed effects and and coworker spillovers contributes around 15%, almost on par with the contribution made by worker-firm sorting.[28]

[27] The decompositions shown abstract from the additional regressors by suppressing their contributions through variance and covariance terms. In addition, we deliberately omit a term relating to the interaction between coworkers and employers.

[28] An economically relevant but distinct question is whether worker and coworker quality are *complementary*

In summary, we find that augmenting the canonical AKM model to incorporate coworker interactions has two main implications for our understanding of the sources of wage inequality. First, each of the contributions from own-type effects, firm effects and worker-firm sorting shrink relative to the standard model, suggesting that these components pick up some of the coworker interactions restricted to be zero in the standard AKM model. Second, coworker sorting represents a quantitatively meaningful source of wage inequality.[29] We conclude the following:
Fact 5. *Having highly skilled coworkers is associated with substantial positive wage gains, and positive assortative matching of workers into teams is about as important a contributor to wage inequality as worker-firm sorting.*

4 Conclusions and policy discussion

In summary, this paper married rich administrative panel data with a comprehensive approach to measuring worker and firm characteristics along multiple dimensions. It shed light on the interplay between worker and firm heterogeneity in shaping economic outcomes.

We conclude with a four observations about future research directions which we believe this study points to. First, while there is a rich literature investigating general-equilibrium firm dynamics given a well-defined notion of the boundary of the firm, a model that integrates such an account with classic notions of assortative matching is missing. On the one hand, such a framework ought to be consistent with the empirical moments that we documented, for instance regarding the relationship between firm size to workforce quality in the cross-section and over time. On the other hand, such a framework would open the door to many questions, including a quantification of the sources of the large-firm wage premium across sectors and an explanation for unequal effects of a failing job ladder in recessions Haltiwanger *et al.* (cf. 2018). It might also allow consideration of ex-ante human capital heterogeneity in the burgeoning literature on wage mark-downs associated with monopsony power, which may have an important bearing on policy implications (Jarosch *et al.*, 2019; Berger *et al.*, 2022, 2023).

Second, our reduced-evidence documenting the importance of coworkers in explaining wage dispersion naturally raises the question if and how the estimated statistical components in the

---

and, hence, whether high-quality workers disproportionately benefit from having high-quality coworkers. Appendix B.4 quantifies the associated moment for wages, that is the cross-partial derivative of wage with respect to own and coworker quality, and finds support for the presence of coworker talent complementarity, consistent with the argument of Freund (2023).

[29]To clarify, the term $2\text{Cov}(\alpha_i, \gamma\bar{\alpha}_{-it})$ is *distinct* from what Song *et al.* (2019) label "worker segregation" in the context of the canonical AKM model, i.e. when $\gamma = 0$ by construction. Worker segregation thus understood pertains to the variance of average firm-level worker effects. In the context of the standard model, greater coworker sorting in this sense means that between-firm inequality is higher but within-firm inequality shrinks correspondingly. According to equation (4), on the other hand, greater coworker sorting also raises *overall* inequality.

coworker-augmented AKM model can be interpreted *structurally* in terms of primitives describing technology, preferences and market structure. Such an exercise would be important to offer an economic interpretation, disentangling the roles of ex-ante firm productivity heterogeneity, human capital heterogeneity, production complementarities between workers and firm and among coworkers, and idiosyncratic tastes for different workplaces.

Third, the correlational evidence that human capital predicts future employment growth among young firms is intriguing and consistent with survey evidence about the importance of talent in shaping start-up outcomes (Gompers *et al.*, 2020). Yet, plausibly causal evidence is lacking. This puts a premium on future research that exploits quasi-random variation in initial workforce skill or builds an empirically rich structural model of so-called "gazelles" (Birch, 1979; Sterk *et al.*, 2021) that helps disentangle alternative sources of ex-ante heterogeneity.

Fourth, while data availability, quality and access is generally more difficult for very small firms (i.e., those with fewer than ten employees, which this paper does not cover), including them in further analysis would yield important insights regarding the robustness of our results in this important segment of the Portuguese economy. Also, extending the analysis to other countries, data access permitting, would provide external validity to the findings in other economic contexts and policy environments.

Turning to a discussion of policy, the paper's results lend further support to the importance of policy action for boosting productivity and labor market outcomes that that recognizes the strong interlinks between the two phenomena. As noted by the latest Economic Survey of Portugal (OECD, 2023), continued upskilling will provide a joint benefit both to the productivity of firms and labor market outcomes of workers. To improve matches between firms and workers, active labor market policies should better target smaller businesses and can include pre-screening programs for job vacancies by public employment agencies. Moreover, costly legal processes add to the cost of hiring workers on permanent contracts, which should continue to be moderated. Regarding the employability of low-wage workers, reducing employer social security contributions could be considered.

In sum, the joint analysis of heterogeneous firms and workers and their interaction on labor markets – and, indeed, interaction among colleagues *within* firms – promises new or improved answers to policy-relevant questions about efficiency and inequality. There is ample scope for empirical research oriented toward causality and novel theoretical models aiding with identification and interpretation, complementing the reduced-form evidence reviewed and documented in this paper.

References

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

Abowd, J. M., Kramarz, F., Pérez-Duarte, S., and Schmutte, I. M. (2018). Sorting Between and Within Industries: A Testable Model of Assortative Matching. *Annals of Economics and Statistics*, (129), 1–32.

Acemoglu, D. and Hawkins, W. B. (2014). Search with multi-worker firms. *Theoretical Economics*, **9**(3), 583–628.

Adenbaum, J. (2022). Endogenous Firm Structure and Worker Specialization. *Job Market Paper*.

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, 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.

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.

Babina, T., Ma, W., Moser, C., Ouimet, P., and Zarutskie, R. (2019). Pay, Employment, and Dynamics of Young Firms.

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.

Battisti, M. (2017). High wage workers and high wage peers. *Labour Economics*, **46**, 47–63.

Berger, D., Herkenhoff, K., and Mongey, S. (2022). Labor Market Power. *American Economic Review*, **112**(4), 1147–1193.

Berger, D. W., Herkenhoff, K. F., Kostøl, A. R., and Mongey, S. (2023). An Anatomy of Monopsony: Search Frictions, Amenities, and Bargaining in Concentrated Markets. In *NBER Macroeconomics Annual 2023, Volume 38*. University of Chicago Press.

Berlingieri, G., Calligaris, S., and Criscuolo, C. (2018). The Productivity-Wage Premium: Does Size Still Matter in a Service Economy? *AEA Papers and Proceedings*, **108**, 328–333.

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.

Birch, D. (1979). The Job Generation Process. *MIT program on neighborhood and regional change*.

Bloom, N., Guvenen, F., Smith, B. S., Song, J., and von Wachter, T. (2018). The Disappearing Large-Firm Wage Premium. *AEA Papers and Proceedings*, **108**, 317–322.

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

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

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

Brown, C. and Medoff, J. (1989). The Employer Size-Wage Effect. *Journal of Political Economy*, **97**(5), 1027–59.

Brown, C. and Medoff, J. L. (2003). Firm Age and Wages. *Journal of Labor Economics*, **21**(3), 677–697.

Burdett, K. and Mortensen, D. T. (1998). Wage Differentials, Employer Size, and Unemployment. *International Economic Review*, **39**(2), 257–273.

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

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., Heining, J., and Kline, P. (2018). Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics*, **36**(S1), S13–S70.

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

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

Choi, J., Goldschlag, N., Haltiwanger, J., and Kim, J. D. (2023). Early Joiners and Startup Performance. *The Review of Economics and Statistics*, pages 1–46.

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*.

Criscuolo, C., Gal, P., Leidecker, T., and Nicoletti, G. (2021). The human side of productivity: Uncovering the role of skills and diversity for firm productivity. *OECD Productivity Working Papers, No. 29*.

Doms, M., Dunne, T., and Troske, K. R. (1997). Workers, Wages, and Technology. *The Quarterly Journal of Economics*, **112**(1), 253–290.

Eeckhout, J. (2018). Sorting in the Labor Market. *Annual Review of Economics*, **10**, 1–29.

Eeckhout, J. and Kircher, P. (2018). Assortative Matching With Large Firms. *Econometrica*, **86**(1), 85–132.

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.

Elsby, M. W. L. and Michaels, R. (2013). Marginal Jobs, Heterogeneous Firms, and Unemployment Flows. *American Economic Journal: Macroeconomics*, **5**(1), 1–48.

Fox, J. T. and Smeets, V. (2011). Does Input Quality Drive Measured Differences in Firm Productivity?*. *International Economic Review*, **52**(4), 961–989.

Freund, L. B. (2023). Superstar Teams: The Micro Origins and Macro Implications of Coworker Complementarities. *Available at SSRN 4312245*.

Gompers, P. A., Gornall, W., Kaplan, S. N., and Strebulaev, I. A. (2020). How do venture capitalists make decisions? *Journal of Financial Economics*, **135**(1), 169–190.

Gulyas, A. (2020). Firm Dynamics With Labor Market Sorting. *CRC TR 224 Discussion Paper Series*, (crctr224_2020_181).

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

Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013). Who Creates Jobs? Small Versus Large Versus Young. *The Review of Economics and Statistics*, **95**(2), 347–361.

Haltiwanger, J. C., Hyatt, H. R., Kahn, L. B., and McEntarfer, E. (2018). Cyclical Job Ladders by Firm Size and Firm Wage. *American Economic Journal: Macroeconomics*, **10**(2), 52–85.

Herkenhoff, K., Lise, J., Menzio, G., and Phillips, G. M. (2018). Production and Learning in Teams. Working Paper 25179, National Bureau of Economic Research.

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

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

Hopenhayn, H. A. (1992). Entry, Exit, and firm Dynamics in Long Run Equilibrium. *Econometrica*, **60**(5), 1127–1150.

Jarosch, G., Nimczik, J. S., and Sorkin, I. (2019). Granular Search, Market Structure, and Wages.

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

Karmakar, S., Sedlacek, P., and Félix, S. (2021). Serial Entrepreneurs and the Macroeconomy.

Lachowska, M., Mas, A., Saggio, R., and Woodbury, S. (2020). Do Firm Effects Drift? Evidence from Washington Administrative Data. Technical Report w26653, National Bureau of Economic Research, Cambridge, MA.

Lazear, E. P. (2004). Balanced Skills and Entrepreneurship. *American Economic Review*, **94**(2), 208–211.

Lentz, R., Piyapromdee, S., and Robin, J.-M. (2022). The Anatomy of Sorting - Evidence from Danish Data. *Working Papers*, (hal-03869383).

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

Lindenlaub, I. and Postel-Vinay, F. (2023). Multidimensional Sorting under Random Search. *Journal of Political Economy*, **131**(12), 3497–3539.

Lochner, B. and Schulz, B. (2022). Firm Productivity, Wages, and Sorting. *Journal of Labor Economics*.

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

Mortensen, D. T. (2003). *Wage Dispersion: Why Are Similar Workers Paid Differently?* The MIT Press.

Mueller, B. and Murmann, M. (2016). The Workforce Composition of Young Firms and Product Innovation - Complementarities in the Skills of Founders and Their Early Employees.

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

OECD (2023). *OECD Economic Surveys: Portugal 2023*. Organisation for Economic Co-operation and Development, Paris.

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

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.

Sterk, V., Sedláček, P., and Pugsley, B. (2021). The Nature of Firm Growth. *American Economic Review*, **111**(2), 547–579.

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature*, **49**(2), 326–365.

Appendix

A Data and methodology

A.1 Data processing

Our basic data cleaning includes the removal of any duplicate person and employer identifiers. Where for a given worker-year multiple spells are recorded, we only retain the job with the highest regular monthly hours.

We focus on manufacturing and services and drop the following STAN-A38 industries: 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). Moreover, by focussing on firms with non-missing entries for value-added, and since the SCIE only covers the non-financial business sectors, we also exclude the finance and insurance industries.

To minimise measurement error or noise we implement the following outlier filtering procedure. We compute annual productivity growth rates and drop the entire firm if in any year it exhibited growth rates in the top/bottom percentile of the productivity growth rate distribution within STAN-A38 industries, or if in every year no growth rate could be computed (e.g., due to radical changes such as mergers or split-ups).

Two more notes are in order. First, because productivity measures were computed in logs, we effectively dropped observations on firms with negative value added. Moreover, when assigning productivity groups, we dropped industry-years where less than ten moving-average firm-level observations were available within STAN A38 x year cells.

Note that the AKM models, both canonical (equation (1)) and augmented (equation (4)), are estimated on an annual worker-level panel, on the basis of which we then construct the firm-level panel used in the majority of analyses. Table A.1 provides summary statistics for this dataset. The statistics differ slightly from the firm-level panel, even when weighting the latter by underlying worker-year observations, due to the additional restrictions imposed on the panel of firms. The differences are, however, small in magnitude.

Variable construction.

Value-added per employee. Value-added (at market prices) corresponds to turnover (sales and services provided) minus cost of goods sold and materials consumed minus External supplies and services (ESF). We compute, for each firm, value-added per employee headcount.

	Mean	Std.
Worker-year obs.	5,507,243	-
Age (years)	40.98	9.72
Tenure (years)	9.84	9.49
Log real hourly wage	1.73	0.51
Female share (%)	37.2	-
Share in low-skill occupations (%)	33.7	-
Share in medium-skill occupations (%)	55.8	-
Share in high-skill occupations (%)	10.5	-

Table A.1: Summary statistics for the worker-level panel

Notes. This table provides summary statistics for the worker-level annual panel (2010-2017). Real values are in 2012 Euros.

Firmsize groups. Firms are bin into three groups based on their number of employees: 10-49; 50-249; and ≥ 250.

Occupational categories. We refer to Criscuolo *et al.* (2021) for a detailed discussion of how two-digit occupations are ranked into low-, medium- and high-skill categories.

B Supplementary results

This appendix contains additional empirical results. Furthermore, several other results, such as an AKM-based wage variance decomposition over time, are reported in Freund (2023) and are available upon request.

B.1 Job creation and job destruction by firm age group

Table B.1 shows job destruction, job creation, and net job creation rate by age-group. Two findings stand out. First, the net job creation is declining in age. Second, most variation is driven by differences in job destruction as opposed to job destruction rates.

B.2 Workforce predicts employment growth: 2004-2017

Table B.2 repeats the analysis in Section 3.3 but estimates equation (3) on a sample spanning 2004-2017 instead of 2010-2017. The results are discussed in the main text.

Firm Group	JD Rate	JC Rate	Net Rate
0-4	0.142	0.200	0.058
5-9	0.146	0.166	0.020
10-19	0.138	0.138	0.000
20-29	0.133	0.117	-0.015
30-49	0.127	0.099	-0.027
50-99	0.125	0.087	-0.038
100+	0.130	0.078	-0.052

Table B.1: Job creation and job destruction by firm age group

Notes. This table reports the job dynamics rates by different firm age groups. JD, JC, and Net stand for job creation, job destruction, and the difference between the two, respectively.

$\ln$(Employment in 5 years)	(1) 0-4y	(2) 5-9y	(3) 10-19y	(4) 20-29y	(5) 30-49y	(6) >50y
FE	0.444***	0.290***	0.240***	0.273***	0.289***	0.229**
	(0.0500)	(0.0332)	(0.0237)	(0.0296)	(0.0463)	(0.0694)
Occupation-FE	0.248***	0.147***	0.138***	0.170***	0.292***	0.172
	(0.0541)	(0.0356)	(0.0279)	(0.0372)	(0.0633)	(0.106)
Within-occ. FE	0.377***	0.246***	0.201***	0.245***	0.226***	0.255***
	(0.0547)	(0.0358)	(0.0252)	(0.0335)	(0.0454)	(0.0716)
Observations	4669	7349	17160	12256	8410	2316

Table B.2: Regressions of 5-year ahead log employment on workforce quality (2004-2017)

Notes. This table indicates the estimated values for β_1 in a set of regressions of five-year ahead employment on current log employment as well as a a measure of worker quality. The regression is run separately for 3 different measures of worker quality – represented by rows – and 6 different firm age groups – indicated by columns. All regressions include industry and firm birth-year fixed effects, in addition to controlling for current log employment. Standard errors are clustered at the firm-level and indicated in parentheses. *** p<0.001, ** p<0.01, * p<0.05. This table presents results for the sample: 2004-2017.

B.3 Decomposition of the large firm wage premium by sector

Table B.3 indicates for each of four sector groups the wage premium by firm size group relative to firms of size 10-19, and decomposes it into AKM components. The results are discussed in the main text.

B.4 Coworker quality complementarity

In Section 3.5 we estimated a coworker-augmented AKM model and used the fixed effects thus obtained to decompose the variance of log wages. In this appendix section, we instead concentrate on the question of *coworker quality complementarity* (Freund, 2023), asking whether high-quality workers *disproportionately* benefit from having high-quality coworkers.

To this end, we replicate the methodology proposed in Freund (2023) and regress worker i's wage in year t on i's own type, average coworker type and, crucially, the interaction between these two covariates. Specifically, we estimate

$$\frac{w_{it}}{\bar{w}_t} = \beta_0 + \beta_x \hat{x}_i + \beta_{x'} \hat{x}_{-it} + \beta_c \left(\hat{x}_i \times \hat{x}_{-it}\right) + \psi_{j(it)} + \nu_{o(i)t} + \xi_{s(i)t} + \epsilon_{it}, \tag{B.1}$$

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.[B.1]

The independent variables are discretized versions of the worker fixed effects (FE) obtained from estimation of equation (1). Specifically, worker i's decile rank in the economy-year specific distribution of WFEs is denoted $\hat{x}_i$. Furthermore, $\hat{x}_{-it}$ is the (unweighted) average quality decile among coworkers. As in the main text, we consider three alternative definitions of the coworker group, namely employer-year, employer-occupation-year and employer-layer-year groups. We also perform the ranking of workers within, respectively, year, occupation-year, and layer-year cells. In addition, we also consider a specification whereby workers are ranked within occupation-year but everyone in the same company is treated as a member of the coworker group.

The coefficient of interest is β_c, which indicates how the effect of having a better coworker vary with your own type. As we treat $\hat{x}$ as continuous in the estimation, β_c specifically indicates how much more the real wage of an individual i rises, as a percentage of the average wage $\bar{w}_t$, with a one-decile increase in coworker quality compared to an individual i' whose rank is one decile lower than that of i. To provide a sense of magnitudes, if β_c is equal to 0.005, this means that the real hourly wage increase from a one decile improvement in the average coworker quality

[B.1] We also include squared terms in $\hat{x}_i$ and $\hat{x}_{-it}$ to address the concern that the interaction term picks up convexity in the return to own or coworker quality. For the sake of readability, these terms are omitted in equation (B.1).

Size category	Log wage	WFE	WFE: occ.	WFE: within-occ.	Firm FE	Observables	Residual
Knowledge-intensive manufacturing							
20-49	0.08	0.04	0.03	0.01	0.03	0.00	0.00
50-99	0.13	0.08	0.04	0.04	0.04	0.00	0.00
100-249	0.22	0.14	0.06	0.08	0.08	0.01	0.00
250-499	0.17	0.11	0.03	0.08	0.05	0.00	0.00
>=500	0.32	0.19	0.04	0.15	0.11	0.02	0.01
Less knowledge-intensive manufacturing							
Size category	Log wage	WFE	WFE: occ.	WFE: within-occ.	Firm FE	Observables	Residual
20-49	0.06	0.03	0.01	0.01	0.03	0.00	0.00
50-99	0.12	0.06	0.03	0.03	0.06	0.00	0.00
100-249	0.19	0.10	0.05	0.05	0.08	0.01	0.00
250-499	0.30	0.16	0.06	0.10	0.13	0.01	0.00
>=500	0.43	0.23	0.08	0.15	0.18	0.01	0.00
Knowledge-intensive services							
Size category	Log wage	WFE	WFE: occ.	WFE: within-occ.	Firm FE	Observables	Residual
20-49	0.15	0.10	0.03	0.07	0.06	0.00	0.00
50-99	0.22	0.14	0.02	0.12	0.08	0.00	0.00
100-249	0.33	0.20	0.03	0.17	0.12	0.00	0.01
250-499	0.34	0.22	0.04	0.18	0.12	-0.01	0.01
>=500	0.37	0.22	0.02	0.20	0.13	0.01	0.01
Less knowledge-intensive services							
Size category	Log wage	WFE	WFE: occ.	WFE: within-occ.	Firm FE	Observables	Residual
20-49	0.07	0.04	0.01	0.03	0.03	0.00	0.00
50-99	0.13	0.08	0.02	0.05	0.06	0.00	0.00
100-249	0.23	0.13	0.05	0.08	0.09	0.00	0.00
250-499	0.22	0.13	0.04	0.09	0.08	0.00	0.00
>=500	0.11	0.06	-0.01	0.08	0.04	0.00	0.00

Table B.3: AKM decomposition of the large firm wage premium, by sector

Notes. This table decomposes the wage premium relative to firms of size 10-19 for 5 other size groups into its AKM components, estimated based on equation (1). This is done separately for 4 different sector groups, classified following the Eurostat indicators on High-tech industry and Knowledge–intensive services aggregation by NACE Rev.2. "Knowledge-intensive manufacturing" comprises high-tech and medium-high tech, "Less knowledge-intensive manufacturing" comprises medium-low tech and low-tech. "WFE" stands for "Worker Fixed Effect."

	Spec. 1	Spec. 2	Spec. 3	Spec. 4
Ranking	Economy-year	Economy-occ.-year	Economy-occ.-year	Economy-layer-year
Coworker group	Firm-year	Firm-year	Firm-occ.-year	Firm-layer-year
$\hat{\beta}_c$	0.0191***	0.0067***	0.0099***	0.0078***
Observations	5,988,548	5,988,548	5,988,548	5,812,372

Table B.4: Coworker wage complementarity estimates

Notes. This table indicates the point estimate of the regression coefficient β_c in regression (B.1), under four different specifications. Standard errors are clustered at the firm-level and indicated in parentheses. *** $p<0.001$, ** $p<0.01$, * $p<0.05$

is 2.5% greater, as a percentage of the average wage, for a worker who is themselves in the top decile as opposed to the fifth decile.

Table B.4 presents the results and is indicative of positive coworker quality complementarities. The magnitude varies across specifications, ranging from 0.0067 to 0.0191. All estimates are statistically significant at the 1% level. This evidence for complementarity, or “supermodularity” in the wage function, is theoretically consistent with positive coworker sorting, the implications of which for wage inequality were highlighted in Section 3.5.