TECHNOLOGICAL CHANGE IN QUANTITIES

Jan Eeckhout, Philipp Kircher, Cristina Lafuente

LIDAM Discussion Paper CORE
2024 / 17

CORE

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: immaq-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/core/core-discussion-papers.html

Technological Change in Quantities*

JAN EECKHOUT, PHILIPP KIRCHER AND CRISTINA LAFUENTE

June 2024

Abstract

Skill-biased technological change has long been linked to rising wage inequality. New technologies also allow firms to expand their scope of their operation. We formalize such quantity-biased technological change and calibrate the model to German matched employer-employee data. The calibration attributes substantial changes in the firm size distribution and in wages to this channel. Quantity-biased technological change spreads out the firm size distribution with a moderating influence on wage inequality within blue and white collar occupations, yet it increases inequality between these occupations. The quantity-bias component in the blue collar occupations alone moderates inequality within and between occupations.

Keywords: quantity-bias, scale-bias, technological change, skill-bias, firm size distribution, wage inequality

JEL: J23, J32, O33

1 Introduction

With the advent of information technology, wage inequality has increased. This technological change is often thought of skill-biased, favoring those with more education, thus leading to larger wage differences with those with less education. Information technology also allows managers to supervise larger sets of workers and thereby to expand their scope of operation. This has the potential to affect not only the size distribution of firms, but also the level and inequality of wages among workers: if more productive firms can now hire larger numbers of workers, more workers can match with more productive firms. This in turn affects the marginal product of workers and the associated remuneration.

*Author information: Jan Eeckhout at University of Pompeu Fabra, jan.eeckhout@upf.edu; Philipp Kircher at Cornell University and Universite Catholique de Louvain and University of Edinburgh, pk532@cornell.edu; Cristina LaFuente at University of Bath, clm96@bath.ac.uk. We thank Gabriele Macci for valuable research assistance. Kircher gratefully acknowledges support through European Research Council Grant ERC-818859.

In this paper, we formalize a model that formalizes this quantity-bias in technological change on top of a standard skill-bias effect. We then quantify it by matching it to changes in the firm size distribution and remuneration in German matched employer-employee data over time. Over the decade 2005 to 2015 quantity-bias was the dominant contributor to the increasing size inequality between firms, but had a moderating influence on wage inequality on the wage distribution with blue and white collar occupations, yet it increased inequality between these occupations.

To understand the impact of improved scale economies on the firms and their workforce requires some notion of firm size. We follow the tradition of Lucas Jr (1978) to rationalize the cross-sectional firm size distribution in the economy through a span-of-control problem. This tradition is appealing as it has at its intellectual heart a communication and control issue that new technologies such as ICT might alleviate. We follow the standard methodology and assume the quantity of production is given by

$$A(x,y)L^{\alpha} \tag{1}$$

where α captures decreasing returns and A is a productivity term that depends on the firm type y and worker type x. The heterogeneity of firms in the cross-section rationalizes a firm size distribution, since more productive firms aim for a larger scale, which is only limited by the scale parameters that capture the span-of-control problem. Technology is quantity biased when α increases and firms are less restricted in achieving a larger scale. We provide a short rationalization of the decreasing returns based on constraints in communication that limit optimal skill usage.

This work is keenly interested not only on the effects of changes in scale economies on firm size, but in particular on its feedback to wage inequality. Therefore, we account for worker heterogeneity, and do so in a way that goes beyond usual "efficiency" units (as a worker with twice the efficiency units always gets twice the pay independent of what happens to firms). We introduce the interaction between firms and workers by allowing productivity A to depend not only on the firm, but also on the ability x of workers that it hires. We allow firms to choose the desired ability, and the number of such workers that it hires.[1]

This generates wage heterogeneity $w(x)$ driven by different marginal products of workers. This marginal product depends not only on the worker type but also on the type of the firm they are matched with. When the scale parameters change and good firms are able to manage a larger set of workers, then workers generally move up to better firm types and achieve a higher marginal product. This not only changes their wage level, but since this effect is heterogeneous across workers, it also changes wage inequality between them as not all workers benefit from this upward mobility to the same degree. This is the crucial link between quantity-biased technological change that allows productive firms to achieve a larger scale and the remuneration

[1] This assumes a single worker type that is hired, which is restrictive even though our micro-foundation delivers this as a result and this is common in related communication problems such as those following Garicano (2000) and Garicano and Rossi-Hansberg (2006) or those in Eeckhout and Kircher (2018).

across different workers. This is distinctly different from effects of skill-biased technological change, which the model captures through change in the marginal product $\partial A/\partial x$ to skill at a given firm.

It is evident that not all workers in a firm undertake the same tasks. To account for this to some degree, we allow for heterogeneity at least at a broad level by separating blue ('b') collar workers and white ('w') collar workers. We assume that workers are attached to the broad occupation, but even within blue (white) collar work there remains substantial heterogeneity in ability. Firms acquire the desired number and ability of workers in each category in a competitive market, and aggregate the contribution of each as in seminal papers on skill-bias such as Katz and Murphy (1992), Caselli and Coleman (2006) or Goldin and Katz (2009):

$$\left[\left(A_b L_b^{\omega_b}\right)^{\sigma} + \left(A_w L_w^{\omega_w}\right)^{\sigma} \right]^{\lambda/\sigma}, \tag{2}$$

through a CES aggregator with substitution parameter σ. Note that in this sense, we nest both the task-based approach of Acemoglu and Autor (2011) within blue/white collar workers (though with endogenous size) with Katz and Murphy (1992) and the heterogeneity between blue and white collar production. This production function reduces exactly to these seminal papers when productivities A_b and A_w are identical for all firms and all workers within an occupational segment so that A_w/A_b captures skill bias across occupations, and when the scale parameters $\omega_b = \omega_w = \lambda = 1$ so that production has constant returns to scale.[2] We are interested in decreasing returns and the effects on the firm size distribution and the remuneration of workers. The link between workers and firms are achieved through dependence of A_s on the firm and worker type in each market occupational segment as outlined in the previous paragraphs, and scale parameters take central stage to capture quantity-bias. The parameter $\omega_i < 1$ induces decreasing returns on each skill level separately as outlined above, while $\lambda < 1$ captures decreasing returns in the combined output, where parts of this can be due to mark-ups induced by monopolistic competition (which we can accommodate in the model) even though we choose our main period of empirical investigation such that mean and variance of markups remain remarkably stable. Quantity-bias on the production side then represents an increase in these scale parameters.

The setup is rich enough to allow us to target the cross-sectional firm size distribution and the relationsship between firm size and profits, the evolution of wages for skilled and unskilled labor by firm size, as well the shifts in the relative use of unskilled versus skilled labor by firm size. Using a large degree of labor heterogeneity within broad occupations allows us to address the large share of wage variation that is within these occupational categories. And we can compare

[2] Caselli and Coleman (2006) also allows for Hicks-neutral capital, which we show not to affect our analysis as long as it is competitively acquired at a common rental price. Like the other seminal papers cited it also distinguished "skilled" and "unskilled" labor, which they classify through educational achievement. In our setup we identify it through the type of work that individuals undertake.

estimated parameters across time periods. Given that some wage heterogeneity within firms remains, we can to some extent consider how wage heterogeneity evolved within and across firms. This richness comes at the cost of more complexity in solving the model, which now involves a system of differential equations with multiple market clearing conditions, and one central contribution of this research is to handle this complexity computationally, and calibrate it to measure the magnitude of the underlying changes in the technology.

For the quantitative exercise we use German matched employer-employee data. In a given year, the production parameters combined with the parameters of the (log-normal) type distributions are estimated to match moments from the data. We compare parameters for the earliest suitable period in our data with the last suitable period in our data, which span more than a decade. The earliest period is 2005 as this follows the completion of the large German labor market reform (the so-called 'Hartz' reforms).[3] The last year of our main analysis consists of 2015, up to which point average markups and markup variance remain remarkably stable in the German economy.

Over the decade that we study, firm sizes become more unequal, and wage inequality expands. The model captures these effects, and attributes most of the change in firm size to quantity-biased technological change. This is also in part responsible for the increased inequality between blue and white collar workers. Technological change would have reduced inequality within occupations, though, in the absence of increased production complementarities and skill-biased change. Absolute effects of quantity-bias on the wage distribution and on median wages dominate those of any other single channel of change in our model, and even the effects on within-occupation inequality are in absolute value roughly as large as those from skill-biased change. This indicates that quantity-bias has the potential to shape inequality as much as traditional forces. Moreover, fostering quantity-bias in blue collar occupation is a useful channel to close wage inequality.

These results are laid out in the following order. We review the literature next. Then Section 2 outlines the model and the sorting condition. Subsequently we calibrate the model and perform counterfactuals that assess different dimensions of change, before discussing extensions and concluding.

Literature. There are various literatures that this works speaks to and on which our analysis builds. Theoretically, our setup extends several literatures. Models of the firm size distribution following the seminal contributions of Lucas Jr (1978) and Hopenhayn (1992) generate firm size differences through a span-of-control problem where the returns to scale are common but

[3]Spanning a period that includes the Hartz reforms raises both conceptual and measurement challenges: the aim of this paper is not to study the role of labor market institutions, nor is it intended to solve labor classification issues that would be more complicated if pre-reform years were included due to the very different notion of full- and part-time work.

productivity of firms differs. Models in this tradition tend not to entail worker heterogeneity beyond efficiency units of labor. Our work allows for span of control in this tradition, but introduces worker heterogeneity and captures the feedback from firm size changes onto wage levels and wage inequality.

Models of assortative matching going back to Becker (1973) have been applied to trade (e.g., Grossman and Maggi (2000)) or compensation (e.g., Gabaix and Landier (2008), Tervio (2008)) but assume one-to-one matching. This has the downside that it is not possible to discuss changes in firm size for labor matching. Technically, these models tend to be much simpler, because assortative matching immediately defines who is matched to whom (the best firm with the best worker, the second best firm to the second-best worker, etc) and the only remaining endogeneous variable is the wage heterogeneity.

This work is closest to that of Eeckhout and Kircher (2018) and its application to trade in Grossman, Helpman, and Kircher (2017) where output depends on the firm, the worker type it chooses, and its size. The main difference is to allow multiple worker types (blue and white collar) within the same firm. Eeckhout and Kircher (2018) provide some preliminary ideas for this case, but do not provide a quantitative implementation. It turns out that multiple types render the computational problem of calibrating such economies massively more difficult, and the main hurdle is market clearing: With only one worker type and assortative matching these models can be solved by guessing the wage of the best worker type, which determines the size of the best firm (which by sorting is matched to the best worker). Then a differential equation establishes the size for all lower-productivity firms and therefore their matched worker types. If there are still workers 'left' unmatched while all firms are assigned, the best worker type was too expensive and the highest firm (and the ones below it) was too small, and a bisection method on the top wage achieves market clearing. With more than one worker type, if all firms are assigned but workers are left, it is not clear how to update the wages of each of the worker types as they interact themselves in the production process.

Other strands of work have assumed matching with linear production. In the trade literature this has been applied to understand the determinants of specialization (see, e.g., Costinot (2009) and Costinot and Vogel (2010), and the overview in Costinot and Vogel (2015)). This literature shows that supermodularity in the productivity of inputs is necessary and sufficient for positive sorting. While their insights are derived in settings where inputs combine linearly to produce output, we show in this paper that one can extend the argument to our setting with decreasing returns to scale. This insight also extends Grossman et al. (2017) where output was directly produced by workers in one profession as in (1) without considering further non-linear aggregation as in (2), though allowing multiple aggregate sectors producting different goods. Already there sorting within a sector required log-supermodularity of the productivity term in (1), but our analysis requires a completely different approach to show that this remains the relevant require-ment even in the presence of within-firm labor aggregation as in (2). While sorting conditions are preserved, the aggregation in (2) does induce a cross-effect in the type and amount of workers from one profession to the other. Eeckhout and Kircher (2018) provide necessary and sufficient sorting conditions for a more general functional forms than (1), but their extension to multiple professions in analogue to (2) only provided necessary but not sufficient conditions for sorting.

Linear production has also been a hallmark of labor market search models with sorting and firm size (see e.g. Bagger and Lentz (2019)), or this literature has used notions where firms are a collection of jobs and search is essentially one-to-one (see the overviews in Chade, Eeckhout, and Smith (2017) and Chade and Kircher (2023)). These frameworks are ill-suited to contemplate the effects of quantity-bias on labor market wage inequality.

On a more applied side one can ask if one can observe technological changes affecting firm size, as well as effects of firms for the remuneration of workers. Empirical work following the seminal contribution of Card, Heining, and Kline (2013) has argued that there is an increasing importance of where a person works for the wages that s/he receives. While there are methodological discussions about the approach (Eeckhout and Kircher (2011); Bonhomme, Lamadon, and Manresa (2019)), the main idea that the role of firms for remuneration is increasing is intriguing. It goes in line with a body of recent work started by Song, Price, Guvenen, Bloom, and Von Wachter (2019) which shows that wage inequality among workers has increased especially between firms. This led us to adopt a model that focuses mostly on wage inequality between firms - we do have wage inequality also within firm but only between large occupational groups (blue collar and white collar).

Finally, ideas of scale-biased economic change have appeared in the literature over time. Already Stevenson (1980) argued that some innovations help large firms more than small ones, and measured this in the electricity-generating sector. More recently, Reichardt (2024) argues that recent technological change might be scale-biased, shifting more weight to large firms. He rationalizes this in an entrepreneurial choice model where entrepreneurs make a choice from a discrete set representing different combinations of size and fixed cost. The paper nicely documents scale-based aspects of historical innovations, but abstracts from any feedback on wage inequality of those who work in such firms (except for the owner-managers). Our model captures a similar effect, but through changes in the returns to scale rather than fixed costs. When firms are less restricted by scale-diseconomies, this helps the productive large firms who can now expand even more, while taking labor from less productive smaller firms who either shrink or disappear. Our focus is also different: We model explicitly the labor side of the model, and the choices regarding quantity and quality and their resulting impact on wage inequality.

2 The Model

In this section we lay out the model through which we will assess changes in the economy, and derive properties of the equilibrium market allocation.

The economy is populated by firms that are heterogeneous in productivity. The mass of firms with type weakly below y is denoted $H_{\mathcal{F}}(y)$, with continuous density $h_{\mathcal{F}}(y)$ on its support $[\underline{y}, \bar{y}] \subset \mathbb{R}_+$. The total mass of firms is then $H_{\mathcal{F}}(\bar{y})$. On the other side of the market there are blue and white collar workers that are heterogeneous in ability. The mass of workers in profession $\mathcal{P} \in \{\mathcal{B}, \mathcal{W}\}$ with type below x_p is $H_p(x_p)$, with continuous density $h_{\mathcal{P}}(x_b)$ on its support $[\underline{x}_{\mathcal{P}}, \bar{x}_{\mathcal{P}}] \subset \mathbb{R}_+$. For simplicity, we assume that workers are attached to their profession.

If a firm of type y hires $L_{\mathcal{P}}$ workers of type $x_{\mathcal{P}}$ within profession $\mathcal{P}$ these produce intermediate labor services

$$S_{\mathcal{P}}(x_{\mathcal{P}}, L_{\mathcal{P}}, y) = \underbrace{\tilde{A}_{\mathcal{P}} \left(\gamma_{\mathcal{P}} x_p^{\sigma_{\mathcal{P}}} + (1-\gamma_{\mathcal{P}}) y^{\sigma_{\mathcal{P}}} \right)^{1/\sigma_{\mathcal{P}}}}_{\Omega_{\mathcal{P}}(x_{\mathcal{P}}, y)} L_{\mathcal{P}}^{\alpha_{\mathcal{P}}}. \tag{3}$$

Here productivity $\Omega_{\mathcal{P}}(x_p, y)$ is a CES-aggregator with substitution parameter $\sigma_{\mathcal{P}} < 1$, and $\alpha_{\mathcal{P}} < 1$ captures decreasing returns in production.

The firm then aggregates the labor services as in Caselli and Coleman (2006) through another CES-aggregator to achieve output

$$\tilde{f}(x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}}, y) = (S_{\mathcal{B}}(x_{\mathcal{B}}, L_{\mathcal{B}}, y)^{\mu} + S_{\mathcal{W}}(x_{\mathcal{W}}, L_{\mathcal{W}}, y)^{\mu})^{\tilde{\lambda}/\mu} \tag{4}$$

with substitution parameter $\mu < 1$ and possibly also decreasing returns on the aggregate captured through $\tilde{\lambda} \leq 1$.

To keep with works such as Caselli and Coleman (2006) we can allow firms to acquire capital at some common rental rate and combine it via a Cobb-Douglas production function with their labor intermediate $\tilde{f}$. This does not capture investments that are specifically aimed to improve returns to scale which we capture in reduced form through changes in the returns parameters, but ensures that we can account for Hicks-neutral capital investments. We can also allow firms to engage in monopolistic competition by setting their price and a representative consumer aggregates the output from different firms in a standard CES utility function. We review in Appendix A.1 that this simply induces profits prior to wage payments of

$$f(x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}}, y) = A \left(S_{\mathcal{B}}(x_{\mathcal{B}}, L_{\mathcal{B}}, y)^{\nu} + S_{\mathcal{W}}(x_{\mathcal{W}}, L_{\mathcal{W}}, y)^{\nu} \right)^{\lambda/\nu} \tag{5}$$

where A is an equilibrium object that is constant across all firms, and $\lambda \leq 1$ is a decreasing returns parameter that captures both the effects of decreasing returns in production as well as through monopolistic competition. In our basic empirical specification we only back out changes in return parameters $\alpha_{\mathcal{B}}, \alpha_{\mathcal{W}}$ and λ across time periods, but do not directly distinguish between market power and production function changes affecting λ over time. Since our main period of investigation features roughly constant mark-ups, this allows us to study changes in production. We also only back out revenue products $A_{\mathcal{P}} \equiv A^{\mu/\lambda}\tilde{A}_{\mathcal{P}}$.

The labor market is competitive. Firms maximize therefore

$$\max_{x_{\mathcal{B}},L_{\mathcal{B}},x_{\mathcal{W}},L_{\mathcal{W}}} f(x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}}, y) - \omega_{\mathcal{B}}(x_{\mathcal{B}})L_{\mathcal{B}} - w_{\mathcal{W}}(x_{\mathcal{W}})L_{\mathcal{W}} \tag{6}$$

where $\omega_{\mathcal{P}}(x_{\mathcal{P}})$ is the competitive equilibrium wage for a worker of type $x_{\mathcal{P}}$ in profession $\mathcal{P}$. Let $G(x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}}, y)$ be the mass of firms with type weakly below y that hire blue (white) collar workers of type weakly below $x_{\mathcal{B}}$ ($x_{\mathcal{W}}$) at size weakly below $L_{\mathcal{B}}$ ($L_{\mathcal{B}}$). We call an allocation G feasible if its marginal over y coincides with $H_{\mathcal{F}}(\cdot)$, and if we sum over the labor demand of all firms for workers below a certain ability this does not add up to more of such workers workers then there are in the economy. That means that for any profession $\mathcal{P} \in \{\mathcal{B}, \mathcal{W}\}$ and any worker ability $x \in [\underline{x}_{\mathcal{P}}, \bar{x}_{\mathcal{P}}]$ it holds:

$$\int_{\{(x_{\mathcal{B}},L_{\mathcal{B}},x_{\mathcal{W}},L_{\mathcal{W}},y)\in\mathbb{R}^5_+ | x_{\mathcal{P}} \leq x\}} L_{\mathcal{P}} dG \leq H_{\mathcal{P}}(x). \tag{7}$$

We can than define a competitive equilibrium of the labor market:

Definition 1 *A competitive equilibrium is then a feasible allocation $G(\cdot)$ and wage functions $\omega_{\mathcal{B}}(\cdot)$ and $\omega_{\mathcal{W}}(\cdot)$ such that*

1. **Profit maximization:** *for any tuple $(x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}}, y)$ in the support of G the elements $(x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}})$ maximize (24) for firm y, and*

2. **Market clearing:** *condition (7) holds everywhere with equality.*

Note that market clearing implies that all workers get hired in equilibrium, which is due to the Inada condition of our production function which ensure that there cannot be a positive measure of workers who are not hired.

2.1 A Simple Microfoundation

Given our focus on information and communication technologies and decreasing returns, it might be worthwhile to provide a simple but novel illustrative model of the communication problem that gives rise to decreasing returns in production and clarifies what our formulation is currently capturing and what it is not.

Assume that there are $L \geq 1$ workers employed by firm y in profession $\mathcal{P}$.[4] Each worker draws a problem $z \in [0, 1]$. The problem needs to be successfully described to and monitored

[4]In particular, our micro-foundation does not translate easily to the problem where a firm hires $L < 1$ workers within a profession. We view this extension to fractional assignments as a minor issue.

by top management in order to proceed productively. This could be due to shared knowledge (e.g., there is a tiny chance that the has catastrophic consequences for the firm and only the upper level management knows and has to "sign off" on the work), or it could be due to effort provision (e.g., only if the problem can be successfully described and monitored the workers can be paid by performance which is necessary to motive them to exert effort).

We think about the state of information and communication technology in the economy as a level $\alpha_{\mathcal{P}}$ up to which problems are easily described and monitored, presumably because these are codified and technology is in place to record data on these dimensions at ease. Consider employees in a supermarket who get paid by the time they work productively, and some draw problems such as the need to open an additional cash register, re-stocking shelves, or are asked by a customer for help. Standard ICT solutions today routinely include the ability to scan employee barcodes at every cash register, and to scan item bar codes for every employee involved in restocking, which makes it easy to condition pay or performance evaluations on these tasks (i.e., workers at registered or in restocking have $z < \alpha_{\mathcal{P}}$). But workers who happen to be asked by a customer for help can easily be confused with taking a break (i.e., here $z > \alpha_{\mathcal{P}}$). This might change as artificial intelligence in camera systems evolves to distinguish customer interactions from lunch breaks (i.e., $\alpha_{\mathcal{P}}$ might increase). For illustration we take the extreme stance that problems that are standardized $(z \leq \alpha_{\mathcal{P}})$ are productive and the employee achieves full productivity $\Omega(x_{\mathcal{P}}, y)$, while problems that are not standardized $(z > \alpha_{\mathcal{P}})$ cannot be communicated and monitored at all and lead to zero output. Labor services $S_{\mathcal{P}}$ for workers in this profession for this firm then add the productivity across all employees encountering problems below $\alpha_{\mathcal{P}}$.

This formulation clarifies the type of information and communication systems we have in mind: these are standardized systems that apply to all firms. This captures systems such as bar code scanning at cash registers, or improvements in artificial intelligence that distinguish customer interactions from breaks. These are general technology improvements that become standard. Our approach does not capture customized investments that make the communication cutoff firm-specific (i.e., $\alpha_{\mathcal{P}}(y)$ is a function of y). We discuss this further in the discussion in Section 6, but point out here that the focus on standardized systems arguably captures a non-trivial part of technology improvement and allows our formulation to be closer to the long-standing tradition that generates firm size difference through productivity heterogeneity instead of heterogeneity in returns to scale (Lucas Jr (1978); Hopenhayn (1992)), as we can recover familiar functional forms and simultaneously handle worker heterogeneity across firms. We will illustrate that next.

If the firm has $L = 1$, this employ concentrates on tasks that are essential to the operation and have been codified. But as additional employees join the firm, they venture into problems that are more complex and less often encountered. Let $\Psi_{\mathcal{P}}(z|L_{\mathcal{P}})$ be the cumulative distribution of problems that the L'th employee in profession $\mathcal{P}$ of the company draws from. If all employees within the same profession and firm have the same ability, we can rationalize the case of exponentially rising labor services with firm size (e.g., labor services of form $\Omega_{\mathcal{P}} L_{\mathcal{P}}^{\alpha_{\mathcal{P}}}$) as a special case through the following insight:

Lemma 1 *Under this microfoundation, if $L_{\mathcal{P}} \geq 1$ workers of ability $x_{\mathcal{P}}$ are employed in profession $\mathcal{P}$ in a firm of type y, they provide labor services as in (3) for any level of $\alpha_{\mathcal{P}} \in (0,1)$ iff the cumulative distribution of problems for the L'th worker has form $F(\alpha_{\mathcal{P}}|L_{\mathcal{P}}) = \alpha_{\mathcal{P}} L_{\mathcal{P}}^{\alpha_{\mathcal{P}}-1}$.*

Proof: We are to show that

$$1 + \int_{1}^{L_{\mathcal{P}}} F(\alpha_{\mathcal{P}}|L) dL = L_{\mathcal{P}}^{\alpha_{\mathcal{P}}}, \tag{8}$$

for any $\alpha_{\mathcal{P}} \in (0,1)$ and any $L_{\mathcal{P}} \geq 1$, where the left hand side represents that workers up to $L = 1$ are concentrated on core problems that can be communicated while additional workers are only productive if their problem remains under the communication threshold. Taking derivatives reveals as necessary condition that $F(\alpha_{\mathcal{P}}|L_{\mathcal{P}}) = \alpha_{\mathcal{P}} L_{\mathcal{P}}^{\alpha_{\mathcal{P}}-1}$. Note that $F(\alpha_{\mathcal{P}}|L_{\mathcal{P}})$ is a proper cumulative distribution since $F(0|L_{\mathcal{P}}) = 0$, $F(1|L_{\mathcal{P}}) = 1$ and $\partial F(\alpha_{\mathcal{P}}|L_{\mathcal{P}})/\partial \alpha_{\mathcal{P}} > 0$ for $\alpha_{\mathcal{P}} \in (0,1)$. Clearly, $F(\alpha_{\mathcal{P}}|L_{\mathcal{P}}) = \alpha_{\mathcal{P}} L_{\mathcal{P}}^{\alpha_{\mathcal{P}}-1}$ satisfied (8). $\square$

Since the main novelty is to have an economy meaningful worker-heterogeneity and to study the feedback of firm-size on wages and wage inequality, would an individual firm still want to hire identical workers within a profession under this micro-foundation? If the firm cannot "order" its employees and therefore does not know in which order they draw their problems, this is indeed the case.[5] To see this, observe that a firm of size $L_{\mathcal{P}}$ the firm has expected output $\Omega_{\mathcal{P}}(x_{\mathcal{P}}, y) \alpha_{\mathcal{P}} L_{\mathcal{P}}^{\alpha_{\mathcal{P}}-1}$ from an employee of type $x_{\mathcal{P}}$. At size $L_{\mathcal{P}}$, for any given amount of labor services $\bar{S}_{\mathcal{P}}$ the firm wants to achieve these labor services at minimal cost. Let $\Phi_{\mathcal{P}}(x_{\mathcal{P}})$ be the amount of workers with type below $x_{\mathcal{P}}$ that the firm hires (with obvious restriction that $\Phi_{\mathcal{P}}(\bar{x}_{\mathcal{P}}) = L_{\mathcal{P}}$), its cost minimization problem becomes

$$\begin{aligned} \min_{\Phi_{\mathcal{P}}(\cdot)} \quad & \int w(x) d\Phi_{\mathcal{P}} \\ s.t. \quad & \int \Omega_{\mathcal{P}}(x_{\mathcal{P}}, y) \alpha_{\mathcal{P}} L_{\mathcal{P}}^{\alpha_{\mathcal{P}}-1} d\Phi_{\mathcal{P}} = \bar{S}_{\mathcal{P}}. \end{aligned}$$

Clearly, putting all mass on a type $x_{\mathcal{P}}$ with maximal $\Omega_{\mathcal{P}}(x_{\mathcal{P}}, y)/\omega(x_{\mathcal{P}})$ is optimal. Therefore, restricting firms to hire only a single worker type in each profession is without loss of optimality. While within a firm-profession combination we can reduce complexity through a single firm type, does this rule out a feedback from size on wage inequality or vice versa? We will see that this feedback persists as different firms specialize on different workers which affects wages and

[5] If the firm can order its employees, it would line up better workers first as their labor services are less likely to get lost, and this might give rise to different worker types within the same firm. If workers only observe the level of their problem when they are working on it, this is not the appropriate model and have not explored it further.

the firm size. We also have not ruled out that different firms that share the same y hire different worker types, which could happen if they are indifferent. Under appropriate sorting conditions discussed next this will not be the case.

2.2 Solution and Conditions for Positive Assortative Matching

A standard feature of most assignment models is that they are much easier to characterize if there is sorting, which loosely means that more productive firms hire more productive workers.

To define this more clearly in our setting, consider an allocation G with support that includes only combinations $(x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}}, y)$ such that $x_{\mathcal{P}} = \mu_{\mathcal{P}}(y)$ and $L_{\mathcal{P}} = \theta_{\mathcal{P}}(y)$ for some functions $\mu_{\mathcal{P}}(\cdot)$ and $\theta_{\mathcal{P}}(\cdot)$. Function $\mu_{\mathcal{P}}(\cdot)$ is called the assignment function, and the sum of functions $\theta_{\mathcal{P}}(\cdot)$ across both professions denotes the size of the firm. The equilibrium exhibits positive assortative matching - or in short "positive sorting" - in occupation $\mathcal{P}$ if the assignment function is strictly increasing. We say that the equilibrium exhibits positive sorting if it features positive sorting in both occupations. In our model more able workers in a given profession earn higher wages, and more productive firms make higher profits. In the data firm size is positively related to higher profits and higher wages for both worker types, which the model can rationalize only under positive sorting. The following shows necessary and sufficient conditions for such sorting, and then describes how the model can be solved under positive sorting.

For a given firm type y, consider any optimal tuple $(\mu_{\mathcal{B}}(y), \theta_{\mathcal{B}}(y), \mu_{\mathcal{W}}(y), \theta_{\mathcal{W}}(y))$, where we suppress the y subscript for brevity in the following. This tuple maximized (24) only if it produces labor services $\bar{S}_{\mathcal{P}}(y) := S_{\mathcal{P}}(\mu_{\mathcal{P}}, \theta_{\mathcal{P}}, y)$ at the lowest cost in both professions. In particular, $(\mu_{\mathcal{P}}, \theta_{\mathcal{P}})$ has to solve

$$\begin{aligned} \min_{x_{\mathcal{P}}, L_{\mathcal{P}}} \quad & \omega_{\mathcal{P}}(x_{\mathcal{P}}) L_{\mathcal{P}} \\ s.t. \quad & \Omega_{\mathcal{P}}(x_{\mathcal{P}}, y) L_{\mathcal{P}}^{\alpha_P} = \bar{S}(y). \end{aligned} \tag{9}$$

The constraint can equivalently be written as

$$\Omega_{\mathcal{P}}(x_{\mathcal{P}}, y)^{1/\alpha_P} L_{\mathcal{P}} = \bar{S}_y^{1/\alpha_P},$$

which resembles the linear constraints familiar from trade models such as in Costinot and Vogel (2015). Substituting this constraint into the objective function and taking logs as a monotone transformation reveals that $(\mu_{\mathcal{P}}, \theta_{\mathcal{P}})$ has to solve

$$(\mu_{\mathcal{P}}, \theta_{\mathcal{P}}) \in \arg\min_{x_{\mathcal{P}}} \left(\ln(\omega_{\mathcal{P}}(x_{\mathcal{P}})) - \frac{1}{\alpha_{\mathcal{P}}} \ln(\Omega_{\mathcal{P}}(x_{\mathcal{P}}, y)) + \frac{1}{\alpha_{\mathcal{P}}} \ln(\bar{S}(y)) \right) \tag{10}$$

$$\Rightarrow (\mu_{\mathcal{P}}, \theta_{\mathcal{P}}) \in \arg\max_{x_{\mathcal{P}}} \left(-\ln(\omega_{\mathcal{P}}(x_{\mathcal{P}})) + \frac{1}{\alpha_{\mathcal{P}}} \ln(\Omega_{\mathcal{P}}(x_{\mathcal{P}}, y)) - \frac{1}{\alpha_{\mathcal{P}}} \ln(\bar{S}(y)) \right). \tag{11}$$

Standard results from monotone comparative statics apply: if the objective function in the maximization problem is strictly supermodular in $x_{\mathcal{P}}$ and y, then higher firm types necessarily choose higher worker types. And if the objective function is strictly submodular around optimal point $(\theta_{\mathcal{P}}(y), y)$ then the equilibrium cannot be positively assortative in profession $\mathcal{P}$. These necessary and sufficient conditions necessarily imply:

Lemma 2 *Strict (weak) log-supermodularity of function $\Omega_{\mathcal{P}}(\cdot,\cdot)$ is a sufficient (necessary) condition for positive assortative matching in profession $\mathcal{P}$.*

Therefore, if $\Omega_{\mathcal{P}}$ is strictly log-supermodular, then $\mu_{\mathcal{P}}(y)$ has to be increasing in y, and by market clearing it has to be continuous and strictly increasing. Since strictly increasing functions are almost everywhere differentiable, this can be summed up by implying $\mu'_{\mathcal{P}}(y) > 0$.

While the necessary and sufficient conditions for sorting can be separated across professions, the exact allocation of talent cannot. In one-to-one matching models where each firm hires one worker, positive sorting immediately fixes the allocation of talent by lining up firms and workers according their rank in the type distribution. In our setting we can only say that

$$\mu_{\mathcal{P}}(y) = \bar{x}_{\mathcal{P}} \;\;\&\;\; \theta_{\mathcal{P}}(y) = \bar{\theta}_{\mathcal{P}} \tag{12}$$

where the top agents are matched but the amount of workers hired at the top firms is not determined. The amount of hiring $\bar{\theta}_{\mathcal{P}}$ at top firms affects how many workers arrive in top firms, or more precisely: how quickly lower-ranked workers have to move down the distribution of firm productivities. To see this, note that market clearing under positive sorting requires that all workers hired by firms above y equal the number of workers above $\mu_{\dagger}$ or $\int_y^{\bar{y}} \theta_{\mathcal{P}}(y') dH\mathcal{F}(y) = \bar{H}_{\mathcal{P}} - H_{\mathcal{P}}(\mu(y))$, which yields in differential form:

$$\mu'_{\mathcal{P}}(y) = \theta_{\mathcal{P}}(y) h_{\mathcal{F}}(y) / h_{\mathcal{P}}(\mu_{\mathcal{P}}(y)). \tag{13}$$

It says that the speed at which worker types change depends on the size of the firms divided by the amount of available workers per firm. This shows that larger firm size, all all else equal, induces a quicker change of workers across firm types.

How large do firms want to be? Equilibrium functions $(\mu_{\mathcal{P}}(\cdot), \theta_{\mathcal{P}}(\cdot))$ for both professions solve profit maximization problem (24), and are characterized by first order conditions with respect to worker type and firm size:

$$\frac{\partial f(\mu_{\mathcal{B}}(y), \theta_{\mathcal{B}}(y), \mu_{\mathcal{W}}(y), \theta_{\mathcal{W}}(y), y)}{\partial x_{\mathcal{P}}} \frac{1}{\theta_{\mathcal{P}}(y)} = \omega'_{\mathcal{P}}(\mu_{\mathcal{P}}(y)), \tag{14}$$

$$\frac{\partial f(\mu_{\mathcal{B}}(y), \theta_{\mathcal{B}}(y), \mu_{\mathcal{W}}(y), \theta_{\mathcal{W}}(y), y)}{\partial L_{\mathcal{P}}} = \omega_{\mathcal{P}}(\mu_{\mathcal{P}}(y)), \tag{15}$$

for both $\mathcal{P} \in \{B, W\}$. Condition (15) determines firm size by the marginal product of the last worker to the wage. Equation (14) describes wage inequality across workers of different ability in the same profession, and depends on the change in average product by worker type. A particular noteworthy insight of this equation is that it highlights the role of matching in shaping wage inequality: the left hand side of (14) depends on the type of firm that the worker matches with (and the other workers it hires). All else equal, matching with a better firm induces a higher marginal product due to the complementarities between firm and worker types, which raises wage inequality between workers. On the other hand, all else equal firms being larger reduces the wage inequality between workers. Both effects kick in when firms are less constraint by decreasing returns, as workers move up to better firms but firms also become larger.

To understand the drivers of firm size more deeply, and to solve the model, it is convenient to totally differentiate (15) and to divide both sides of it by (14) for each profession $\mathcal{P}$. Using (13) to substitute out $\mu'_{\mathcal{P}}(y)$ and rearranging the two resulting equations for $\theta'_{\mathcal{P}}(y)$ in profession $\mathcal{P}$ and $\theta'_{-\mathcal{P}}(y)$ in the other occupation "-$\mathcal{P}$" then yields:

$$\theta'_{\mathcal{P}} = \frac{\begin{bmatrix} -\left[f_{L_{-\mathcal{P}}L_{-\mathcal{P}}}f_{L_{\mathcal{P}}x_{\mathcal{P}}} - f_{L_{\mathcal{P}}L_{-\mathcal{P}}}f_{L_{-\mathcal{P}}x_{\mathcal{P}}}\right]\theta_{\mathcal{P}}\mathcal{H}_{\mathcal{P}} - \left[f_{L_{-\mathcal{P}}L_{-\mathcal{P}}}f_{L_{\mathcal{P}}x_{-\mathcal{P}}} - f_{L_{\mathcal{P}}L_{-\mathcal{P}}}f_{L_{-\mathcal{P}}x_{-\mathcal{P}}}\right]\theta_{-\mathcal{P}}\mathcal{H}_{-\mathcal{P}} \\ -\left[f_{L_{-\mathcal{P}}L_{-\mathcal{P}}}\right]f_{L_{\mathcal{P}}y} + f_{L_{\mathcal{P}}L_{-\mathcal{P}}}f_{L_{-\mathcal{P}}y} + f_{L_{-\mathcal{P}}L_{-\mathcal{P}}}f_{x_{\mathcal{P}}}\mathcal{H}_{\mathcal{P}} - f_{L_{\mathcal{P}}L_{-\mathcal{P}}}f_{x_{-\mathcal{P}}}\mathcal{H}_{-\mathcal{P}} \end{bmatrix}}{f_{L_{-\mathcal{P}}L_{-\mathcal{P}}}f_{L_{\mathcal{P}}L_{\mathcal{P}}} - f_{L_{\mathcal{P}}L_{-\mathcal{P}}}f_{L_{-\mathcal{P}}L_{\mathcal{P}}}},$$

where subscripts denote partial derivatives to the variable (except for the indicator $\mathcal{P}$ which denotes a profession and $-P$ which denotes the other profession), where $\mathcal{H}_{\mathcal{P}}(\mu_{\mathcal{P}}(y), y) := h_{\mathcal{F}}(y)/h_{\mathcal{P}}(\mu_{\mathcal{P}}(y))$ denotes the local ratio of firms to workers, and were we dropped the arguments $(\mu_{\mathcal{B}}(y), \theta_{\mathcal{B}}(y), \mu_{\mathcal{W}}(y), \theta_{\mathcal{W}}(y), y)$ of function f and its partial derivatives.[6] By reversing subscripts $\mathcal{P}$ and $-\mathcal{P}$ we obtain a similar expression for $\theta_{-\mathcal{P}}(y)$.

Even though expression (2.2) provides little intuition, it suffices as a viable strategy to solve the model: Assume the necessary and sufficient condition for positive sorting, and use (12) with some guesses for $\bar{\theta}_{\mathcal{B}}$ and $\bar{\theta}_{\mathcal{W}}$ as starting points. Then (13) and (2.2) form a differential equation system that determines firm size and matching at all lower y. Follow that differential equation until $\underline{y}$, where it has to hold that $\mu_{\mathcal{B}}(\underline{y}) = \underline{x}_{\mathcal{B}}$ and $\mu_{\mathcal{W}}(\underline{y}) = \underline{x}_{\mathcal{W}}$. If either of these fails, the initial guesses for $\bar{\theta}_{\mathcal{B}}$ and $\bar{\theta}_{\mathcal{W}}$ have to be updated until the endpoint conditions are met. Finally, wages can be backed out from (15), which also makes clear that the initial guesses about $\bar{\theta}_{\mathcal{P}}$ effectively are guesses about the wages that the most able workers obtain. This outlines how the model can be solved if the data makes it obvious that positive sorting is required.

To round off the theoretical discussion, it might be useful to provide slightly more intuition for the content of equation (2.2). For this purpose, imagine that there is no interaction between the output of the two professions, so that any cross-partials that have both a variable indexed by

[6] See Appendix A.4 for the derivation.

$\mathcal{P}$ and by $-\mathcal{P}$ are zero.[7] Then the expression simplifies massively to:

$$\theta'_{\mathcal{P}} = \frac{\left[f_{L_{\mathcal{P}}x_{\mathcal{P}}} - \frac{f_{x_{\mathcal{P}}}}{\theta_{\mathcal{P}}}\right]\frac{\theta_{\mathcal{P}}}{\mathcal{H}_{\mathcal{P}}^{-1}} + f_{L_{\mathcal{P}}y}}{-f_{L_{\mathcal{P}}L_{\mathcal{P}}}}.$$

The denominator is positive because of our decreasing returns assumptions, and serves mostly as a normalization. Whether better firms are becoming larger ($\theta'_{\mathcal{P}}(y) > 0$) and by how much depends then first of all on the question how much better firms gain more from being large (i.e., on $f_{L_{\mathcal{P}}y}$). But better firms also hire better workers, and how firm size changes depends on the marginal gain with such workers of being larger beyond their average product which represent the increase in wages for such workers (i.e., it depends on $f_{L_{\mathcal{P}}x_{\mathcal{P}}} - \frac{f_{x_{\mathcal{P}}}}{\theta_{\mathcal{P}}}$ which represents $f_{L_{\mathcal{P}}x_{\mathcal{P}}} - \omega'_{\mathcal{P}}$). This is multiplied by the speed a which workers change (i.e, $\theta_{\mathcal{P}}/\mathcal{H}_{\mathcal{P}}^{-1}$, which equals $\mu'_{\mathcal{P}}(y)$). Finally, it depends on how large firms would like to be relative to the local availability of workers to firms ($\frac{\theta_{\mathcal{P}}}{\mathcal{H}_{\mathcal{P}}^{-1}}$). These terms are intuitively important determinants of firm size. Expression (2.2) takes them into account, but also takes into account the local availability of workers in the other profession and their impact on size in the current profession through substitution or complements.

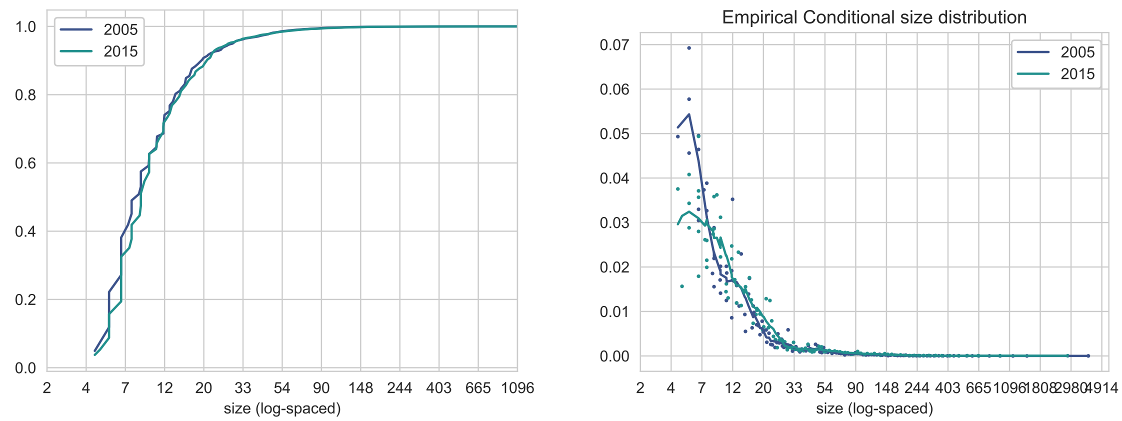
3 Data

We use linked employer-employee data from Germany, also known as the LIAB dataset. This dataset combines three information sources: the Institute for Employment Research (IAB) yearly *Establishment Panel* survey, the Establishment History Panel (BHP) administrative records and the administrative records of workers. From the first source we obtain data on revenue, costs and firm composition. From the latest source, we get individual daily wage and other job characteristics. These allows us to create a representative panel of firms in Germany for which we can calculate the average wage of workers by type of job, firm size and value added.

We select the years 2005, 2010, 2015 and 2019, using the cross-sectional version of the LIAB. We select full-time workers and exclude workers in training. This is an important sample selection choice, as the size of establishment (firms) should be interpreted as the size of the full-time, regular workforce. We select firms with at least 5 full-time workers with at least one white collar and one blue collar worker. We separate workers by their job skill requirements. In particular, we classify jobs as “white collar" those whose jobs require “complex task involving special knowledge" or “highly complex tasks", and “blue collars" those with “low-complexity routine tasks" and “skilled, technical tasks". These categories are related but not quite overlap with the cognitive/manual separation common in the literature, but provide a better split than education,

[7] This arises for our setting when $= \nu = 1$ so that output is linear in each professions labor services.

Figure 1: Changes in firm distribution, Germany



The empirical CDF conditions on establisments larger than 5 workers, with at leas tone worker of each type. The Second panel shows the histogram of the data, fitted with a polynomial for ease of interpretation.

since in the German system has relative low numbers of college graduates. Below is a plot of the changes in the observed firm size distribution and a table of descriptive statistics for our sample.

	2005					2015				
	firm size	log wage blue collar	log wage white collar	log profit	share white collar	firm size	log wage blue collar	log wage white collar	log profit	share white collar
mean	12.36	9.93	10.32	13.03	0.21	12.93	10.04	10.46	13.10	0.21
var	18.56	0.10	0.17	0.63	0.05	23.19	0.11	0.17	0.66	0.06
min	5.00	9.80	9.96	12.34	0.09	5.00	9.83	10.03	12.08	0.12
p25	6.57	9.86	10.16	12.48	0.14	7.00	9.96	10.34	12.65	0.14
p50	8.41	9.90	10.33	12.84	0.15	9.00	10.03	10.46	12.93	0.16
p75	12.82	10.00	10.47	13.28	0.18	13.07	10.13	10.59	13.39	0.18
p99	64.59	10.19	10.61	15.07	0.26	67.79	10.30	10.75	15.37	0.24
max	3947.48	9.96	10.51	19.61	0.40	2822.75	10.03	10.61	19.66	0.41

Notes: Weighted data using the weights in the establishment survey. Wages and profits are annual and deflated using the GDP deflator.
These data corresponds to 127 bins, 50 establishments per bin.

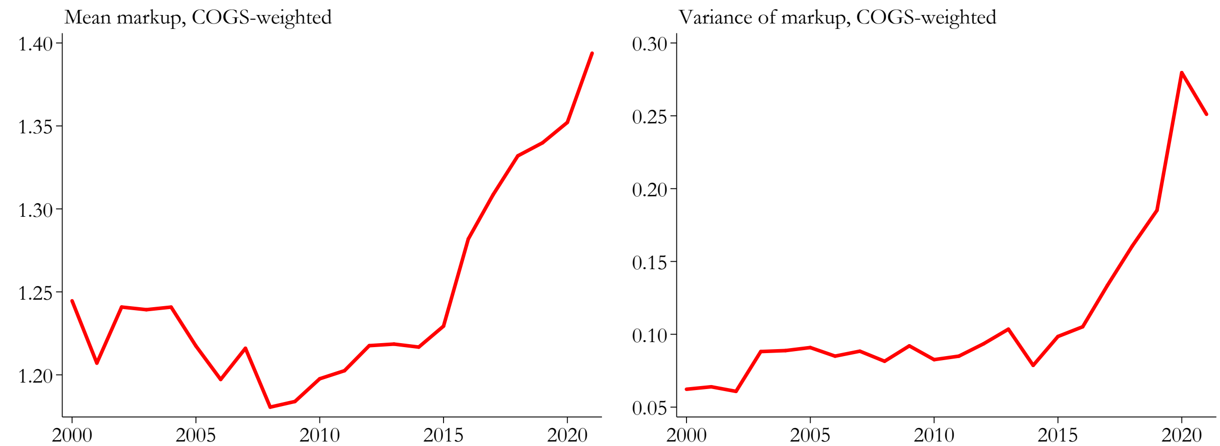
Table 1: Descriptive staticstics of the binned data (unweighted)

It is worth noting that the wage data is accurate but is right censored, as its original purpose is to calculate social security contributions, for which all income up to a point are liable. This is a well-known issue with this dataset (see, e.g., Card, Heining, and Kline (2013)), affecting over 60% of workers in our sample. To overcome this issue, we follow conventional practice in the literature (Dustmann, Ludsteck, and Schönberg (2009)) and input censored wages under the assumption that the error term in a wage regression is normally distributed.

Our main analysis focuses on the decade between 2005 and 2015. The data goes back to the 1990s, but we choose 2005 as starting point since Germany went through a sequence of labor market reforms (the so-called Hartz reforms) prior to 2005 which changed both market environ-

Figure 2: Mark-ups in Germany over time (cost-weighted).

(a) Average mark-up over time (b) Variance of mark-ups over time



ment as well as nature of permissible jobs in the economy.

We choose the end point of our main analysis a decade later, both because a decade is a standard unit long enough for some technological change to materialize but also because 2015 marks the end of a rather stable period in terms of market power in Germany. To illustrate the rationale for our timing choice, we follow the production function approach by De Loecker and Eeckhout (2018) to back out market power for German firms over time. The left panel of Figure 2 shows the cost-weighted average mark-up over time. Up to 2015 average mark-ups stay relatively constant, averaging 22% both in 2005 and 2015. Moreover, the dispersion of mark-ups remains very stable at a low level until 2015. This indicates a remarkably stable time in terms of market power, allowing us to abstract from it in our main calibration. After 2015 average mark-ups and their variance explode, indicating a much larger role for market power changes in that time period.[8]

Both 2005 and 2015 fall into times of economic expansion, following the crash of the ".com" bubble of 2003 and the Great Recession of 2008, respectively. The Great Recessions was comparably mild in Germany, and the decade is marked by an overall decrease in unemployment.

[8] Appendix A.2 confirms this insight for sales-weighted mark-ups.

4 Quantitative Exercise

We parametrise the production function as

$$Y = \left(\left[L_{\mathcal{W}}^{\alpha_{\mathcal{W}}} \tilde{A}_{\mathcal{W}} \left((1-\gamma_{\mathcal{W}}) y^{\frac{\sigma_{\mathcal{W}}-1}{\sigma_{\mathcal{W}}}} + \gamma_{\mathcal{W}} x_{\mathcal{W}}^{\frac{\sigma_{\mathcal{W}}-1}{\sigma_{\mathcal{W}}}} \right)^{\frac{\sigma_{\mathcal{W}}}{\sigma_{\mathcal{W}}-1}} \right]^{\frac{\rho-1}{\rho}} + \left[L_{\mathcal{B}}^{\alpha_{\mathcal{B}}} \tilde{A}_{\mathcal{B}} \left((1-\gamma_{\mathcal{B}}) y^{\frac{\sigma_{\mathcal{B}}-1}{\sigma_{\mathcal{B}}}} + \gamma_{\mathcal{B}} x_{\mathcal{B}}^{\frac{\sigma_{\mathcal{B}}-1}{\sigma_{\mathcal{B}}}} \right)^{\frac{\sigma_{\mathcal{B}}}{\sigma_{\mathcal{B}}-1}} \right]^{\frac{\rho-1}{\rho}} \right)^{\frac{\lambda\rho}{\rho-1}}$$

which combines equations (3) and (4) and normalizes $A = 1$ as it cannot be separately identified from $A_{\mathcal{B}}$ and $A_{\mathcal{W}}$.

The skill levels $x_{\mathcal{B}}, x_{\mathcal{W}}$ and y are assumed to have a log-normal probability distribution in the population:

$$x_{\mathcal{P}} \sim \mathcal{LN}(1, s_{\mathcal{P}}); y \sim \mathcal{LN}(1, s_y)$$

for each type of profession $\mathcal{P} = \{\mathcal{B}, \mathcal{W}\}$.[9] Notice that we normalize the means of the distribution to be equal to one, and adjust the variance to make sure this is always the case. Doing this we ensure that $\tilde{A}_{\mathcal{P}}$ and $\gamma_{\mathcal{P}}$ uniquely pin down the mean of the skill level - that is, that the mean skill levels (and wages) are not dependent on the type distribution. Instead, they depend on parameters of the production function. The variance of the type distribution is still important because the relative density of the firm and worker distribution are key for labour market clearing and (therefore) the firm size distribution.

We therefore have the following 12 parameters to estimate:

$$\{\tilde{A}_{\mathcal{W}}, \gamma_{\mathcal{W}}, \alpha_{\mathcal{W}}, \sigma_{\mathcal{W}}, s_{\mathcal{W}}, \tilde{A}_{\mathcal{B}}, \gamma_{\mathcal{B}}, \alpha_{\mathcal{B}}, \sigma_{\mathcal{B}}, s_{\mathcal{B}}, \rho, s_y\}$$

Given the low levels of mark-ups we set $\lambda = 1$.

We calibrate these parameters for each year separately, using the Simulated Method of Moments. For a given set of parameters we solve the model and calculate a series of moments relating to the firm size distribution and correlations between firm size, wages, profits and type composition of the workforce.

[9] It is well known in matching models like this one that one can label types by their rank in the type distribution - which renders this label uniform on the unit interval - adjust the production function to map ranks back into productivity types. In this sense the type distributions can be viewed as adding three parameters ($s_{\mathcal{B}}, s_{\mathcal{W}}, s_y$). We normalize the means precisely because average productivity is already picked up by parameters $\tilde{A}_{\mathcal{P}}$ and $\gamma_{\mathcal{P}}$ for workers and analogous expressions by firms. So we use

$$h(x_{\mathcal{P}}) = \frac{C_{\mathcal{P}}\sqrt{2}}{\sqrt{\pi x_i^2 s_{\mathcal{P}}^2}} \cdot e^{-\frac{\left(\log(x_i) - \left(1 - \frac{s_{\mathcal{P}}^2}{2}\right)\right)^2}{2 s_i^2}}$$

which is the pdf of a log-normal distribution, centered at e and scaled by $C_{\mathcal{P}}$, which is given by the data (average number of workers of each type per firm).

In particular, the model delivers a unique mapping between firm size, wages, profits and the composition of the firm. In the data, there are clear correlations but there is also considerable variation that our model will not be able to replicate. We therefore use the sum of the squared residuals between the data and the implied mapping from our model from firms size to (a) wage of blue collar workers (b) wage of white collar workers (c) profits and (d) share of white collars in the firm. Additionally, we target 9 deciles of the firm size distribution. In total, our objective function is the sum of 5 sums of squared residuals, one for the correlation of firm size with each variable (a)-(d) plus the fit to firm size distribution, weighted by the variance of each variable.[10]

The algorithm to solve the model, as outlined in the discussion of equation (15), consists of a shooting protocol that iterates over guesses of the size of the largest firm. Converge is guaranteed in the one worker type case, because of the established uniqueness of the solution. However, with two types of workers the algorithm requires to update only one type of guess at a time. That is, given the guess for type $\bar{\theta}_{\mathcal{P}}$, we integrate the differential equation system and update the guess for $\bar{\theta}_{-\mathcal{P}}$, the other worker type. When the market clears for the $\bar{\theta}_{-\mathcal{P}}$ type, then we can infer if the guess for $\bar{\theta}_{\mathcal{P}}$ type is too high or too low and update. The algorithm converges to the solution where both markets are matched, provided we give reasonable initial guesses for both types.[11]

The following table provides our best estimates:

	$\tilde{A}_W$	$\tilde{A}_B$	γ_W	γ_B	σ_W	σ_B	α_w	α_b	ρ	λ	s_W	s_B	s_y
2005	0.672	0.322	0.037	0.038	0.941	0.869	0.171	0.278	0.390	1	0.401	0.516	0.527
2015	0.791	0.281	0.019	0.047	0.401	0.963	0.210	0.352	0.474	1	0.556	0.385	0.528

Table 2: Estimated parameters

Figures 3 and 4 show the basic fit of our model the empirical data. It trades the wage distributions for blue and white collar workers well, and captures the profits for firms except for some slight overshooting in 2005 profits. It also captures well that larger firms have a greater share of white collar workers, though evidently it misses the large dispersion present in the data. Based on this data we often refer to white collar as "high skill" as the wages in this occupation are overall higher as is seen in the second panel. The model also captures well the share of small and

[10] The distance of the firm distribution is weighted by the variance of firm size in the data, which the distances from (a)-(d) variables are weighted by their variance in the data. That is, the SSR between the predicted firm size to profit ratio is weighted by the variance of profits in the data. This makes the objective function place a higher weight on more precise correlations and less on less precise ones. The firm size distribution is weighted by the variance of firm size.

[11] If we instead updated the guesses of both types at the same time, the algorithm would not converge and instead circle between guesses. This also forfeits the use of Newton-like methods (based on distance from the objective) to update the guesses. Instead, we exploit the fact that, given the guess for one type $\mathcal{P}$, there is one guess for the other $-\mathcal{P}$ that clears its market. This in turn gives us a measure of the true distance of the guess of the $\mathcal{P}$ guess from optimal one that clears its market.

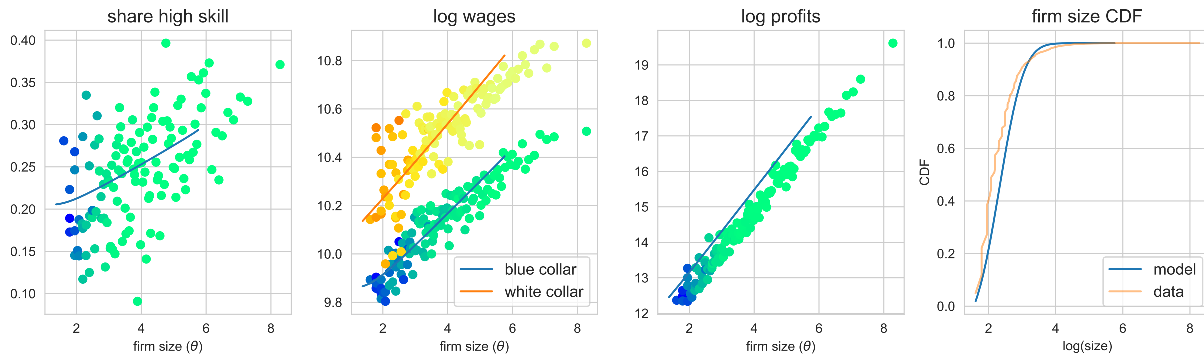


Figure 3: 2005: relation between size and white collar ("high skill") in panel 1, size and low ages for each occupation in panel two, size and profits in panel three, and the firm size distribution in panel 4. It compares model fit (solid lines in first three panels and blue curve in last) to data (dots in first three panels and orange line in last).

large firms in the market, but generates somewhat too many mid-size firms in both time periods.

The fit for the firm size distribution is also summarized in Table 3, while the fit for the wage distribution is captured in 4. It confirms the overall good fit, with overpredicting the medium firm size. Since the model is quite parsimoneous and still has to capture whole distributions across four dimensions, it seems to capture most of the salient dimensions of the data.

	data 2005	model 2005	data 2015	model 2015
p25	6.57	7.77	7.00	8.03
p50	8.41	10.89	9.00	11.52
p75	12.82	15.65	13.07	17.16
p99	64.61	39.14	67.84	47.83
90-10 ratio	3.395	3.615	3.474	3.996

Table 3: Model fit of the firm size distribution

5 Sources of Change and Policy Implications

We will move the model as a laboratory to explore the effects of changing only particular parameters from the 2005 calibration to the 2015 one. This allows us to isolate the effects of each of the changes that happened, and provides a lense through which to assess their respective effects.

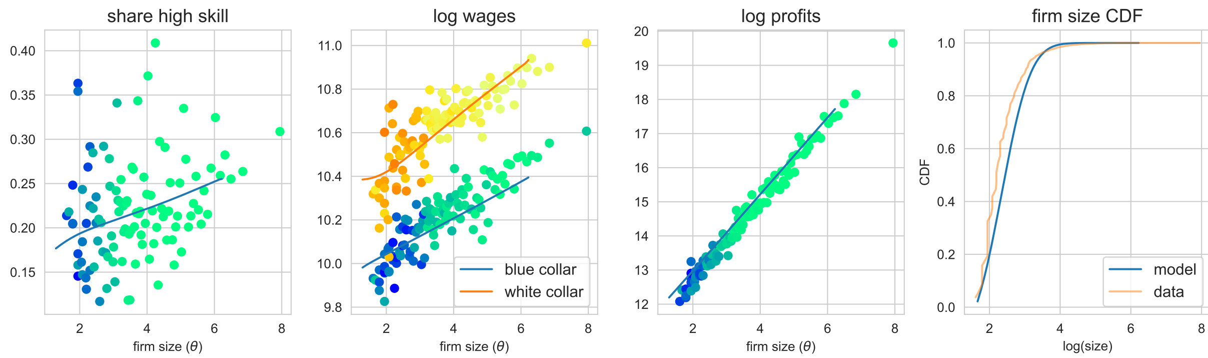


Figure 4: 2015: relation between size and white collar ("high skill") in panel 1, size and low ages for each occupation in panel two, size and profits in panel three, and the firm size distribution in panel 4. It compares model fit (solid lines in first three panels and blue curve in last) to data (dots in first three panels and orange line in last).

	Blue Collar workers				White Collar workers			
	2005 model	2005 data	2015 model	2015 data	2005 model	2005 data	2015 model	2015 data
p25	9.94	9.86	10.07	9.96	10.27	10.16	10.48	10.34
p50	9.99	9.90	10.10	10.03	10.33	10.33	10.51	10.46
p75	10.04	10.00	10.14	10.13	10.39	10.47	10.55	10.59
p99	10.16	10.19	10.23	10.30	10.54	10.61	10.65	10.75

Table 4: Model fit of the wage distribution

5.1 Quantity-biased technological change

At the heart of our paper is the idea that modern technology increases the returns to scale in production, which shifts employment to more productive firms who can expand, which reduces the workforce left for less productive firms which shrink.

To illustrate this, we leave all parameters at their 2005 levels, except that we change $\alpha_{\mathcal{P}}$ to its 2015 levels in both the blue and white collar professions ($\mathcal{P} \in \{B, \mathcal{W}\}$). Table 5 documents the changes for firm size. Relative to the calibration for 2005 (first column), quantity biased-change alone indeed increases the size of the largest percentile of firms and its spread relative to the 25th percentile (column two). The contraction of firms happens below the 25th percentile, i.e., at very small firms. The changes at the largest percentiles capture most of the change in firm size in the model, as can be seen with the comparison with the third column, though the model misses the very largest firms in the data of 2015 (last column).

	best fit 2005	only Δ $\alpha_{\mathcal{P}}$	best fit 2015	data 2015
p25	7.77	7.68	8.03	7.00
p50	10.89	10.90	11.52	9.00
p75	15.65	16.04	17.16	13.07
p99	39.14	43.06	47.83	67.84

Table 5: Results on firm size distribution from changing $\alpha_{\mathcal{P}}$ to 2015 levels, keeping all other parameters at 2005 levels.

These firm size changes spill over to worker renumeration, as documented in Table 6. Relative to the 2005 calibration, wages rise for workers at all percentiles because marginal products do not fall as quickly when firms expand and workers move up the productivity distribution of firms. This happens for blue collar workers (compare the 2005 model in column one with the counterfactual in column two) and white collar workers (compare column 4 and column 5). Interesting spread between the 75th percentile wage and 25th percentile wage decline with quantity-biased technological change for both professions, and so does the 99th vs 25th percentile difference. So wage inequality within each occupation declines, even though the full 2015 calibration implies an increase in inequality within each profession (see column 3 and 6) in line with increased wage inequality in the data (not shown here). On the other hand, inequality between white and blue collar occupations widens with quantity-biased technological change from 3.4% to 5.2% at the median wage, while in the full 2015 calibration this between-occupation inequality increases much less.[12]

Overall, quantity-biased technological change can account for the large change in firm size, it

[12] Figure 5 in appendix A.5 illustrates how our model with just quantity-biased technological change compares to the actual 2015 data.

	Blue collar workers			White collar workers		
	2005 model	only $\Delta\ \alpha_{\mathcal{P}}$	2015 model	2005 model	only $\Delta\ \alpha_{\mathcal{P}}$	2015 model
p25	9.94	10.22	10.07	10.27	10.74	10.48
p50	9.99	10.26	10.10	10.33	10.79	10.51
p75	10.04	10.30	10.14	10.39	10.84	10.55
p99	10.16	10.40	10.23	10.54	10.96	10.65

Table 6: Results on wage distribution from changing $\alpha_{\mathcal{P}}$ from 2005

benefits worker wages in both occupations, but decrease within-occupation inequality because workers now work for more equal firms. In a loose sense, everyone now works for good firms.

5.1.1 Quantity-biased technological change by occupation

These impact of quantity-biased technological bags the question how a change in each occupation matters. To understand this, we changing one $\alpha_{\mathcal{P}}$ at a time.

Table 7 illustrates that changes in the quantity-bias for white collar workers is more important for the change in the firm size distribution than changes for blue collar workers. This is the case even though the absolute and relative change in $\alpha_{\mathcal{P}}$ is larger for blue collar workers. The reason is that white collar workers were so constraint before, and since they are complements with blue collar workers, this amplifies.

	best fit 2005	only changing α_B	only changing α_W	best fit 2015
p25	7.77	7.79	7.72	8.03
p50	10.89	10.96	10.88	11.52
p75	15.65	15.85	15.83	17.16
p99	39.14	40.21	41.21	47.83

Table 7: Results on firm size distribution on changing one $\alpha_{\mathcal{P}}$ at a time from 2005

Table 8 documents that each occupation mostly benefits in terms of wage levels from changes in its own returns to scale, though there are some positive spillovers from increased returns to scale of the other occupation because of the complementarities between the two occupations.

So from the perspective of reducing wage inequality between groups, the change in scale for blue-collar workers is much more effective since they earn less across the wage distribution.

The impacts of quantity-bias in the calibrated model are economically large both for the firm size distribution as well as for wages. To judge its magnitude relative to other changes, we also investigate the partial contributions of other important variables in the next sections. We will foremost concentrate on skill-bias technological change, which in the model has a clear within-

	Blue collar workers				White collar workers			
	2005 model	$\Delta\ \alpha_B$ only	$\Delta\ \alpha_W$ only	2015 model	2005 model	$\Delta\ \alpha_B$ only	$\Delta\ \alpha_W$	2015 model
p25	9.94	10.16	9.99	10.07	10.27	10.54	10.46	10.48
p50	9.99	10.20	10.04	10.10	10.33	10.60	10.51	10.51
p75	10.04	10.24	10.09	10.14	10.39	10.66	10.56	10.55
p99	10.16	10.34	10.21	10.23	10.54	10.80	10.68	10.65

Table 8: Results on wage distribution from changing one $\alpha_{\mathcal{P}}$ at a time from 2005

occupation and across-occupation component that are analyzed in the next two sections. We then briefly touch on the remaining changes related to changing complementarities and spreads in the type distributions.

5.2 The contribution of intra-occupationalal skill-bias

One notion of skill-bias refers to the relevance of worker skills in the production function. When $\gamma_{\mathcal{P}}$ increases, it is more relevant to have a higher worker type. This can be best seen by considering the extreme of $\gamma_{\mathcal{P}} = 0$, which implies that all workers are equally productive in generating output. As this parameter increases, this shifts importance to the worker type, away from the firm type. Obviously this constitutes a within-occupation skill-bias, without obvious implications for firm size.

Changing only $\gamma_{\mathcal{P}}$ in both occupations indeed has little effect on firm size (see Table 9), yet it "rotates" wages around their occupational median (see Table 10). Since $\gamma_{\mathcal{P}}$ increases in blue and decreases in white collar occupations, this "steepens" the wage schedule for blue collar occupations and "flattens" it for white collar occupations. In terms of absolute magnitudes the changes for within-occupation wage inequality are comparable to those under quantity-bias, though obviously they have a different sign in blue collar jobs. Since the medians are not affected, it does not create additional across-occupation inequality at the median.

	best fit 2005	only $\Delta\ \gamma_{\mathcal{P}}$	best fit 2015
p25	7.77	7.76	8.03
p50	10.89	10.88	11.52
p75	15.65	15.64	17.16
p99	39.14	39.36	47.83

Table 9: Results on firm distribution from changing only skill-bias $\gamma_{\mathcal{P}}$ to 2015 levels.

	Blue collar workers			White collar workers		
	2005 model	only Δ $\gamma_{\mathcal{P}}$	2015 model	2005 model	only Δ $\gamma_{\mathcal{P}}$	2015 model
p25	9.94	9.93	10.07	10.27	10.30	10.48
p50	9.99	9.99	10.10	10.33	10.33	10.51
p75	10.04	10.05	10.14	10.39	10.36	10.55
p99	10.16	10.20	10.23	10.54	10.44	10.65

Table 10: Results on wage distribution from changing only skill-bias $\gamma_{\mathcal{P}}$ to 2015 levels.

5.3 Factor productivity and inter-occupational skill bias

The obvious other driver of inequality arises because of factor productivity changes in $\tilde{A}_{\mathcal{P}}$. This has both a simple level effect when factor productivity changes uniformly, and distributional effects when it changes more for one occupation than another.

In the calibration, factor productivity for white collar work increase, while it decreases for blue collar work. This spreads out firm sizes more (see Table 11), though much less than due to quantity-biased change. It arises here because white collar workers are relatively scarce (partially because of their low returns to scale) and the higher productivity in the white collar sector allows better firms to leverage across a larger number of blue collar workers.

	best fit 2005	only Δ $\tilde{A}_{\mathcal{P}}$	best fit 2015
p25	7.77	7.78	8.03
p50	10.89	10.93	11.52
p75	15.65	15.82	17.16
p99	39.14	40.59	47.83

Table 11: Results on wage distribution from changing $\tilde{A}_{\mathcal{P}}$ from 2005 to 2015

In terms of wage inequality, factor productivity has little effect on wage inequality within each occupation, but changes inter-occupational inequality substantially. To get a sense of the direction we should expect, consider our model but with all workers and firms being identical and no decreasing returns, which reduces it to Goldin and Katz (2009). We know for that model that the log-wage difference across occupations changes in the log-change of relative factor productivity with slope $(\rho - 1)/\rho$. With $\rho < 1$, this means that the larger output in one sector actually leads to relatively lower wages. This happens here (see Table 12) and across-occupational inequality reduces and actually reverses. White collar jobs are becoming so productive that they are less scarce and achieve lower remuneration.[13]

[13]Both the increase in $\tilde{A}_{\mathcal{W}}$ and the decrease in $\tilde{A}_{\mathcal{W}}$ in isolation have the effect of decreasing wages in the white

These effects are lower in absolute magnitude than the median wage effects induced by quantity-bias, though with differential sign for white collar workers. That reconciles this with our usual perception on renumeration across sectors: if one summes the effects at the median from quantity-bias (Table 6) and factor productivity (Table 12), white collar occupations remain higher paid overall.

	Blue collar workers			White collar workers		
	2005 model	only $\Delta\ \tilde{A}_{\mathcal{P}}$	2015 model	2005 model	only $\Delta\ \tilde{A}_{\mathcal{P}}$	2015 model
p25	9.94	10.07	10.07	10.27	9.94	10.48
p50	9.99	10.12	10.10	10.33	10.00	10.51
p75	10.04	10.17	10.14	10.39	10.06	10.55
p99	10.16	10.29	10.23	10.54	10.20	10.65

Table 12: Results on wage distribution from changing $\tilde{A}_{\mathcal{P}}$ from 2005 to 2015

5.4 Other changes: worker-firm complementarities, cross-type complementarities and type-distributions

Here we briefly review the impact of the remaining changes related to complementarities and type distributions.

Worker-firm complementarities $\sigma_{\mathcal{P}}$ have very minor effect on the firm size distribution and on inter-occupational wage ineqality at the median. The main consequence is increased wage dispersion among white collar wages, in a similar order of magnitude as that associated with quantity-bias. There is hardly any effect on wage inequality among blue collar workers. This is summarized in Tables 13 and 14 in Appendix A.5.[14]

Cross-type complementarity ρ across the services of blue and white collar workers has virtually no effect on firm sizes (Table 15 in Appendix A.5) despite the fact that it uniformly reduces wages among blue and among white collar workers (Table 16 in Appendix A.5).[15] The latter occurs because each output is more reliant on the other, making each individually less valuable at the margin to the firm. This hurts white collar workers less as their output is scarcer due in part to their lower returns to scale, which widens inequality between both groups. Effects are in absolute value smaller than those of quantity-bias, which acts in the opposite direction.

The distribution of skills has spead out more for white collar while it got more compressed for blue collar workers, and the spread in the firm type distribution remained virtually unchanged. This increased the spread in the firm size distribution as better firms now find more good white-

collar sector and increasing wages in the blue collar sector. Results omitted for brevity.

[14]This is also depicted in in Figure 11 .

[15]This is also summed up in Figure 13.

collar workers to hire, which complements with their blue collar workforce. The effect is quantitatively substantially smaller than the direct effect of quantity-biased change. This raised average wages and inequality within the white collar occupation, but decreased and compressed wages in the blue collar occupation. Though only the spread of the skill distribution is changed in this counter-factual, this still has average effects on wages in each occupation.[16] This is explained by the sorting in the model: in the extreme where all workers are identical they all get the average benefit from pairing with firms, but they loose the ability to sort and the associated gains. These effects can be seen in Appendix A.5) in Table 17 for firm size and in Table 18 for wages.[17]

5.5 Taking Stock

In our calibration, among all parameter changes in production over the 2005-2015 decade, quantity-bias had by far the largest effects on the firm size distribution. In absolute value it also had the largest effects on median wages in each occupation, and has among the largest impact on wage inequality across and within sectors. It increased wages for both occupations, while inducing a compression of wages within each sector because workers become more evenly matched to the good firms. Since white collar wages rise more on average, it increased across-occupation wage inequality.

These wage effects are in absolute value of similar magnitude as the effects of skill-biased technological change within and across sectors, though they act in opposite direction across occupations and within white collar jobs.

Quantity-bias affects wages by raising skill demand as firms are less constraint by decreasing returns, which raises pay in the affected occupations. In our setting it reduces inequality even within an occupation as more of the workers move into good firms. If such technological change can be fostered through government intervention, skill-bias for blue collar workers achieves not only efficiency gains but also equity gains as it reduces both within and across-occupational inequality. In that sense it is very different from automation, which tends to take tasks away from workers and reduces labor demand and pay.

6 Discussion and Future Evolution of Quantity-Biased Change

Our approach to quantity-bias takes the stance that the technology for communication and supervision is developed outside the firm and available at the same level for all firms. In our model, but in the quantitative part and in the micro-foundation, this implies that communication threshold $\alpha_{\mathcal{P}}$ applies to all firms, governing their scale economy in the same way. This links to the existing literature using related ideas, and captures the idea that outside entities beyond

[16] As explained earlier, the average is normalized to zero as any effects are picked up in factor productivity $\tilde{A}_{\mathcal{P}}$.

[17] This is also summed up in Figure 13.

the control of any single firm develop such systems. As eluded to before, cash registers now come regularly with the ability to scan employee codes, and so do scanners for stocking new material. This is technology is available to all firms.

In this interpretation it is interesting to ask which general-level technologies will arise in the future. Much of artificial intelligence is driven by reinforcement learning, where learning is most successful in places where most data is produced. In our model, this is exactly at the task threshold $\alpha_{\mathcal{P}}$, which divides tasks that are already easy to communicate from those that are not. Among the tasks that are not yet amenable to easy communication and monitoring, the largest part falls right on this threshold. Whether the number of workers in blue or white collar is larger, all else equal, determines where AI will progress quicker in shifting out the threshold. Given that α_W remains far below α_B in our calibration, there are more workers on the white collar threshold and this alone should shift attention to increases in α_W.

Obviously our division into white and blue collar is somewhat limiting, and future work might want to consider additional categories of workers, either because they fulfill more differentiated production tasks or because ICT has already progressed more for them as their tasks are more "routine" and can be easier communicated and monitored. In contrast to standard routine-bias, this is a good thing in quantity-bias because adding workers has less decreasing returns which raises labor demand in this sector in particular at high-productivity firms. Adding more types is conceptually straightforward though computationally somewhat difficult. The sorting condition in Lemma 2 and the differential equations and boundary conditions (12) to (15) in Section 2.2 remain unchanged, just the number of occupations $\mathcal{P}$ needs to be expanded.

The framework can easily be extended so that market power λ is type-dependent, for example because final demand is modeled as in Matsuyama and Ushchev (2017). Rending λ a function of firm type y might be necessary to address the data after 2015 where not only the average mark-up but also the dispersion of mark-ups explodes. Again, the techniques in this paper including the differential equations and boundary conditions (12) to (15) in Section 2.2 extend to this setting once sorting is guaranteed, and the conditions for sorting are unchanged as the proof of Lemma 2 is not affected. Obviously the exact way to parametrize the dependence of market power parameter λ on firm type needs further reflection.

A different view on quantity-bias arises when firms themselves invest in producing information and communcation technology. Then the threshold $\alpha_{mathcal P}$ becomes itsself a function of the firm type y, either exogeneously or through some optimal investment decision. In the exogeneous version, some firms are becoming larger not only when they are productive, but also when they have particularly favorable scale parameters. In the endogeneous version, if more productive firms invest more in communication and monitoring or have an advantage in using artificial intelligence because of the larger amount of data they produce inhouse, they can achieve a larger size both because of productive advantage and because of investments into favorable scaling technology. Solving the model under assortative matching remains the differential equations and boundary conditions (12) to (15), but the sorting conditions in Lemma 2 need to be reconsidered.

7 Conclusion

One consequence of technological change and the advent of digital technology is the ability to better communicate, supervise, and monitor. This paper takes the stance that this enables firms to supervise larger numbers of workers, as illustrated in our simple micro-foundation. In the quantitative exercise we attribute substantial effects to quantity-biased technological change over the decade 2005 to 2015: it was the dominant contributor to increasing size inequality between firms, and had a moderating influence on wage inequality within blue and white collar occupations, yet increased inequality between these occupations. Absolute effects on wages both within and across occupations due to quantity-bias alone are quantitatively large, even when compared to effects implied by skill-biased change.

The main goal of this paper is to highlight the importance of changes to the scale of firms, and its feedback to the remuneration of workers. We provide a model that can be scaled in a number of dimensions, such as adding more occupations. Our micro-foundation lays out a simple task model that is tractable enough to allow adaptations to more complicated settings. We extend sorting conditions to settings with decreasing returns. And we provide a first quantitative account that quantity-biased technological change is a first-order phenomenon not just for the distribution of firm size, but also for wage inequality between workers.

A Appendix

A.1 Appendix: Market Power and Hick-Neutral Capital

A.1.1 Market Power and Markups:

The goal of this subsection is to show that a firm with output F generates revenues of form

$$\text{constant} * (F)^{\lambda}$$

in a monopolistically competitive market, for some parameter $\lambda < 1$ and a constant that identical for all firms and which they take as given. So monopolistic competition can be incorporated into our setup through factor λ. Moverover,

$$\lambda = \frac{1}{markup},$$

where the markup is the firm's price over marginal cost. So information on average markups can discipline the choice of λ.

To show this, we consider a model of monopolistic competition in the Dixit-Stiglitz convention and review some of its basics. In particular, denote the output produced by firm y by $F(y)$. Consider a final goods producer who combines varieties of intermediate inputs into a final good Y according to CES production function

$$Y = \left[\int_y F(y)^{1/\gamma} dy\right]^{\gamma} \tag{16}$$

where $\sigma = \gamma/(\gamma - 1)$ represents the elasticity of substitution and $\gamma > 1$. The final goods producer faces a retail price P_r that is it takes a gives, and which could be determined endogeneously in a competitive goods market. Its profit maximization problem is given by

$$\max_{F(.)} P_r \left[\int_y F(y)^{1/\gamma} dy\right]^{\gamma} - \int_y P(y) F(y) dy \tag{17}$$

where $P(y)$ is the price of intermediate input firm y. The first order condition to (17) with respect to $F(y)$ requires almost everywhere

$$P_r \gamma \left[\int_y F(y)^{1/\gamma} dy\right]^{\gamma - 1} F(y)^{\frac{1}{\gamma} - 1} - P(y) = 0. \tag{18}$$

Raising both sides of (16) to the power $(\gamma - 1)/\gamma$ and substitution into (18) reveals

$$P_r \gamma Y^{\frac{\gamma - 1}{\gamma}} F(y)^{\frac{1 - \gamma}{\gamma}} - P(y) = 0, \tag{19}$$

which can be rearrange to display the demand function.[18] More relevant for our purpose, normalizing the price index $P_r = 1$, this can be rearranged to give us to the equilibrium price as a function of firm-specific output $F(y)$:

$$P(y) = F(y)^{\frac{1-\gamma}{\gamma}} Y^{-\frac{1-\gamma}{\gamma}}. \tag{20}$$

Rewriting the revenue function for the firm yields

$$\begin{aligned} & R(F(y)) \\ = \; & P(y)F(y) \\ = \; & Y^{-\frac{1-\gamma}{\gamma}} F(y)^{\frac{1}{\gamma}} \end{aligned} \tag{21}$$

So the revenue function equals to a concave transformation of output $F(y)^{\lambda}$ for some parameter $\lambda = 1/\gamma < 1$, multiplied by a constant that all firms take as given (which equals $Y^{-\frac{1-\gamma}{\gamma}}$).

Firm y with cost $C(F)$ then chooses output $F := F(y)$ according to optimization problem

$$\max_F R(F) - C(F)$$

with first order condition

$$\begin{aligned} & R'(F) = C'(F) \\ \Leftrightarrow \quad & Y^{-\frac{1-\gamma}{\gamma}} \frac{1}{\gamma} F^{\frac{1}{\gamma}-1} = C'(F) \\ \Leftrightarrow \quad & \frac{1}{\gamma} P(y) = C'(F) \\ \Leftrightarrow \quad & \gamma = P(y)/C'(F). \end{aligned} \tag{22}$$

The second line follows from differentiating (21) with respect to $F(y)$, and the third line from substitution of (20). So (22) revelas that the mark-up of price over marginal cost is equal to γ, which thereby is equal to $1/\lambda$. This relates the markup to the cuvature of the revenue function.

A.1.2 Hickneutral Capital in addition to Market Power

The goal of this subsection is to show that a firm that produces f with labor but then combines it with Hicks-neutral rented capital in a Cobb-Douglas aggregator to produce output optimally

[18] Rearranging yields a demand for each good of the form $F(y) = \left(\frac{P(y)}{P_r}\right)^{\frac{\gamma}{1-\gamma}} Y$. This equation says that the share of firm i in production is inversely proportional to its relative price. More expensive items are less demanded.

acts as if maximizes

$$\text{constant * } (f)^{\tilde{\lambda}} - \textit{ labor cost}$$

even in a monopolistically competitive market, for some parameter $\tilde{\lambda} < 1$ and a constant that identical for all firms and which they take as given. So the actual maximization problem of the firm looks identical to that of the firms in the main body of the paper, except for a particular specification of the outside curvature parameter.

To see this, assume that firms can hire Hick-neutral capital K at rental cost rK. Output of a firm is $f^{1-\beta}K^{\beta}$, where f is our production function (which depends on the firm type, the high and low skilled workers type, and the quantities of these workers). Importantly, one can solve for the optimal K as if f is already determined. Using the solution to that, we can solve the firms ex-ante problem where it determines the level of f, taking into account how this is later affects capital K. We also maintain that firms are monopolistically competitive as in the previous paragraph, with substitution parameter $\lambda := 1/\gamma < 1$ between varieties in the representative consumer's preferences.

The maximization problem of a firm of type y is then

$$\max_{K, x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}}} \tilde{A} \left[(\tilde{f}(x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}}, y)^{1-\beta} K^{\beta} \right]^{\lambda} - Kr - x_{\mathcal{B}} L_{\mathcal{B}} - x_{\mathcal{W}} L_{\mathcal{W}}$$

where the first term represents the revenues in the monopolistically competitive market and $\tilde{A} := Y^{1-\gamma}$ is the usual multiplier capturing the effects of all other firms in the monopolistically competitive market, while the last three terms capture the factor costs. Evidently this is equivalent to

$$\max_{x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}}} \left\{ \max_{K} \{ \tilde{A} \left[(\tilde{f}(x_{\mathcal{B}}, L_{\mathcal{B}}, x_{\mathcal{W}}, L_{\mathcal{W}}, y)^{1-\beta} K^{\beta} \right]^{\lambda} - Kr \} - x_{\mathcal{B}} L_{\mathcal{B}} - x_{\mathcal{W}} L_{\mathcal{W}} \right\}. \tag{23}$$

A first step in this firm's problem is to solve the interior problem for given output f :

$$\max_{K} \tilde{A} \tilde{f}^{(1-\beta)\lambda} K^{\beta\lambda} - Kr \tag{24}$$

Foc:

$$\beta\lambda \tilde{A} \tilde{f}^{(1-\beta)\lambda} K^{\beta\lambda - 1} = r \tag{25}$$

$$K = \left(\frac{r}{\beta\lambda \tilde{A} \tilde{f}^{(1-\beta)\lambda}} \right)^{1/(\beta\lambda - 1)} \tag{26}$$

substituting (25) back into the objective function (24) gives optimal profits of

$$\begin{aligned} & \tilde{A}\tilde{f}^{(1-\beta)\lambda}K^{\beta\lambda} - \beta\lambda\tilde{A}\tilde{f}^{(1-\beta)\lambda}K^{\beta\lambda} \\ = \; & (1-\beta\lambda)\tilde{A}\tilde{f}^{(1-\beta)\lambda}K^{\beta\lambda} \end{aligned} \tag{27}$$

Then substituting (26) into the optimal profit (27) gives

$$\begin{aligned} & (1-\beta\lambda)\tilde{A}\tilde{f}^{(1-\beta)\lambda}\left(\frac{r}{\beta\lambda\tilde{A}\tilde{f}^{(1-\beta)\lambda}}\right)^{\beta\lambda/(\beta\lambda-1)} \\ = & \left[(1-\beta\lambda)\tilde{A}\left(\frac{r}{\beta\lambda\tilde{A}}\right)^{\beta\lambda/(\beta\lambda-1)}\right]\tilde{f}^{(1-\beta)\lambda}\tilde{f}^{(1-\beta)\lambda\beta\lambda/(1-\beta\lambda)} \\ = & \left[(1-\beta\lambda)\tilde{A}\left(\frac{r}{\beta\lambda\tilde{A}}\right)^{\beta\lambda/(\beta\lambda-1)}\right]\tilde{f}^{(1-\beta)\lambda[1+\beta\lambda/(1-\beta\lambda)]} \\ = & \underbrace{\left[(1-\beta\lambda)\tilde{A}\left(\frac{r}{\beta\lambda\tilde{A}}\right)^{\beta\lambda/(\beta\lambda-1)}\right]}_{:=\hat{A}}\tilde{f}^{(1-\beta)\lambda/(1-\beta\lambda)} \end{aligned}$$

Cearly exponent

$$\tilde{\lambda}(\beta,\lambda) = (1-\beta)\,\lambda/(1-\beta\lambda) \tag{28}$$

is in $(0,1)$ and represents the additional curvature in the revenue function that comes from the introduction of captial and monopolistic competition. We can therefore represent the firms problem (23) as

$$\max_{x_{\mathcal{B}},L_{\mathcal{B}},x_{\mathcal{W}},L_{\mathcal{W}}} f(x_{\mathcal{B}},L_{\mathcal{B}},x_{\mathcal{W}},L_{\mathcal{W}},y) - x_{\mathcal{B}}L_{\mathcal{B}} - x_{\mathcal{W}}L_{\mathcal{W}}$$

with adjusted "production" function

$$f(x_{\mathcal{B}},L_{\mathcal{B}},x_{\mathcal{W}},L_{\mathcal{W}},y) = \hat{A}f(x_{\mathcal{B}},L_{\mathcal{B}},x_{\mathcal{W}},L_{\mathcal{W}})^{\tilde{\lambda}}. \tag{29}$$

Clearly this falls within the same class of production functions we have considered, just now some of the parameters capture not only production, but also the effects of capital and market power.

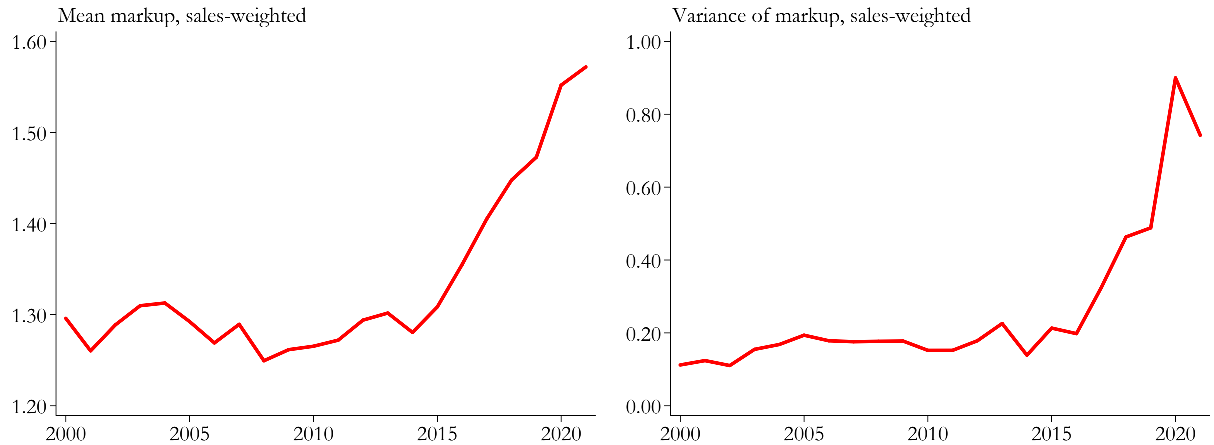
A.2 Additional Figures on Markups in Germany

Figure 2 showed first and second moments for cost-weighted markups over time in Germany. Here, Figure 5 show analogous figures for sales-weighted markups. The broad insights remain: average markups are very stable over time from 2005 to 2015, with stable and low variance. Both

Figure 5: Mark-ups in Germany over time (sales-weighted).

(a) Average mark-up over time

(b) Variance of mark-ups over time



the level and the variance of markups explodes after 2015.

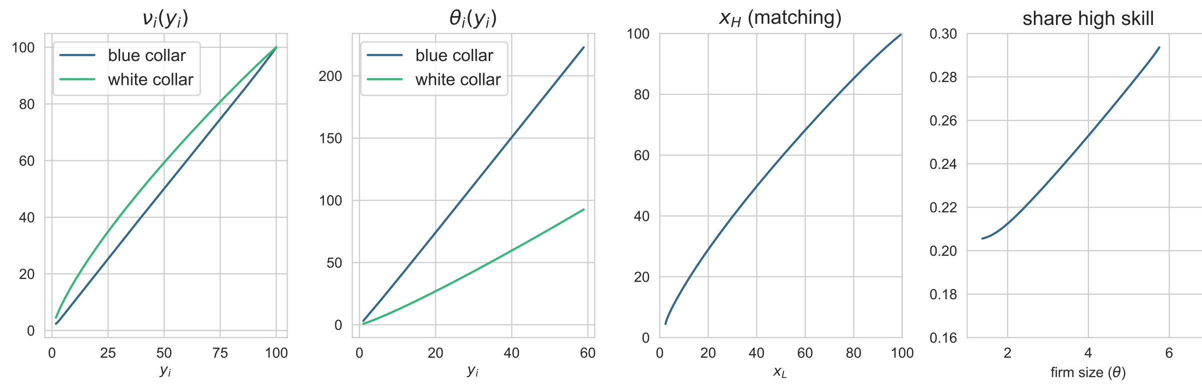


Figure 6: 2005 solution

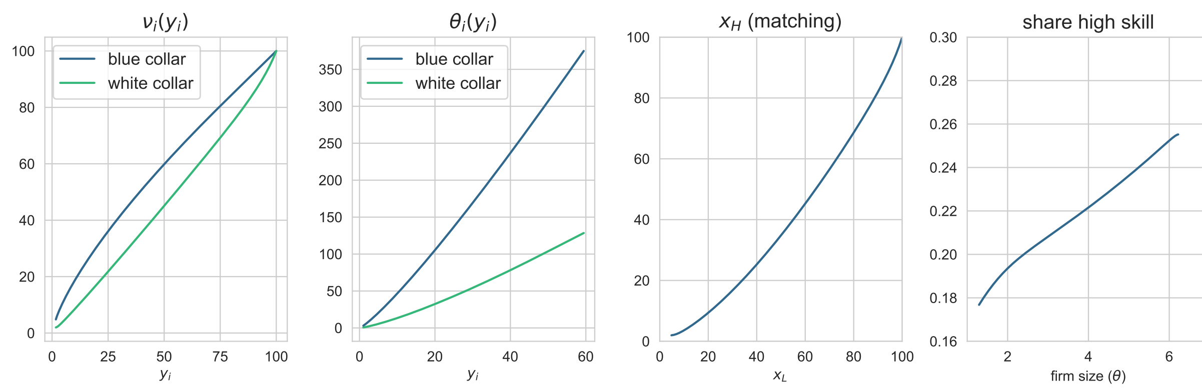


Figure 7: 2015 solution

A.3 Graphical Representation of Equilibrium Allocations

Figures 6 and 7 illustrate the equilibrium allocations.

A.4 Appendix: Differential Equation

Here we derive the differential equation for firm sizes. In the following, subscripts denote partial derivatives to the variable (except for the indicator $\mathcal{P} \in \{\mathcal{B}, \mathcal{W}\}$ which denotes a profession and $-P \in \{\mathcal{B}, \mathcal{W}\}/\{\mathcal{P}\}$ which denotes the other odcupation). Recall the first order conditions for a firm in profession $\mathcal{P}$:

$$\begin{aligned} f_{x_\mathcal{P}} \frac{1}{\theta_\mathcal{P}} &= \omega'_\mathcal{P} \\ f_{L_\mathcal{P}} &= \rightarrow_\mathcal{P} \end{aligned}$$

Full differentiation of the second condition with respect to y :

$$f_{L_\mathcal{P} x_\mathcal{P}} \mu'_\mathcal{P} + f_{L_\mathcal{P} L_\mathcal{P}} \theta'_\mathcal{P} + f_{L_\mathcal{P} x_{-\mathcal{P}}} \mu'_{-\mathcal{P}} + f_{L_\mathcal{P} L_{-\mathcal{P}}} \theta'_{-\mathcal{P}} + f_{L_\mathcal{P} y} = \omega'_\mathcal{P} \mu'_\mathcal{P}$$

Using market clearing $\mu'_\mathcal{P} = \theta_\mathcal{P} \mathcal{H}_\mathcal{P}$, and similar for other occupation:

$$f_{L_\mathcal{P} x_\mathcal{P}} \theta_\mathcal{P} \mathcal{H}_\mathcal{P} + f_{L_\mathcal{P} L_\mathcal{P}} \theta'_\mathcal{P} + f_{L_\mathcal{P} x_{-\mathcal{P}}} \theta_{-\mathcal{P}} \mathcal{H}_{-\mathcal{P}} + f_{L_\mathcal{P} L_{-\mathcal{P}}} \theta'_{-\mathcal{P}} + f_{L_\mathcal{P} y} = \omega'_\mathcal{P} \theta_\mathcal{P} \mathcal{H}_\mathcal{P}$$

Deviding through by other first order condition to get rid of $\omega'_\mathcal{P}$: $\frac{\theta_\mathcal{P}}{f_{x_\mathcal{P}}}$

$$f_{L_\mathcal{P} x_\mathcal{P}} \theta_\mathcal{P} \mathcal{H}_\mathcal{P} + f_{L_\mathcal{P} L_\mathcal{P}} \theta'_\mathcal{P} + f_{L_\mathcal{P} x_{-\mathcal{P}}} \theta_{-\mathcal{P}} \mathcal{H}_{-\mathcal{P}} + f_{L_\mathcal{P} L_{-\mathcal{P}}} \theta'_{-\mathcal{P}} + f_{L_\mathcal{P} y} = f_{x_\mathcal{P}} \mathcal{H}_\mathcal{P}$$

For other profession we obtain similarly

$$f_{L_{-\mathcal{P}} x_{-\mathcal{P}}} \theta_{-\mathcal{P}} \mathcal{H}_{-\mathcal{P}} + f_{L_{-\mathcal{P}} L_{-\mathcal{P}}} \theta'_{-\mathcal{P}} + f_{L_{-\mathcal{P}} x_\mathcal{P}} \theta_\mathcal{P} \mathcal{H}_\mathcal{P} + f_{L_{-\mathcal{P}} L_\mathcal{P}} \theta'_\mathcal{P} + f_{L_{-\mathcal{P}} y} = f_{x_{-\mathcal{P}}} \mathcal{H}_{-\mathcal{P}}$$

Multiplying the latter by $\frac{f_{L_\mathcal{P} L_{-\mathcal{P}}}}{f_{L_{-\mathcal{P}} L_{-\mathcal{P}}}}$ and deducting it from the former eliminate $\theta'_{-\mathcal{P}}$ and we get (after again multiplying the resulting equation by $f_{L_{-\mathcal{P}} L_{-\mathcal{P}}}$) :

$$\begin{aligned} &\left[\begin{array}{c} \left[f_{L_{-\mathcal{P}} L_{-\mathcal{P}}} f_{L_\mathcal{P} x_\mathcal{P}} - f_{L_\mathcal{P} L_{-\mathcal{P}}} f_{L_{-\mathcal{P}} x_\mathcal{P}}\right] \theta_\mathcal{P} \mathcal{H}_\mathcal{P} + \left[f_{L_{-\mathcal{P}} L_{-\mathcal{P}}} f_{L_\mathcal{P} L_\mathcal{P}} - f_{L_\mathcal{P} L_{-\mathcal{P}}} f_{L_{-\mathcal{P}} L_\mathcal{P}}\right] \theta'_\mathcal{P} \\ + \left[f_{L_{-\mathcal{P}} L_{-\mathcal{P}}} f_{L_\mathcal{P} x_{-\mathcal{P}}} - f_{L_\mathcal{P} L_{-\mathcal{P}}} f_{L_{-\mathcal{P}} x_{-\mathcal{P}}}\right] \theta_{-\mathcal{P}} \mathcal{H}_{-\mathcal{P}} + \left[f_{L_{-\mathcal{P}} L_{-\mathcal{P}}}\right] f_{L_\mathcal{P} y} - f_{L_\mathcal{P} L_{-\mathcal{P}}} f_{L_{-\mathcal{P}} y} \end{array} \right] \\ = \;& f_{L_{-\mathcal{P}} L_{-\mathcal{P}}} f_{x_\mathcal{P}} \mathcal{H}_\mathcal{P} - f_{L_\mathcal{P} L_{-\mathcal{P}}} f_{x_{-\mathcal{P}}} \mathcal{H}_{-\mathcal{P}} \end{aligned}$$

Rearranging for $\theta'_\mathcal{P}$ delivers:

$$\theta'_{\mathcal{P}} = \frac{\left[\begin{array}{c} -\left[f_{L_{-\mathcal{P}}L_{-\mathcal{P}}} f_{L_{\mathcal{P}}x_{\mathcal{P}}} - f_{L_{\mathcal{P}}L_{-\mathcal{P}}} f_{L_{-\mathcal{P}}x_{\mathcal{P}}}\right] \theta_{\mathcal{P}} \mathcal{H}_{\mathcal{P}} - \left[f_{L_{-\mathcal{P}}L_{-\mathcal{P}}} f_{L_{\mathcal{P}}x_{-\mathcal{P}}} - f_{L_{\mathcal{P}}L_{-\mathcal{P}}} f_{L_{-\mathcal{P}}x_{-\mathcal{P}}}\right] \theta_{-\mathcal{P}} \mathcal{H}_{-\mathcal{P}} \\ -\left[f_{L_{-\mathcal{P}}L_{-\mathcal{P}}}\right] f_{L_{\mathcal{P}}y} + f_{L_{\mathcal{P}}L_{-\mathcal{P}}} f_{L_{-\mathcal{P}}y} + f_{L_{-\mathcal{P}}L_{-\mathcal{P}}} f_{x_{\mathcal{P}}} \mathcal{H}_{\mathcal{P}} - f_{L_{\mathcal{P}}L_{-\mathcal{P}}} f_{x_{-\mathcal{P}}} \mathcal{H}_{-\mathcal{P}} \end{array}\right]}{f_{L_{-\mathcal{P}}L_{-\mathcal{P}}} f_{L_{\mathcal{P}}L_{\mathcal{P}}} - f_{L_{\mathcal{P}}L_{-\mathcal{P}}} f_{L_{-\mathcal{P}}L_{\mathcal{P}}}}$$

No cross-partials between $\mathcal{P}$ and -$\mathcal{P}$ variables:

$$\theta'_{\mathcal{P}} = \frac{\left[f_{x_{\mathcal{P}}} - f_{L_{\mathcal{P}}x_{\mathcal{P}}} \theta_{\mathcal{P}}\right] \mathcal{H}_{\mathcal{P}} - f_{L_{\mathcal{P}}y}}{f_{L_{\mathcal{P}}L_{\mathcal{P}}}}$$

A.5 Appendix: More on counterfactuals.

This section lists more graphs and figures for the counterfactuals in Section 5.

A.5.1 More on quantity-bias

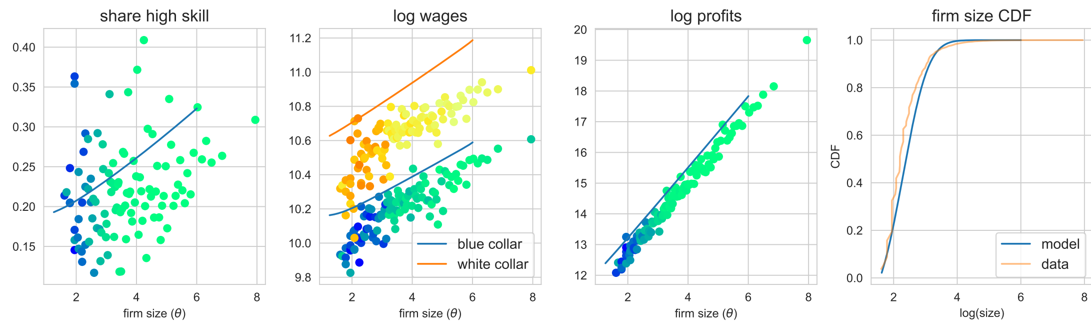


Figure 8: Fit to the 2015 data when only span-of-control ($\alpha_{\mathcal{P}}$) is changed to 2015 levels, keeping all other parameters at 2005 levels.

A.5.2 More on intra-occupation skill-bias

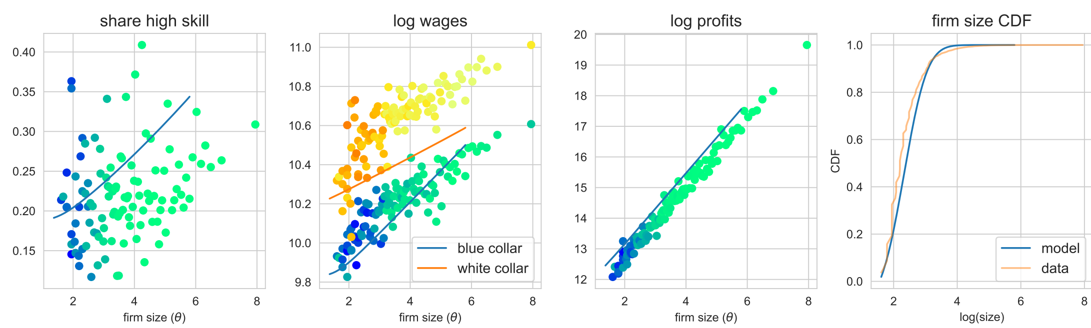


Figure 9: Fit to the 2015 data when only skill complementarity ($\gamma_{\mathcal{P}}$)

A.5.3 More on TFP and inter-occupation skill-bias

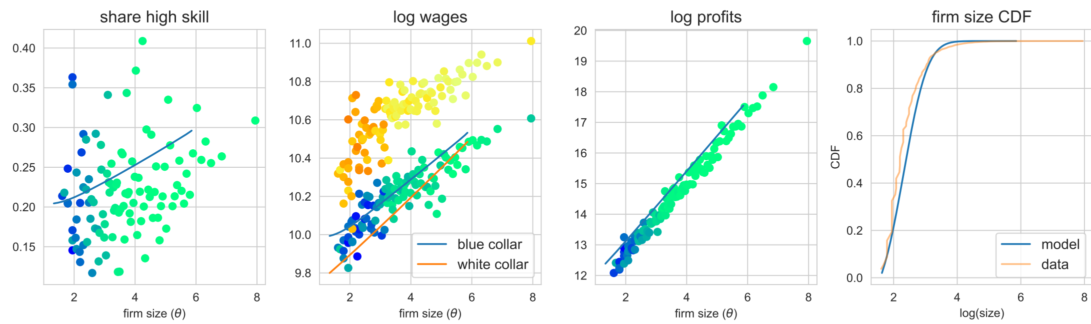


Figure 10: Fit to the 2015 data when only TFP ($\tilde{A}_{\mathcal{P}}$)

A.5.4 More on worker-firm-complementarity (changing $\sigma_{\mathcal{P}}$)

	best fit 2005	only Δ $\sigma_{\mathcal{P}}$	best fit 2015
p25	7.77	7.87	8.03
p50	10.89	10.99	11.52
p75	15.65	15.69	17.16
p99	39.14	37.82	47.83

Table 13: Results on firm size distribution from changing $\sigma_{\mathcal{P}}$ from 2005 to 2015

	Blue collar workers			White collar workers		
	2005 model	only Δ $\sigma_{\mathcal{P}}$	2015 model	2005 model	only Δ $\sigma_{\mathcal{P}}$	2015 model
p25	9.94	9.94	10.07	10.27	10.24	10.48
p50	9.99	9.99	10.10	10.33	10.32	10.51
p75	10.04	10.03	10.14	10.39	10.41	10.55
p99	10.16	10.15	10.23	10.54	10.67	10.65

Table 14: Results on wage distribution from changing $\sigma_{\mathcal{P}}$ from 2005 to 2015

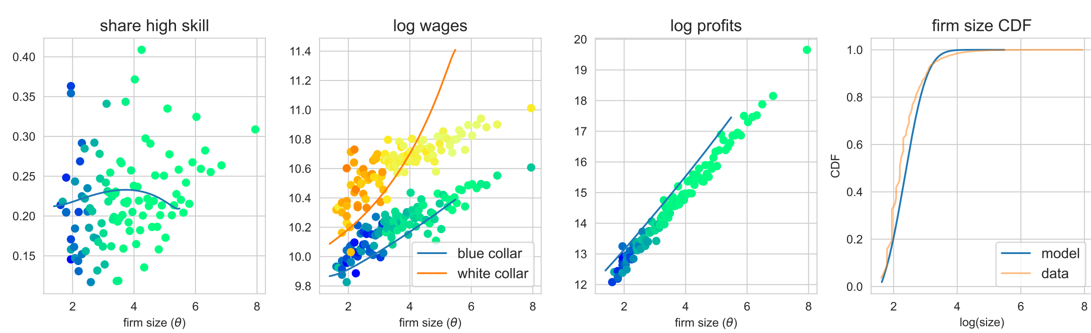


Figure 11: Fit to the 2015 data when only the firm-worker complementarities ($\sigma_{\mathcal{P}}$) from 2005 to 2015 values.

A.5.5 More on changing cross-type complementarity ρ

	best fit 2005	only $\Delta\ \rho$	best fit 2015
p25	7.77	7.78	8.03
p50	10.89	10.89	11.52
p75	15.65	15.65	17.16
p99	39.14	39.14	47.83

Table 15: Results on wage distribution from changing only ρ from 2005 to 2015

	Blue collar workers			White collar workers		
	2005 model	only $\Delta\ \rho$	2015 model	2005 model	only $\Delta\ \rho$	2015 model
p25	9.94	9.71	10.07	10.27	10.17	10.48
p50	9.99	9.76	10.10	10.33	10.23	10.51
p75	10.04	9.81	10.14	10.39	10.28	10.55
p99	10.16	9.94	10.23	10.54	10.43	10.65

Table 16: Results on wage distribution from changing only ρ from 2005 to 2015

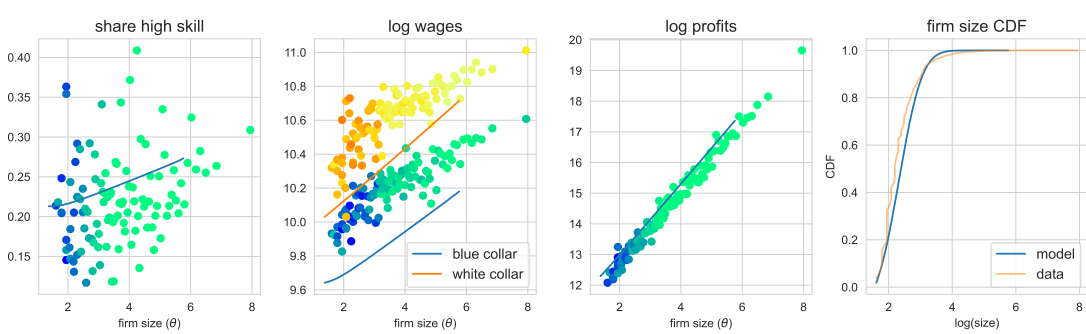


Figure 12: Fit to the 2015 data when only changing type complementarity (ρ) from 2005 to 2015 value, all other variable remain at 2005 value

A.5.6 Changing the distribution of skill only

	best fit 2005	only $\Delta\ s_{\mathcal{B}}, s_{\mathcal{W}}, s_y$	best fit 2015
p25	7.77	8.10	8.03
p50	10.89	11.47	11.52
p75	15.65	16.60	17.16
p99	39.14	42.02	47.83

Table 17: Results on wage distribution from only changing skill distributions from 2005 to 2015

	Blue collar workers			White collar workers		
	2005 model	only $\Delta\ s_{\mathcal{B}}, s_{\mathcal{W}}, s_y$	2015 model	2005 model	only $\Delta\ s_B, s_W, s_P$	2015 model
p25	9.94	9.86	10.07	10.27	10.31	10.48
p50	9.99	9.90	10.10	10.33	10.39	10.51
p75	10.04	9.93	10.14	10.39	10.47	10.55
p99	10.16	10.03	10.23	10.54	10.67	10.65

Table 18: Results on wage distribution from only changing skill distributions from 2005 to 2015

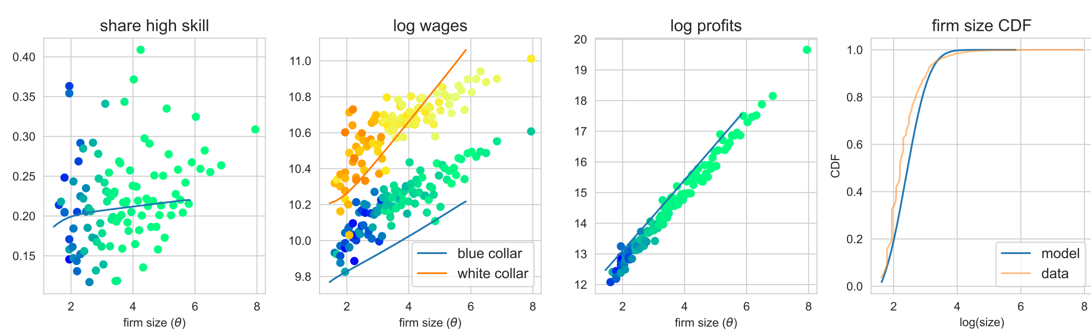


Figure 13: Fit to the 2015 data when only changing skill distribution dispersion from 2005 to 2015, holding all other parameters at 2005 level

References

ACEMOGLU, D. AND D. AUTOR (2011): "Skills, tasks and technologies: Implications for employment and earnings," in *Handbook of labor economics*, Elsevier, vol. 4, 1043–1171.

BAGGER, J. AND R. LENTZ (2019): "An empirical model of wage dispersion with sorting," *The Review of Economic Studies*, 86, 153–190.

BECKER, G. (1973): "A Theory of Marriage I," *Journal of Political Economy*, 81, 813–846.

BONHOMME, S., T. LAMADON, AND E. MANRESA (2019): "A distributional framework for matched employer employee data," *Econometrica*, 87, 699–739.

CARD, D., J. HEINING, AND P. KLINE (2013): "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly journal of economics*, 128, 967–1015.

CASELLI, F. AND W. J. COLEMAN (2006): "The world technology frontier," *American Economic Review*, 96, 499–522.

CHADE, H., J. EECKHOUT, AND L. SMITH (2017): "Sorting through search and matching models in economics," *Journal of Economic Literature*, 55, 493–544.

CHADE, H. AND P. KIRCHER (2023): "Matching with Search Frictions," *Online and Matching-Based Market Design*, 402.

COSTINOT, A. (2009): "An elementary theory of comparative advantage," *Econometrica*, 77, 1165–1192.

COSTINOT, A. AND J. VOGEL (2010): "Matching and inequality in the world economy," *Journal of Political Economy*, 118, 747–786.

——— (2015): "Beyond Ricardo: Assignment models in international trade," *economics*, 7, 31–62.

DE LOECKER, J. AND J. EECKHOUT (2018): "Global market power," Tech. rep., National Bureau of Economic Research.

DUSTMANN, C., J. LUDSTECK, AND U. SCHÖNBERG (2009): "Revisiting the German wage structure," *The Quarterly journal of economics*, 124, 843–881.

EECKHOUT, J. AND P. KIRCHER (2011): "Identifying sorting—in theory," *The Review of Economic Studies*, 78, 872–906.

——— (2018): "Assortative Matching with Large Firms," *Econometrica*, 86, 85–132.

GABAIX, X. AND A. LANDIER (2008): "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics*, 123, 49–100.

GARICANO, L. (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of political economy*, 108, 874–904.

GARICANO, L. AND E. ROSSI-HANSBERG (2006): "Organization and inequality in a knowledge economy," *The Quarterly journal of economics*, 121, 1383–1435.

GOLDIN, C. AND L. F. KATZ (2009): *The race between education and technology*, harvard university press.

GROSSMAN, G. M., E. HELPMAN, AND P. KIRCHER (2017): "Matching, sorting, and the distributional effects of international trade," *Journal of political economy*, 125, 224–264.

GROSSMAN, G. M. AND G. MAGGI (2000): "Diversity and trade," *American Economic Review*, 90, 1255–1275.

HOPENHAYN, H. A. (1992): "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, 1127–1150.

KATZ, L. F. AND K. M. MURPHY (1992): "Changes in relative wages, 1963–1987: supply and demand factors," *The quarterly journal of economics*, 107, 35–78.

LUCAS JR, R. E. (1978): "On the size distribution of business firms," *The Bell Journal of Economics*, 508–523.

MATSUYAMA, K. AND P. USHCHEV (2017): "Beyond CES: three alternative classes of flexible homothetic demand systems," *Global Poverty Research Lab Working Paper*.

REICHARDT, H. (2024): "Scale-Biased Technical Change and Inequality," Tech. rep., Working Paper.

SONG, J., D. J. PRICE, F. GUVENEN, N. BLOOM, AND T. VON WACHTER (2019): "Firming up inequality," *The Quarterly journal of economics*, 134, 1–50.

STEVENSON, R. (1980): "Measuring technological bias," *The American Economic Review*, 70, 162–173.

TERVIO, M. (2008): "The Difference that CEOs Make: An Assignment Model Approach," *American Economic Review*, 98(3), 642–668.