DEPARTMENT
OF ECONOMICS

Team Production and Firm Dynamics with Search and Matching Models

Gabriele Macci

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Economics
of the European University Institute

Florence, 10 October 2023

European University Institute
Department of Economics

Team Production and Firm Dynamics with Search and Matching Models

Gabriele Macci

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Economics
of the European University Institute

Examining Board

Prof. Philipp Kircher, Cornell University, Supervisor
Prof. Russell W. Cooper, EUI, Co-Supervisor
Prof. Pieter Gautier, Vrije Universiteit Amsterdam
Prof. Leo Kaas, Goethe University Frankfurt

Researcher declaration to accompany the submission of written work Department Economics - Doctoral Programme

I Gabriele Macci certify that I am the author of the work "Team Production and Firm Dynamics with Search and Matching Models"I have presented for examination for the Ph.D. at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the Code of Ethics in Academic Research issued by the European University Institute (IUE 332/2/10 (CA 297).

The copyright of this work rests with its author. Quotation from it is permitted, provided that full acknowledgement is made. This work may not be reproduced without my prior written consent. This authorisation does not, to the best of my knowledge, infringe the rights of any third party.

I declare that this work consists of approximately 32.000 words.

Statement of inclusion of previous work:

I confirm that chapter 1 was jointly co-authored with Ms. Cristina Lafuente, and I contributed 70% of the work.
I confirm that chapter 2 was jointly co-authored with Ms. Cristina Lafuente, and I contributed 60% of the work.
I confirm that chapter 1 draws upon an earlier article I published "Corrigendum to Capital Investment in “Assortative Matching with Large Firms”", Econometrica, 2021, vol. 89(4), pages o11-o14.

Signature and date:
March 30, 2023

Abstract

This thesis investigates various aspects of the effects of technological change, capital subsidies, and firm dynamics on labor reallocation, wage inequality, the gender pay gap and, coworker learning in Italy.

The first chapter uses a matching model that partitions the workforce into white and blue-collar workers, and firms hire both types while optimizing their labor composition based on a quality-quantity trade-off. Technological change leads to increased capital investment, affecting labor composition and wage distribution. Results from simulation modeling on Italian data show that worker-capital complementarities drive inequality changes within firms, while worker-capital complementarities, worker-teammates complementarity, and compositional effects account for shifts in the between-firms component.

The second chapter further investigates the impact of capital subsidies on the gender pay gap, highlighting the distribution of the pay gap between and within firms over time. A theoretical model explains the connection between pay gap distribution, capital intensity and skill pay gap. Reducing capital rental prices can decrease gender pay gap by allowing female workers to sort into better-matched jobs at the expense of an increased skill gap.

The last chapter explores coworker learning opportunities, acknowledging firms as the ultimate arrangers of a worker's group of colleagues. A firm dynamics general equilibrium model with human capital spillovers among colleagues is used. The model human capital allocation is efficient, as it balances learning and production complementarities based on firm productivity level. Using employer-employee data, the study reveals stylized facts on coworker spillovers conditional on firm growth and size, emphasizing the importance of a large pool of small growing firms to generate coworker learning opportunities.

Acknowledgements

I am indebted to EUI and its community for providing me with such an amazing learning environment. First of all, I am indebted with Philipp and Russell for the constant support. I also want to thank my colleagues, especially Cristina, all the rest of the faculty and the admin team of VLF, and the numerous speakers I interacted with during these many years.

A special thank goes to my wife Chiara and all my family.

Last, I thank Fondazione Rodolfo De Benedetti and INPS for making me available the data used for this thesis.

Contents

Chapter 1

Firm Wage Inequality: Capital-Skill Complementarity and Labor Reallocation

Cristina Lafuente & Gabriele Macci

Abstract

This paper studies how technological change affects labor reallocation in the economy and how the latter feeds into changes over wage inequality. In our model, the workforce is partitioned into white and blue-collars and, within each category, workers are heterogeneous in their quality. Firms operate by hiring both white and blue-collars and differ one another on their labor composition on two margins: the quality and the intensity of their labor factors. In addition, each firm chooses capital optimally to fit with its labor force. Technological change encourages firms to invest in more capital, which feeds back into their labor composition decisions, and ultimately in the distribution of wages across firms in the economy. We compare our simulated model to Italian data for 1998 and 2001. The model qualitatively replicates the data increased (decreased) inequality within (between) firms. We decompose inequality changes into three channels, namely worker-capital complementarity, worker-teammates complementarity and a compositional effect. Worker-capital complementarities are the dominant source of changes in inequality for the within firm component while all three channels are important to account for the shifts in the between firms component of inequality.

1.1 Introduction

How does technological change affect the reallocation of labor across firms? And how does this translate into wage inequality? Despite many studies have found that more polarized labor composition across establishments explains most of the rise in wage inequality over the last few decades —see Song et al. (2019) for US from 1981 to 2013, Card et al. (2013) for

Germany from 1985 to 2009 and, Skans et al. (2007) for Sweden from 1985 to 2000— none of those studies analyses how labor reallocates across firms as a consequence of technological change.[1] However, labor allocations, and in particular the workforce composition of a firm, is directly affected by technological change because the latter alters the technical rate of substitution *i)* among labor and capital, *ii)* among different occupations and, *iii)* among quality and quantity within each occupation.[2] In this paper, we develop and test a model of labor (re-)allocation and wage determination that explicitly allows for technological change to affect workforce composition within and between firms in the economy.

The literature on rising wage dispersion and its firm-level decomposition, so far, has been primarily empirical. Even though a comparison of the results across-countries is not straightforward due to different methodologies and national data sources, a recurrent finding is that firm-level changes in labor composition explain most of the variation in wage dispersion (see for instance Song et al. 2019 and the references therein). Here, we advance an interpretation for this regularity by proposing a new theoretical framework that links together firm hiring decisions with technological change.

The interaction between labor reallocation and capital investment is at the core of our interpretation. For about half a century, technological change led to a substitution of labor for capital (Acemoglu and Autor, 2011). The thesis of our paper is that not only the new capital has been biased in favor of white-collar occupations —so as to affect the relative quantity demanded of white-collar workers versus blue-collar workers—, but it has also been *talent-biased*. In other words, for a given occupation, the degree at which capital has substituted for labor has been inversely proportional to the talent of the workers employed. As a consequence, lower-quality workers (less talented), either white or blue-collar, were more frequently displaced and tended to match together, ending up working for the same firms. This has determined an increase in between firm segregation over the quality of their workforce. This quality segregation has caused an upsurge in overall inequality due to rising between firm inequality.

The seminal paper of Krusell et al. (2000) showed how skill-biased technological change can lead to an increase in inequality overall in the economy. The aim of our work is to detail such an increase at an higher resolution. With higher resolution, we mean looking at firm-level evolution of inequality. To do this, we will follow the set-up proposed by Eeckhout and Kircher (2018)[3] augmented by two capital factors, namely structures and equipment. In line with Krusell et al. (2000), technological change is proxied by the faster growth of equipment assets relative to structures and labor supply.

Trade literature (see for instance Costinot and Vogel 2010, Sampson 2016 and, Grossman et al. 2017) has vastly documented the repercussions of sorting and matching on inequality. A major contribution of this literature is to develop a framework that distinguishes two forces driving changes into inequality: *i)* variations in wages (the price of factors) and *ii)* variations in the allocation of workers (the quantity and quality of factors). This project also starts from the same step, decomposing inequality changes into those two forces. However, as a further step with respect to this literature, the presence of capital in the model allows

[1]Sector-level studies like Haltiwanger et al. (2022) and Briskar et al. (2022) argue that industries differences in wage inequality are first-order compared to within-sector firm-level dispersion.

[2]Generally, the quality and the quantity of work required to perform a job are not perfect substitutes unless the job can be represented in terms of efficiency units. Therefore, in general, there is a trade-off between hiring more versus hiring better workers (Eeckhout and Kircher, 2018).

[3]Instead of the original two-sided market for workers and managers, we consider a formulation where workers are on both sides of the market.

to trace back how those two forces varied as a consequence of technological change. To do this, we develop a structural model that explicitly accounts for changes in labor composition across firms. In fact, just adopting a reduced-form approach, as in Song et al. (2019), would prove limited in disentangling how technological change influenced wages because it does not consider how labor reallocates as a response to the technological progress. We show that in a sufficiently rich framework, where a worker marginal product depends also on their own colleague type and quantity, labor reallocation plays a substantial role in shaping the wage profile of the economy after a change in capital investment.

The model is not properly calibrated, yet we report the outcomes from an informed simulation. The model parameters are chosen after comparing to an Italian administrative dataset from Veneto region (Veneto Workers History), for the years 1998 and 2001. We use moments from 1998 for a first calibration of the model to show its qualitative properties.[4] In particular, we use the calibrated model to decompose the effects of wage inequality originating from a decrease in the price of equipment. We keep the 1998 parameters constant and simulate a decrease in the price of equipment which proxies technological change occurring between 1998 and 2001. We do this in the spirit of Krusell et al. (2000) as a proxy to isolate the effects of the technological change channel on wage inequality. First, as in the data, the simulated wages increase in their variance when equipment becomes cheaper. Also qualitatively similar to the data, the wage dispersion between firms lowers while within firms gets higher. Second, we use the theoretical structure of our theory to account for how much of these changes are attributable to capital-labor complementarities and teammates complementarities, the residual being a compositional effect. We argue that capital complementarities are first order in explaining the rise in inequality occurring within firms. At the same time, our theory suggests that all three channels are relevant for explaining the decrease in between firm wage dispersion.

The rest of the paper proceeds as follows. Section 1.2 introduces the model and discusses at length the production function specification adopted for the quantitative exercise. In particular, section 1.2.3 describes how to decompose variations in wage inequality using our model, the exact formulas are then reported in appendix A.2. In section 1.3 we describe our data, how the equilibrium looks like for the calibrated version of our model and, a structural decomposition of the wage variance. Finally section 1.4 concludes and outlines future research directions.

1.2 Model

Our environment can be regarded as a specific application of the framework conceptualized in Eeckhout and Kircher (2018). Their paper is the first to investigate a firm trade-off between hiring more versus better workers in a competitive economy with two-sided heterogeneity. Such framework is especially suitable for studying inequality (Grossman et al., 2017) because a worker's marginal product is not just a function of his own type but also depends on the type and quantity of his colleagues. Consequently, fairly general wage distributions can be obtained when the production structure is rich enough.

The production function we consider fits well into the logic of Krusell et al. (2000) who

[4] Please note that the simulation in its current stage is informed by the data but still distant in several regards. As such, the quantitative conclusions of the simulation are intended for illustration purposes. A full estimation with a better quantitative fit to the data will be available in the next update of this project.

study wage inequality evolution as a consequence of technological change. The main success of Krusell et al. (2000) has been replicating the evolution of US skill-premium with a relatively parsimonious model. They partition the labor force into blue and white collars, and look at how each labor group's respective marginal product is affected by its complementarity with capital equipment.[5] Motivated by their success, our model assumes heterogeneity within both blue and white-collars and exploits the methodology developed by Eeckhout and Kircher (2018) to determine how wage dispersion evolves within and between firms as a reaction to an incentive in investing more into capital equipment.

1.2.1 Environment

The environment is static and can be regarded as a stationary economy in which any intertemporal decision has implicitly collapsed into a present value formulation. The environment abstracts from any type of friction, in particular labor search frictions and capital adjustment costs.[6]

Production is organized into *teams*. A team is a collection of blue and white-collar workers equipped with a certain amount of capital. There are two types of capital, namely equipment, q, and structures, k. Both types of capital are in perfectly elastic supply. In other words, at a certain price i_q (i_k for structures) a team is served as much equipment (structures) as it demands. We distinguish equipment capital from capital structures because their returns are strikingly different and, as a consequence of technological change, the former was accumulated much faster than the latter during the twentieth century (Krusell et al., 2000).

The workforce is partitioned into two groups, blue and white-collar workers. Blue-collar workers are denoted by u and they are heterogeneous in their type x, which is distributed according to a generalized CDF $\mathcal{U}(x)$ such that $\int_{\underline{x}}^{\bar{x}} \mathrm{d}\mathcal{U}(x) = \bar{U}$, where $\bar{U}$ is the total mass of blue-collar workers and $\underline{x}$ and $\bar{x}$ are, respectively, the worst and the best type of blue-collar workers. Similarly, white-collar workers are denoted by s and they are heterogeneous in their type y, which is distributed according to a generalized CDF $\mathcal{S}(y)$ such that $\int_{\underline{y}}^{\bar{y}} \mathrm{d}\mathcal{S}(y) = \bar{S}$, where $\bar{S}$ is the total mass of white-collar workers and $\underline{y}$ and $\bar{y}$ are, respectively, the worst and the best type of white-collar workers. We assume that in each team there is only one x and one y-type, though we do not make any restriction on their quantities (i.e. the u and the s are unrestricted).[7] To sum up, a team is a collection of labor and capital factor quantities $\{u, s, k, q\}$ and labor factor qualities $\{x, y\}$.

For a given team, the production structure closely mimics Krusell et al. (2000). Produc-

[5] Actually, the partition in Krusell et al. (2000) is into skilled and unskilled workers. So their production function distinguishes workers on the basis of their education. On the contrary, we split them on the basis of their occupation. In a hypothetical world in which unskilled workers are only blue collars and skilled workers are only white collars, the two partitions perfectly overlap. We argue that partitioning the workforce based on their occupation is more interesting when looking at firm-level data because it relates more closely to the actual tasks performed by the worker on their job.

[6] Extending the model to the case of competitive search is relatively straightforward (see Eeckhout and Kircher 2018). On the contrary, adding capital adjustment costs would require a dynamic framework as the firm problem cannot be regarded as stationary anymore.

[7] Actually, it would be sufficient to assume that in each team there is only one x-type as the fact that there is a single y-type follows as a result if the production function displays constant returns to scale in quantities (see Eeckhout and Kircher 2018, lemma 1).

tion can be seen as a two-layer hierarchy of "problems" that the team needs to solve in order to produce output (Garicano, 2000).[8] First, there is a "cognitive problem" that the team can only solve using a combination of white-collar work, s, and equipment capital q. Second, more labor is needed to implement the cognitive solution, and this requires the participation of blue-collar work, u. Finally, the production activity takes place into a physical place so the previous two steps are combined with capital structures, k. The second similarity with Krusell et al. (2000) is that we assume that factors substitute for each other at constant elasticity rates. Given this assumption and the production steps just outlined, we propose the following production function:

$$\text{CD}\left(k, \text{CES}\left(u, \text{CES}\left(q, s\right)\right)\right), \tag{1.1}$$

where CES and CD stands for Constant-Elasticity-of-Substitution and Cobb-Douglas, respectively. Notice that, as technically required in Eeckhout and Kircher (2018), both CD and CES functions are concave and expression (1.1) displays constant returns to scale in $\{u, s, k, q\}$.

Since structures do not play a central role for our argument, we set a unitary elasticity of substitution between k and the other factors. On the other hand, the elasticity rates of the two CES functions are crucial for determining with whom a blue or a white-collar individual teams up and what is going to be their marginal product. Under this functional specification we say there is "white-collar-biased technological change" when equipment substitutes for blue-collar workers at an higher rate than for white-collar workers.[9]

We depart from Krusell et al. (2000) in relaxing the assumption that all teams in the economy substitute at the same rate labor for capital. In particular, we design the elasticity rates so that they display two properties. First, the higher the worker type (either blue or white-collar), the less is substitutable with equipment. Second, the higher the type of the white-collar worker in the team, the more a blue-collar worker is substitutable with equipment. If one interprets a type as a measure for talent, then property one says that technological change is "talent-bias", in the sense that the higher a worker type, the more they are complement with equipment. Property two, says that the higher the type of a white collar worker, the more facilitated is the adoption of equipment in substitution for blue-collar labor. To formalize, consistent with these two properties, we assume white-collar workers and blue-collar workers substitute with equipment with respective elasticities of

$$\varepsilon_{q,s} = \frac{1}{1-\rho+y} \quad \text{and} \quad \varepsilon_{q,u} = \frac{1}{1-\sigma+x-y}. \tag{1.2}$$

The parameters ρ and σ are common across all teams and denote the baseline substitution rate when worker types are normalized to the value of 0, allowing for a direct comparison with Krusell et al. (2000).[10] On the other hand, equipment is more complement with higher x and y types as $\varepsilon_{q,s}$ decreases in y and $\varepsilon_{q,u}$ decreases in x. Finally, higher values in y tend

[8]Alternatively, one could think of these problems as a series of tasks in the spirit of Acemoglu and Autor (2011).

[9]This is exactly what Krusell et al. (2000) document in their study on the aggregate US economy, there is skill-biased technological change (see also footnote 5 above). Sharing the same elasticity for white-collar workers and equipment vs. blue-collar workers seems sensitive to assume since both equipment and white-collar workers contribute to solve the team cognitive problems.

[10]As a benchmark, Krusell et al. (2000) estimate elasticities of 1.67 and .67, respectively, for blue and white-collar jobs for the aggregate US economy (see also footnote 5 above).

to make equipment more substitute with blue-collar work.

We propose the following specification for the nested production function described in (1.1):

$$F(x,y,u,s,k,q) = A_k k^{\alpha}\left[\psi\tilde{u}(u,x,y)^{\tilde{\sigma}(x,y)} + (1-\psi)\left(\lambda\tilde{s}(s,x,y)^{\tilde{\rho}(y)} + (1-\lambda)(A_q q)^{\tilde{\rho}(y)}\right)^{\frac{\tilde{\sigma}(x,y)}{\tilde{\rho}(y)}}\right]^{\frac{(1-\alpha)}{\tilde{\sigma}(x,y)}}, \quad (1.3)$$

where

$$\begin{aligned}
\tilde{\rho}(x,y) =& \rho - y\\
\tilde{\sigma}(x,y) =& \sigma - x + y\\
\tilde{u}(u,x,y) =& u(\delta_u\phi_u(x)^{\gamma} + (1-\delta_u)\phi_s(y))^{\frac{1}{\gamma}}\\
\tilde{s}(s,x,y) =& s(\delta_w\phi_u(x)^{\gamma} + (1-\delta_w)\phi_s(y))^{\frac{1}{\gamma}}
\end{aligned}$$

and $\phi_u(x)$ and $\phi_s(y)$ are positive and increasing in their arguments. Function F denotes a team's total output, A_k and A_q are capital productivity measures, α, ψ, and λ contribute to determine the income shares of the four factors $\{u,s,k,q\}$. The type of workers $\{x,y\}$ play two roles in (1.3). First, worker types affect elasticities as described above. Second, worker types affect the efficiency units value of labor quantities u and s, through functions $\tilde{u}$ and $\tilde{s}$. Notice these two functions are CES aggregator in $\phi_u(x)$ and $\phi_s(y)$, with elasticity parameter γ. Intuitively, when $\gamma < 0$, the worker types $\{x,y\}$ are complement and otherwise are substitute. This interaction between x and y is an important force in guaranteeing the equilibrium allocation is PAM or NAM. [11] Finally notice that the expression in (1.3) as constant returns to scale in $\{u,s,k,q\}$.

1.2.2 Equilibrium

To define a competitive equilibrium for this economy we need to specify prices and allocations. Let us denote with $\{\omega^b(x), \tilde{\omega}^w(y)\}$ the wages of blue-collar and white-collar workers whose type is $\{x,y\}$ respectively. To spell out allocations, we need to keep track of which workers are employed in the same team. To do this, let us define with $\mu(x)$ the *type* y of white-collar workers who are matched with blue-collar workers of type x. For a given function $\mu(x)$, we can express the wage of a white-collar worker $y = \mu(x)$ as $\omega^w(x) = \tilde{\omega}^w(y)$. Let us also define $\theta(x)$ as the *quantity* of x-type workers that are matched to one unit of white-collar worker of type $y = \mu(x)$. Notice since (1.3) displays constant returns in $\{u,s,k,q\}$ we can normalize s to 1 and define the intensity-unit of production function $f(x,y,\theta,k,q) = \frac{1}{s}F(x,y,\frac{u}{s},1,k,q)$ for $\theta = \frac{u}{s}$.[12] Finally denote with $q(x)$ and $k(x)$ the amount of capital equipment and structures used by a team that employs x-type workers. Using this notation, a team $\{x,y,\theta,k,q\}$ is formed whenever the following expression is

[11] Despite we do not have an analytical proof, numerically whenever we set $\tilde{u} = u$ neither the PAM or NAM condition are fulfilled.

[12] With a little abuse of notation, we keep writing k and q although they should also be intended as normalized by units of white-collar work, that is k/s and q/s.

maximized in equilibrium:

$$\max_{\{x,y,\theta,k,q\}} \quad f(x,y,\theta,k,q) - \left[\omega^b(x)\theta + \omega^w(y) + i_k k + i_q q\right], \tag{1.4}$$

and maximizing (1.4) implies that no x or y-type workers find it optimal to leave their own team to join a different production team.

Definition 1.1. *Given i_q and i_k, a competitive equilibrium is a tuple $\{\omega^b(x), \omega^w(x), \mu(x), \theta(x), k(x), q(x)\}$ such that:*

1. *$\{x, \mu(x), \theta(x), k(x), q(x)\}$ is a solution to (1.4) for every $x \in [\underline{x}, \bar{x}]$, given $\omega^b(x)$ and $\omega^w(x)$;*
2. *labor demand for x and y-types does not exceed the endowment $\mathcal{U}(x)$ and $\mathcal{S}(y)$ for any sub-interval of $[\underline{x}, \bar{x}]$ and $[\underline{y}, \bar{y}]$;*

Let us start characterizing the equilibrium with noticing that a constant return technology in factor quantities, implies that the value of the team maximization problem in (1.4) is zero in equilibrium. To make any further progress, we need for the two sides of the labor market to match assortatively, either positively (PAM, $\mu'(x) > 0$) or negatively (NAM, $\mu'(x) < 0$). Define as $\hat{F}(\cdot)$ an auxiliary production function that already encompasses the optimal capital decisions.[13] Then Eeckhout and Kircher (2018) show that a necessary and sufficient condition to have PAM in equilibrium is:

$$\hat{F}_{xy}\hat{F}_{us} > \hat{F}_{xs}\hat{F}_{yu}, \tag{1.5}$$

which must hold for the whole domain of $\{x, y, u, s\}$. Vice versa, there is NAM when the sign of (1.5) is reverted over all the domain. In appendix A.1 and Eeckhout et al. (2021) are reported the conditions on the original production function F that guarantee condition 1.5 is fulfilled. Equilibrium wages and allocations are readily obtainable under PAM or NAM according to Proposition 2 in Eeckhout and Kircher (2018). In particular, notice that the economy is competitive and so wages equal worker marginal products for their own team. A type-x blue-collar in equilibrium gets $\omega^b(x) = f_u(x, \mu(x), \theta(x), k(x), q(x))$ and a type-y white-collar matched to $x' = \mu^{-1}(y)$ receives $\omega^w(x') = f_s(x', \mu(x'), \theta(x'), k(x'), q(x'))$. The schedule form $\mu(x)$ and $\theta(x)$ are derived after solving the following partial differential equations:

$$\text{PAM:} \qquad \mu'(x) = \frac{\mathcal{H}(x)}{\theta(x)}, \qquad \theta'(x) = \frac{\mathcal{H}(x)\hat{F}_{yu} - \hat{F}_{xs}}{\hat{F}_{us}}$$

$$\text{NAM:} \qquad \mu'(x) = -\frac{\mathcal{H}(x)}{\theta(x)}, \qquad \theta'(x) = -\frac{\mathcal{H}(x)\hat{F}_{yu} + \hat{F}_{xs}}{\hat{F}_{us}}$$

where $\mathcal{H}(x) = \frac{\mathrm{d}\mathcal{U}(x)}{d\mathcal{S}(\mu(x))}$.

[13] Formally, one defines $\hat{F}(\cdot)$ as:

$$\hat{F}(x,y,u,s) = \max_{\{k,q\}} F(x,y,u,s,k,q) - i_k k - i_q q,$$

where optimal k and q are functions of the remaining input values, $\{x, y, u, s\}$.

Notice that despite the equilibrium $\{\mu, \theta\}$ are unique, the values for $u(x)$ and $s(x)$ are not. This is because there are constant returns in factor quantities. Consequently, this implies that we cannot pin down the exact number of teams for a given x, but only their total mass. For instance, suppose that x and y match in equilibrium and so there is a measure $\mathrm{d}\mathcal{W}(\mu(x))$ of teams employing x and y. It remains undetermined in the model whether this mass $\mathrm{d}\mathcal{W}(\mu(x))$ just comprises a single atomistic team or a potentially infinite a multitude of them.

In spite of this caveat, the wage dispersion decomposition outlined in the next section applies anyhow. This is an important result for our argument as it means that using our notion of "teams" is enough for explaining the wage dispersion variation within and between firms, despite the absence of a well-defined notion of a firm in the model. Intuitively, this result holds true because two identical firms (same factor quantities and same wages) can be regarded as a single firm twice as large. In either case, total output is the same because of constant returns and so are individual wages and total wage variance. In the case of two separate firms, the across firm component of a wage variance decomposition is zero, as the two firms have identical wage schedules and worker composition. In the case of a single firm, the across firm wage variance is zero by definition.

1.2.3 Wage dispersion changes decomposition

As each team only comprises a certain x and y-type and there is a continuum of different types in the economy, then there exist a continuum of different teams. For each team $\{x, \mu(x), \theta(x), k(x), q(x)\}$ we can compute the average wage, $\omega(x)$, and its variance, σ^2:

$$\omega(x) = \frac{\theta(x)\omega^b(x) + \omega^w(x)}{\theta(x) + 1} \tag{1.6}$$

$$\sigma^2(x) = \left[\omega^w(x) - \omega^b(x)\right]^2 \frac{\theta(x)}{\theta(x) + 1}. \tag{1.7}$$

Using (1.6) and (1.7), it is possible to derive total wage dispersion in the economy. Since for every x there is a density $\mathrm{d}S\mathcal{W}(\mu(x))$ of teams, then average wage and total variance in the economy are:

$$\omega = \int_{\underline{x}}^{\bar{x}} \omega(x)\mathrm{d}S\mathcal{W}(\mu(x)) \tag{1.8}$$

$$\sigma^2 = \int_{\underline{x}}^{\bar{x}} \left[\omega^w(x) - \omega\right]^2 \mathrm{d}S\mathcal{W}(\mu(x)) + \int_{\underline{x}}^{\bar{x}} \left[\omega^b(x) - \omega\right]^2 \mathrm{d}U\mathcal{B}(x), \tag{1.9}$$

rearranging (1.9) and exploiting the fact that under PAM we have $\int_{\underline{x}}^{x'} \theta(x)\mathrm{d}\mathcal{W}(\mu(x)) = \int_{\underline{x}}^{x'} \mathrm{d}\mathcal{B}(x)$,[14] one can decompose total wage dispersion into variation occurring *within* teams, W, and *between* teams, B. We have:

$$\sigma^2 = W + B = \int_{\underline{x}}^{\bar{x}} \sigma^2(x)\mathrm{d}S\mathcal{W}(\mu(x)) + \int_{\underline{x}}^{\bar{x}} \left[\omega(x) - \omega\right]^2 (\theta(x) + 1)\mathrm{d}S\mathcal{W}(\mu(x)). \tag{1.10}$$

[14] Under NAM we have $\int_{\mu(x)}^{\mu(x')} \theta(x)\mathrm{d}S\mathcal{W}(x) = \int_{x'}^{x} \mathrm{d}\mathcal{B}(x)$ and decomposition in (1.10) goes through unchanged.

Notice this is the model counterpart of empirical measures of wage inequality decomposition across firms, as for instance computed in Song et al. (2019) and Briskar et al. (2022).

Changing the primitives of the model is going to change total wage dispersion for the economy. As several forces contribute to that change, the model can help us in identifying them. First, changes in the primitives of the model imply a *direct effect* in a worker marginal product and so on their wage. However, when primitives change, also team composition $\{x, \mu(x), \theta(x), k(x), q(x)\}$ endogenously responds in equilibrium. Such an equilibrium response entails an *indirect effect* on wages. This is the case because a worker marginal product does not only depend on their own type but also on *i)* the level of capital per worker in the team (*capital deepening complementarities*) and *ii)* their colleague type μ and relative proportion θ (*teammates complementarities*). Furthermore, for a given level of total wage dispersion, reallocating workers across teams is also going to affect the share of wage dispersion within and between firm . This is the *compositional effect*.

As an illustration of these three effects, suppose to compare economy A and B which are identical but for the return of equipment where $i_{q,B} < i_{q,A}$. In this example, the direct effect on wages is null because i_q does not directly affect the marginal product of labor. To look into the indirect effect on wages, let us start by fixing team composition as in economy A and denote these teams as $\{x, \mu_A(x), \theta_A(x), k_A(x), q_A(x)\}$. Any of these team can rent equipment at a cheaper rate, so they find optimal to use more in equipment in B than in A. In addition, as equipment and structures are complementary in the model, they also find optimal to use more structures. With a little abuse of notation, we can represent these production units as $\{x, \mu_A(x), \theta_A(x), k_B(x), q_B(x)\}$. Since $\{k_B(x), q_B(x)\} > \{k_A(x), q_A(x)\}$, then wages will be higher in economy B and we refer to the increased wage dispersion as the indirect effect stemming from capital deepening complementarities. We now turn to changes in labor composition.

Suppose for simplicity the allocation in both economies is PAM with $\mu_A''(x) < 0$ and $\theta_A'(x) = 0$ for every x. In words, all teams in economy A have the exact same proportion of white and blue collars, but top white-collars are disproportionately matched with better blue-collars. As the elasticity of substitution for blue collars is $\frac{1}{1-\sigma-x+\mu(x)}$, it will be the case that lower x blue-collars are substituted more intensively with equipment. In equilibrium for the economy B, this translates in $\theta_B(x)$ being downward sloping and $\mu_B(x)$ being more concave. In words, top white-collar workers in economy B work in production units with more equipment and relatively less blue-collars than in economy A. On the other hand, bottom white-collars are in production units more blue-collar intensive in economy B than in A. The resulting changes in wage stemming from variations in $\{\mu, \theta\}$ are referred to as the indirect effect of teammates complementarities.

To conclude this example, the compositional effect can be studied by keeping fix wages and, thus, total wage variance. For instance consider wages of economy A and compare the within-between firm wage dispersion decompositions that result from allocating workers according to $\{\mu_A, \theta_A\}$ and $\{\mu_B, \theta_B\}$.

More generally, the variation in wage dispersion from economy A to economy B is $\sigma_B^2 - \sigma_A^2$ which can be separated into changes in dispersion within and between teams:

$$\sigma_B^2 - \sigma_A^2 = (W_B - W_A) + (B_B - B_A), \tag{1.11}$$

both the within and between team changes can be decomposed into three types of effects.

Change in total wage dispersion (eq. A.12)					
Within team change (eq. A.13)			Between team change (eq. A.15)		
Composition effect (eq. A.14, line 1-2)	Indirect effect (eq. A.14, line 3-4)		Composition effect (eq. A.16, line 1-2)	Indirect effect (eq. A.16, line 3-4)	
	Capital deepening complementarities (eq. A.17 and A.18, line 1-2)	Teammates complementarities (eq. A.17 and A.18, line 3-4)		Capital deepening complementarities (eq. A.19 and A.20, line 1-2)	Teammates complementarities (eq. A.19 and A.20, line 3-4)

Table 1.1: Decomposition of total wage dispersion into changes stemming from worker-capital complementarities and worker-teammates complementarities (indirect effects) and relocation of workers across firms (compositional effect). This table does not account for changes in wage dispersion due to changes in the parameters of the production function or in the generalized CDFs $\mathcal{U}$ and $\mathcal{S}$. The corresponding equations can be found in appendix A.2.

First, direct changes in wages due to changing marginal products. Second, changes in wages due to changing complementarities either related to capital deepening or teammates. Third, holding wages fixed, there is a zero-sum reshuffle of within and between team variance components. In appendix A.2 we provide an analytical decomposition of the indirect and compositional effects.[15] In table 1.1, we summarize how to decompose total wage dispersion in the absence of direct effects, i.e. when only compositional, capital deepening complementarity and teammates complementarity effects are in place. Arguably, the conceptual novelty of our work lies in developing a theory to account for the *compositional* and *teammates-complementarity* effects as the literature has traditionally focuses on the direct and the capital-complementarity effects.

1.3 Quantitative Exercise

1.3.1 Data

We use Italian employer-employee data (VWH) matched with companies balance-sheet data (AMADEUS); and restrict our attention to compare year 1998 against 2001. The advantage of focusing on a short time-span is that it is more realistic to assume that the underlying production technology has not gone through significant changes. At the same time, Veneto underwent a fast capital accumulation during those years (Card et al., 2014) and this allows us to highlight more precisely the impact of capital usage on firm wage inequality.[16] Finally, using as a baseline 1998 allows us to measure wages more precisely as part-time vs. full-time information only started to be recorded in that year.

The VWH is a matched employer-employee database obtained from administrative records of the Italian social security system (see Tattara and Valentini 2007 and Bartolucci et al. 2018). It contains information on private sector employees in the Veneto region of Italy[17]

[15] There is still no analytical framework to decompose the direct changes, i.e. those stemming from variations in the production function parameters or in the generalized CDFs $\mathcal{U}$ and $\mathcal{S}$. In appendix A.2, we discuss how this framework can be developed.

[16] A limitation from our model is considering the economy as static. In this spirit, we acknowledge that capital and labor market decisions taken in 1998 were taken in the forward-looking perspective and as such 2001 is unlikely to be an "independent" observation. At the same time, it must be noted that Veneto's labor market is one of the more dynamic of Italy (Contini and Trivellato, 2005) so at least for the labor factor this seems to be less of a concern .

[17] Specifically, for the provinces of Vicenza and Treviso.

Descriptive statistics for the employer-employee sample

	1998	2001
Number of firms	2,693	3,580
Number of workers	114,471	154,351
Number of equivalent workers[a]	100,857	135,483
% blue-collars	70.41	68.47
Firm size	37.452	37.845
	(84.927)	(109.429)
Blue-white collars ratio	4.023	3.972
	(4.752)	(4.480)
ln(Tangible assets per white-collar)	4.468	4.493
	(1.218)	(1.276)
ln(Intangible assets per white-collar)	1.676	1.920
	(1.432)	(1.511)
Average ln(daily wage)	4.685	4.704
	(0.203)	(0.177)
Wage std. per firm	0.239	0.250
	(0.114)	(0.111)
Total wage variance[b]	100.0	107.95
% within firms	45.12	49.57
% between firms	54.88	50.43

Table 1.2: Standard deviation in parenthesis. All monetary values are in 2003 euro. [a]: Total number of days worked divided by 360; all the statistics except for the first two lines are weighted for firm size as measured by equivalent employees. [b]: Total wage variance in 1998 is normalized to 100 and total wage variance in 2001 is normalized to the labor force number of 1998. Wage variance decomposition is reported after substituting individual wages with the average wage in the respective firm-occupation cell. See appendix A.3 for the raw decomposition.

from 1975 to 2001. From the available information, it is possible to compute daily wages for every employee in a given firm and for a certain year. However, it is only from 1998 that observations contain information about part-time vs. full-time allowing us to compare more precisely daily wages across workers. Each record is associated with a job qualification — manager, white collar or blue collar— which is used in section 1.3.2 to identify the different segments of the labor supply.

The AMADEUS database is administered by Bureau Van Dijk, it covers the period 1993 - 2007 and contains data of comparable financial and business information on Europe's largest (by total assets) 520 000 public and private companies. We use the reported amounts of tangible and intangible assets to map into the model capital structures and equipment, respectively. The definition of equipment capital in our model does not completely overlap with the accounting definition of intangible assets. Yet, we argue identification is preserved as intangible assets are a factor of production that usually complements with white-collar better than with blue-collar occupations.

For our analysis we restrict our attention to full-time employees with a blue or white-collar job qualification working at the same employer for at least a quarter. Among white-collar workers, we also include managers as they naturally fit into the implementation component of our production function. On the side of firms, we only consider firms employing both blue and white-collar workers. Since only a subset of companies in AMADEUS report their assets, we drop from the sample those firms without asset information for both 1998 or 2001. The resulting sample is not representative for the whole population of companies as it comprises firm on average larger and with a lower blue-white-collars ratio than for the

Calibration values

Parameter name	Description	Value for 1998	Value for 2001
$[\underline{x}, \bar{x}]$	Blue-collar type bounds	$[-0.01, 0.01]$	
distribution of $\mathcal{U}(x)$	Type distribution for blue-collars	Uniform	
$[\underline{y}, \bar{y}]$	White-collar type bounds	$[-0.10, 0.10]$	
distribution of $\mathcal{S}(x)$	Type distribution for white-collars	Uniform	
$\mathcal{U}(\bar{x})/\mathcal{S}(\bar{y})$	Blue-white ratio in the economy	4.0	
α	income share of structure	0.015	
ψ	income share of blue-collars	0.83	
λ	income share of white-collars	0.88	
σ	Avg CES for blue-equipment	0.40	
ρ	Avg CES for white-equipment	-0.50	
A_k	Structures productivity	130.0	
A_q	Equipment productivity	1.5	
$\phi_u(x)$	Blue-collar productivity	$10e^{4x}$	
$\phi_s(y)$	White-collar productivity	$10e^{0.8y}$	
γ	CES of blue-white productivity	-10	
δ_u	Blue share of blue-productivity	0.9	
δ_s	Blue share of white-productivity	0	
i_q	Rental rate of equipment	0.22	0.20
i_k	Rental rate of structures	0.098	

Table 1.3: Model parameter values for calibration to 1998 and 2001.

whole population (for more details on our sample selection see appendix A.3).

A set of descriptive statistics for the sample is reported in table 1.2. For the purpose of comparability across firms, we define an equivalent worker as an employee who has reported a full-year of working days. On average, firm size is about 37 and there are seven blue-collars every three white collars. The relative intensity of intangible over tangible assets grows from 6.1% in 1998 to 7.6%, arguably a significant shift. Average wages and wage dispersion are higher in 2001, however the dispersion in standard deviations is higher in 1998.[18] Overall, this reflects into an increase in total wage dispersion, driven by the within component.[19]

1.3.2 Calibration

We need to calibrate the production function, the rental price for the two types of capital and the type distribution for the blue and white-collars. The share parameters $\{\alpha, \psi, \lambda\}$ are set so to generate plausible factor shares. The average elasticity of substitution parameters $\{\sigma, \rho\}$ between capital and labor are set as in Krusell et al. (2000).To discipline the productivity-related parameters $\{A_k, A_q\}$, and functions$\{\phi_u(x), \phi_s(y)\}$ we aim to generate plausible daily returns for each factor of production. In particular, we refer to Card et al. (2014) for setting the rental rate of structures and equipment (see table 5 column 3 therein). Card et al. (2014)

[18]The yearly composite change in the standard deviation is 1.51% for this period which nearly doubles the long-term trend of 0.79% documented by Briskar et al. (2022) for 1985-2018.

[19]A usual finding in the literature of within between firm wage decompositions is that the between component is the main factor shifting inequality. However, it is also common that the role of the within component grows when looking at the top quantiles of the firm size distribution. Notice that our sample firm size average is 37 that is just below the top decile of the full distribution of firm size. Further, using nearly the universe of firms in the economy, notice that alsoBriskar et al. (2022) document an approximate decrease of 10% in the share of between firms wage inequality between 1985 and 2018.

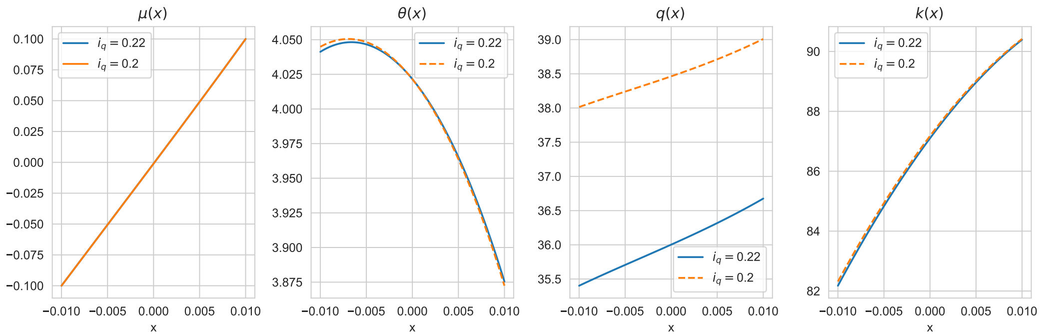


Figure 1.1: Equilibrium labor and capital allocation for 1998 ($i_q = .22$) and 2001 ($i_q = .20$). Parameter values are reported in table 1.3.

estimates an average user-cost for intangibles between 1997 and 2001 of 0.2. Accordingly, we set $i_{q,1998} = 0.22$ and $i_{q,2001} = 0.20$. The functions $\{\phi_u, \phi_s\}$ are exponential so that we can normalize the support for types around zero, and generate an increasing convex return for $\{x, y\}$ types. The parameter γ primarily reflects into the assortativity direction of the labor market, as $\gamma < 0$ implies $\tilde{u}, \tilde{s}$ supermodular in $\{x, y\}$. The choice of $\{\delta_u, \delta_s\}$ reflect a hierarchical vision where production spillovers flow in a one-way direction only (Garicano, 2000). In particular, we assume $\delta_s = 0$, that is the productivity of white-collar work is not directly affected by the x type of the workers they are matched to. On the other hand, $\delta_u < 1$ as white-collar type y affects the blue-collar workers productivity. Finally notice that the type distributions are latent variables and are hard to estimate from data.[20] We simply assume that types are uniformly distributed and that the blue-collars total mass is roughly 80% of total labor force.[21] We report in table 1.3 our model specification for the numerical simulation.

When solving the model in accordance to this specification, we obtain that the equilibrium displays PAM and θ is overall downward sloping but compressed around 4 (see figure 1.1), which is reflected in the almost 45 degree slope of μ. In other words, more productive white and blue-collars work together and there is an higher proportion of white-collars for higher values of x. Both equipment and structures are increasing in worker types with capital intensity for structure being about three times that for equipment.

Table 1.4 compares a series of moments from the data and the simulations. Moments include averages and correlations with appendix A.4 comparing additional correlations and percentiles of variables' distributions. Despite the average level of simulated variables aligns well with the model, the proposed calibration is not yet able to match other moments. In particular, the correlation among variables in the model is too high when compared to the data counterpart, and some come even with the wrong sign.[22] An important feature of the

[20] One could use an AKM approach Abowd et al. (1999) to obtain the worker types from a worker-firm fixed effects regression. Yet these fixed effects are hard to translate into theoretical types distributions (see Eeckhout and Kircher, 2011 among others).

[21] Since we are centering the type distributions around zero, it is reasonable on a first approach to use a symmetric distribution. Since we are being conservative on the spread of the support, a uniform distribution simplifies calculations without great cost. We plan to relax this assumption in the next update of the project.

[22] In general, given the high number of degrees of freedom, estimating the parameters poses a significant challenge. In particular, the absence of noise in the model and PAM means that wages of both types of workers are bound to be highly correlated, unless we accept non-monotonicity of wages in worker types. Under monotonic wages, it must be the case then that equipment and capital are non-monotonic. However,

Comparison of data and simulation moments

	1998		2001	
	data	model	data	model
avg(ω)	4.685	4.702	4.704	4.703
avg(σ)	0.239	0.232	0.250	0.235
avg(θ)	4.023	4.000	3.972	4.000
avg(k)	4.468	4.464	4.493	4.465
avg(q)	1.676	3.584	1.920	3.650
corr(w^s, w^u)	0.491	1.000	0.464	1.000
corr(θ, ω)	-0.324	-0.940	-0.403	-0.945
corr(θ, σ)	-0.246	-0.920	-0.335	-0.926
corr(θ, k)	0.282	-0.885	0.250	-0.892
corr(θ, q)	0.065	-0.939	-0.009^a	-0.951

Table 1.4: ω is the average log wage per firm; σ is the firm wage dispersion; θ is the firm ratio between blue- and white-collar workers (in equivalent units); k is the log of structures (tangible assets in the data) per equivalent white-collar worker in the firm; q is the log of equipment (intangible assets in the data) per equivalent white-collar worker in the firm. The data coefficients are for the cross-section years reported on the headings of the table. The model counterpart coefficients are calculated for the functions $\{\omega(x), \sigma(x), \theta(x), k(x), e(x)\}$ when solving for the model specifications reported in table 1.3. The data moments are computed weighting firms by their employment equivalent size. ^{a}: Not significant at the 0.1 level.

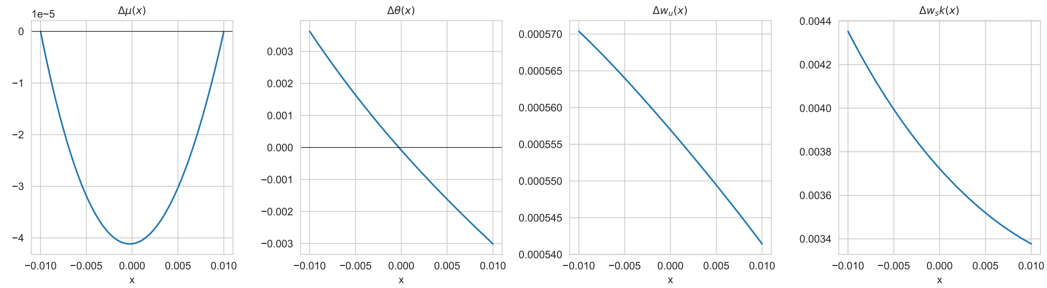


Figure 1.2: Equilibrium differences between 1998 and 2001 in labor allocation and log wage. The expression $\Delta\mu(x) = \mu_{2001}(x) - \mu_{1998}(x)$ represents the change in the type of y that worker x is matched to. Similarly it goes for $\theta(x)$. The last two panels represent the changes in log wages for blue-collar workers of type x and skilled workers of type $\mu_{year}(x)$. Parameter values are reported in table 1.3.

data that the model is able to replicate qualitatively is a *decreasing* θ. A promising direction to bring the values of simulated correlations closer to the data is choosing parameters combinations where θ and/or equipment is not monotonic — like the one we proposed, though quantitatively too little. Despite these shortcomings, we use this calibration to illustrate how the model reacts when the price of equipment is lowered – but recall that this is not yet a full-fledged quantitative exercise, so the observations that follow need to be taken with caution.

First of all, notice that the model simulation for 1998 and 2001 only differs in the value of i_q, as a consequence the allocation of labor and capital is very similar across the two specification except for equipment itself. Nevertheless, there is a change in allocation between the two simulated economies as shown in the first two panels of figure 1.2. When equipment

parameter combinations that result in non-monotonic choices for equipment result in poorer fit in levels. In the next update of the project, we aim to get an estimation that matches both levels and correlations.

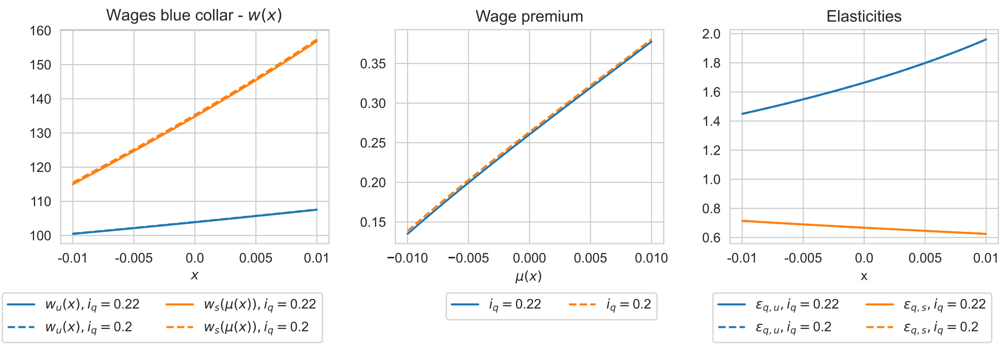


Figure 1.3: Equilibrium wages and elasticities of substitution for 1998 ($i_q = .22$) and 2001 ($i_q = .20$), notice that elasticities virtually overlap. Parameter values are reported in table 1.3.

is cheaper, every production unit becomes more capital intensive. Better teams use cheaper equipment to substitute for blue collar workers, as the change in θ shows. However, notice that because top white collar workers are more efficient at utilising equipment, they increase q *less* than teams with less efficient white collars. So equipment intensity increases across the board, resulting in higher substitution of blue-collars, but differences in capital intensity across teams decrease.

As a consequence, top blue-collars in 2001 are in lower demand in top white-collar teams than in 1998 and blue-collar ratio decreases for top x. On the other side of the distribution, where bottom white-collar are less complement with equipment, the demand for blue-collar worker rises – this is reflected in the changes in θ in the second panel of figure 1.2. The matching is also (slightly) affected: the minimum of $\Delta\mu(x)$ coincides with $\Delta\theta(x)$ crossing the zero line. Mechanically, left of the crossing point of $\Delta\theta$ blue-collars match with less white collars (θ goes up) and consequently μ_{2001} is less steep than μ_{1998}. After that point, things reverse: blue-collars get matched with more white collars (θ decreases) and consequently μ_{2001} is steeper than μ_{1998}. This is reflected in the changes of μ in the first panel. Effectively, equipment is replacing blue collar workers in quantity and quality within teams. Overall, the differences in equilibrium μ and θ for 1998 and 2001 can be summarized by looking at equilibrium elasticities in the last panel of figure 1.3. While we assumed that capital is "talent-biased", that is higher types of x and y *individually* are more complement to capital than lower types, it does not have to be the case that higher x and y types are also more complement to capital *in equilibrium* as captured in equation (1.2) of the model. Indeed, the equilibrium elasticity of substitution is decreasing in y while it increases in x. Intuitively, since the equilibrium is PAM, higher x types match with higher y types that make capital adoption easier and favor the substitution of blue-collar worker with more capital and this channel dominates the bias of capital for talent.

We now turn our attention to wages and factor shares that are reported in figure 1.3 and 1.4. Wages are increasing in worker types, reflecting the positive slope of the functions ϕ_u and ϕ_s. However, the wage profile for white-collar workers is steeper and this reflects in an increasing wage premium over the type distribution. This is also confirmed when looking at factor shares. The share paid out for capital equipment and structures is constant across worker types while blue-collar shares decreases for higher percentiles of x.

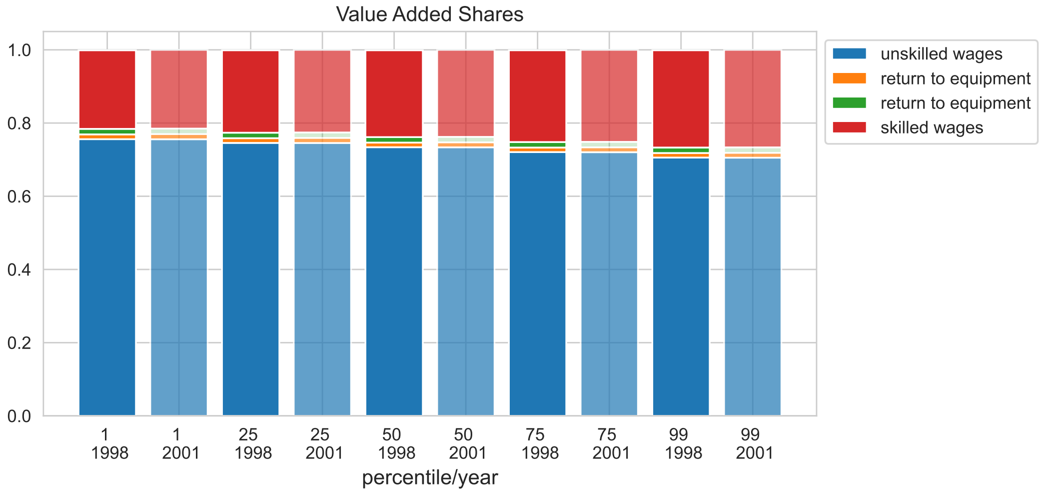


Figure 1.4: Factor shares for labor and capital in 1998 and 2001. Percentiles refer to the distribution of the blue-collar x types. Parameter values are reported in table 1.3.

1.3.3 Results

Total wage variance between 1998 and 2001 simulated model has increased by 2.04% (7.95% is the corresponding data moment). In table 1.5 we decomposed this increase in within and between firm changes. The within component is dominant as it accounts for 100.97% of the total change. The between component change is negative, coherently with decreasing rises in wages in the last two panels of figure 1.2.

The structure of the theory outlined in section 1.2.3 allows us to further decompose the changes in inequality into a compositional effect and two indirect effects, i.e. capital complementarity and teammates complementarity. Perhaps unsurprisingly, given the nature of the exercise was changing i_q only, we find that the capital complementarity channel is the dominant force behind the within firm component rise. This result is aligned with previous theories studying technological change that use occupation but no further heterogeneity in worker types (Krusell et al., 2000). If there was no heterogeneity in x and y in the economy, a representative firm would emerge and changes in inequality would occur within this firm. In this representative economy, if capital is biased in favor of white-collar workers, a decrease in the price of equipment would increase within firm wage inequality 1-to-1 with economy-wide wage inequality. However, such a reading misses where the change is coming from: better paying (higher white collar type) teams are being more efficient than lower paying teams at substituting blue collars for capital. So it is qualitatively not true that within firm dispersion only comes from the substitution of blue collar workers with equipment, but also from white collar workers making a more efficient use of equipment. The magnitude of the changes in allocations under our current specification are very small, but this is likely a by-product of the current parameter combination. Despite small, these changes in allocation represent an indication of a likely direction of change for when the estimation of the model is completed.

Finally, notice that the between component change in inequality is only mildly dominated by the capital complementarity channel. The compositional effect and the teammates

Decomposition of the simulated change in total wage variance from 1998 to 2001.

Component	Value(*100)	in %
Tot. change	0.129	2.04
– Within component	0.131	100.97
— capital complementarity	0.135	104.19
— teammates complementarity	-0.002	-1.66
— compositional effect	-0.002	-1.56
– Between component	-0.001	-0.97
— capital complementarity	-0.002	-1.85
— teammates complementarity	0.000	0.34
— compositional effect	0.001	0.54

Table 1.5: Simulated data. Notice that percentage values are expressed in terms of the total change of the first line. Values are rescaled by 100 to increase readibility.

complementarity channel counteract for nearly half the effect of capital complementarity on between firm changes in wage dispersion.

1.4 Conclusion

In this paper we propose a novel channel to account for when studying the effects of technological change on wages. On top of increased marginal return of labor due to complementarities with capital, we argue that changes in capital investment entail a relocation of the workforce. In a sufficiently rich environment, as the one with two-sided continuous heterogeneity we described, a worker marginal product depends also on their colleagues and, so, is affected by a relocation of the workforce. We call this channel the teammates complementarity effect.

We consider the consequences of a relocation of the workforce, jointly with changes in labor-capital complementarity, to shed new light on the debate about increasing wage inequality within vs. between firms. In doing so, we decompose changes in wage dispersion into changes in marginal product and compositional effects. We account for two possible sources of variation in a worker marginal product, capital complementarity and teammates complementarity. In addition, we establish that labor relocation involves a compositional effect, reshuffling part of the within firm wage dispersion into between firm dispersion or vice versa.

As future steps for this paper, we intend to further analyze the data and understand how to compare our findings with the rest of the empirical literature. We also need to calibrate the model more accurately as it looks totally off at the moment.

Future research directions, possibly as extensions of the paper or as autonomous projects, would for instance be the followings. Turn the model into a dynamic set-up where capital decisions are made on an explicit intertemporal horizon. Let the workers choose, and possibly transit along their career, between blue and white-collar jobs; as this is likely a relevant dimension to consider when studying the long-run effects of technological change. Microfound the production function, which is taken as given for now, as the result of worker interactions within teams; this way offering a link with the literature on hierarchical organizations (Garicano, 2000), on task-based production (Acemoglu and Autor, 2011) and the introduction of robots (Acemoglu and Restrepo, 2022).

Chapter 2

The Gender Pay Gap Between and Within Firms: Implications from Capital-Skill Complementarity and Sorting

Cristina Lafuente & Gabriele Macci

Abstract

In this project, we intend to quantify the impact of capital subsidies on the gender pay gap. The first step is documenting how the gender pay gap is distributed between and within firms over time in Italy. The second step proposes a theoretical model that explains how the distribution of the gender pay gap is linked to firms. The *levels* of the gender pay gap distribution are explained with a competitive sorting and matching model with taste-based gender discrimination. The *changes* in the distribution are explained by augmenting the model with capital investment. Holding gender discrimination fixed, a reduction in the rental price of capital reduces the gender pay gap by letting female workers sort into better matched jobs. Finally we discuss how to use a calibrated version of the model to compare counterfactual scenarios with different levels of capital investment subsidies.

2.1 Introduction

The gender pay gap has always attracted a lot of attention in the profession and, in the last few decades, among policy-makers. There has been a huge collective effort explaining and proposing policy measures to mitigate it (see Blau and Kahn, 2017; Rubery and Koukiadaki, 2016, among many others). In this project, we contribute to this large literature by focusing on firms. Our focus on the distribution of the gap between firms is motivated by the finding of Card et al. (2016), Sorkin (2017) and Casarico and Lattanzio (2022) —who use the same Italian administrative data as us—, that the larger gender pay gaps occur in better-paying jobs and high wage-premium firms. Further, since firms are a major driver of changes in inequality (see for instance Abowd et al., 1999; Song et al., 2019), policies affecting the firm wage distribution are likely to affect the gender pay gap as well.

In this project, we explore the impact of *capital deepening* on the gender pay gap. A long tradition (see Krusell et al. 2000) argues that capital deepening affects unequally workers across different occupations.[23] We argue that increases in capital deepening also affect the gender pay gap, through two channels. Firstly, if women are increasingly over-represented among new graduate cohorts, then more capital adoption will reduce the gender pay gap, as more women will tend to be in jobs that are more likely to be complements with capital (Bhalotra et al., 2022). Secondly, since the total surplus that can be generated is higher for more capital-intensive jobs, the cost of mismatch is increasing in capital investment. A higher cost of mismatch makes taste-based gender discrimination more expensive and, accordingly, reduces the gender pay gap by improving labor market sorting.

First, we use Italian administrative data to document what is the firm-level relationship between the gender pay gap and capital deepening between 1998 and 2018. The relationship between the two variables is generally negative and non-linear. Despite the *gender pay gap* has reduced over time, this has happened unequally across occupations. In particular, the gap has closed faster for white-collar workers than for blue-collar workers. At the same time, we document that virtually the whole increase in the *skill pay gap* within firms is attributable to female white-collar workers increasingly earning more than their female blue-collar colleagues. In light of these findings, we argue that capital-skill complementarity is associated to a decrease in the gender pay gap while magnifying wage dispersion across women.

Second, we propose a competitive matching model of the labor market to discipline our empirical findings. The use of an equilibrium matching model means that we treat gender inequality in a subtle way. Since in the model workers receive their marginal product as wages, a naive inspection of the data would conclude that there is no discrimination in the labour market, since all wage differentials are explained by productivity. However, matching models have the capacity to look at wages in a different dimension– that is, through allocative choices in an heterogeneous agent environment. If there is discrimination in the labor market, this would alter the equilibrium allocation of workers, resulting in skilled women forming teams with less skilled men than in an equilibrium without discrimination (as in Cuberes et al. 2021). As this misallocation results in less total output, it lowers the marginal product of women compared to the one that would prevail if they were matched with better male coworkers.

The rich heterogeneity of the model remains tractable as the matching environment is similar to Eeckhout and Kircher (2018) with two minor modifications. First, capital investment is added in the model following Eeckhout et al. (2021) and Krusell et al. (2000). Second, as a baseline, we introduce gender differences only through taste-based discrimination (Becker, 1957). However, we also discuss the implications on the equilibrium matching and gender gap of having gender-specific capital complementarities (Sánchez et al., 2020; Marjit and Oladi, 2022).

Finally, we discuss how to estimate our model using Italian employer-employee data and use it for assessing counterfactual industrial policies and their unintended consequence on

[23] In particular, lower skilled workers are easier to substitute with capital than higher skilled ones, increasing the college-premium.

the gender wage gap. For instance, subsidized capital depreciation policies incentivize firms to be more capital intensive and increase total surplus per worker. How this would affect the gender gap is partially analyzed in Card et al. (2016), by comparing workers that *remain* in the same firm over time. They find that the gender gap is unaffected by changes in potential surplus over time. However, their approach is limited in the scope as it rules out by design labor reallocation across firms. We argue our structural model is well-suited in revising this question as it accounts for both the direct effect of changes in surplus and the indirect effect stemming from workforce reallocation across-firms.

The rest of the paper proceeds as follows. Section 2.2 describes the data and presents how the gender pay gap, the skill gap and capital deepening evolved over time across firms. Section 2.3 describes the matching framework in a constructive way, first showing how to introduce discrimination between men and women, and then also adding occupations and capital. Section 2.4 briefly discusses how to develop a quantitative exercise to assess the indirect consequences of capital deepening on the geneder pay gap, and section 2.5 concludes.

2.2 Empirical analysis

In this section we describe how the gender and skill pay gap have evolved in Italy over two decades. Gender pay gap has generally reduced, especially for white-collar jobs while the skill gap has increased especially among women. Finally we link these changes to the increased in capital per worker (capital deepening) that has occured during the same time period.

We used Italian administrative data provided by INPS, the Italian Institute of Social Security. We used the employer-employee database to construct firm-level statistics on wage inequality and the CERVED data to link firm payroll information with the book value of their capital. Given the large amount of data available, we limited our analysis to sample of years in the last two decades, namely 1998, 2003, 2008, 2013, 2016 and, 2018.[24] Because the focus of the paper is to analyses secular trends in wage inequality and capital adoption, this sample gives us a good time frame to study these changes.

In the dataset, every employment spell, regardless of its duration, is stored as a separate record. Records are at yearly frequency, therefore multiple-year spells are recorded as a sequence of observations. For each observation we observe the total salary, the number of days under contract during the year, the occupation (trainee, blue-collar, white-collar and manager), and the part-time fraction in case of non full-time workers.[25] As a baseline, we restrict our analysis to blue and white-collar workers full-time employed at a single employer during the year. This restriction guarantees us the sample of men and women we used is as much as possible homogeneous and that wages are readily comparable. Given the nature of our research question, we focus our attention to firms reporting (i) at least one worker for each gender and occupation and (ii) positive fixed assets (either tangible or intangible) in their balance sheet statement.[26]

[24] The reason we include 2016 here is because 2013 is a year affected by the Great Recession, and in particular the Eurozone crisis.

[25] In the data there are two groups of white-collar workers, qualified and unqualified. We pool them together.

[26] These restrictions are related - firms reporting capital are more likely to be bigger, and bigger firms are more likely to have the 4 categories of workers on payroll. For this reason, a simpler restriction is to focus on firms with at least 20 full-time employees. This turns out to be a more practical approach than applying the two restrictions together, since capital reporting is quite noisy for the first two waves. For this reason

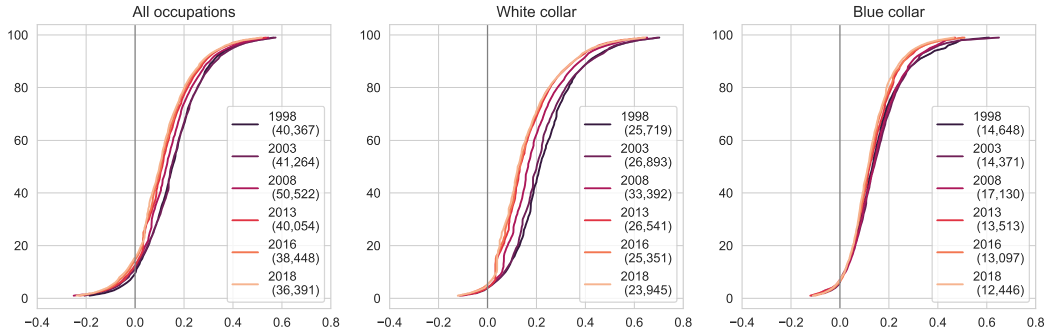


Figure 2.1: Gender gap is defined as the difference between the log average wage of men versus women within a firm. All samples are restricted to firms employing at least 20 full-time employees. On the vertical axis there are percentiles of the gap distribution. In parenthesis is the number of observations.

The wage gap is measured as the log difference between the average wage of men and women in a firm. In figure 2.1 we report the empirical CDFs for the gender wage gap in Italy between 1998 and 2018. Let us first focus on the aggregate gender gap. Notice the gender wage gap is generally positive, i.e. men earning more than women. This is true both at the mean, as most values lie on the right of 0 in the x-axis, and at the median, as the 50th percentile on the y-axis has a positive value on the x-axis. This is remarkable, but recall that we are not controlling for occupational composition. Part of the explanation could be related to seniority or occupational differences within firm – which we are trying to explain in this paper. These composition effect likely accounts for the few firms where the gap is negative as well. Yet, these differences must be large enough to tip most of the firms into the positive. Second, the gap is not uniform across firms: if it was, the CDF would look like a straight vertical line. Instead, it goes from slightly over zero at the 20th percentile to over 20% at the 80th, and reaches almost 60% at the upper tail. Third, the gap overall has reduced over time as the lines moved leftward. However, the reduction has not occurred at the same speed across firms. For later years, the line has become steeper in its central part, suggesting a similar average wage gap has prevailed across more firms. At the same time, though, the extremes of the line moved little over time indicating that firms with more extreme gaps have changed their wage policies the less.

Turning to the firm gender gap *by occupation*, we can observe that most of the decrease in the gap has occurred for white-collar jobs. The reduction of the gap for white-collar workers has been large and widespread, as the whole CDFs moved leftward starting from the bottom 10th percentile onward. On the other hand, the reduction in the gap for blue collar workers has mostly been driven by firms with very large gaps to start with. Only the top of the CDFs moved leftward for this group of workers. It is also interesting to notice that the overall gap is smaller (and more concentrated) than for white collars: if we focus on the 80th percentile, the gender pay gap is very close in all years to 20%. For white collars, it was more than 30%. This indicates that wage differentials are driven by skilled workers and firms that employ them.

We now turn to the skill gap, which is measured as the log difference between the average

---

the positive capital restriction is currently not enforced for the first part of our analysis. The share of firms employing at least 20 employees and reporting positive capital in their statement ranges between 58 and 63% in the years we considered.

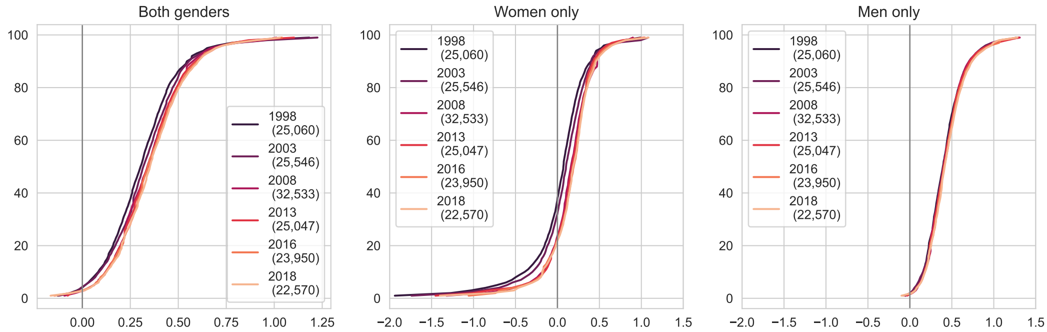


Figure 2.2: Gender gap is defined as the difference between the log average wage of white-collar versus blue-collar workers within a firm. All samples are restricted to firms employing at least 20 full-time employees. On the vertical axis there are percentiles of the gap distribution.

wage of white-collar and blue-collar workers in a firm. In figure 2.2 we report the empirical CDFs for the skill wage gap in Italy between 1998 and 2018. Let us first focus on the skill gap pooling gender together. Notice the skill wage gap is generally positive, i.e. in most firms white collar workers earn more than blue collar workers. This is true at the mean, and at the median and at very low levels. Second, there is substantial heterogeneity in the skill gap across firms, but overall is quite compressed with top 20th-percentile firm having a dispersion of only .25 points larger than bottom 20th-percentile firm. Third, the gap overall has uniformly increased over time as the lines translated rightward.

Turning to the firm skill gap *by gender*, we can observe that virtually all of the increase in the gap has occurred for women's jobs. The increase of the gap for women workers has been substantial particularly for central percentiles. It is also remarkable that as many as 40% of firms had a "negative" skill gap from women. Recall that we are not controlling here for tenure or other personal characteristics. Assume that blue-collar jobs were more likely being occupied for women in the past; if wages increase with tenure, we would expect a firm with some experiences blue collar females share firm with a few young white-collars. This is consistent with both (i) women being placed in low-skill occupations within firms and (ii) white collar women being discriminated against. These are the mechanisms that we will aim to capture in the model. On the other hand, the increase in the gap for men has been very small. Notice that at the end of the sample, in 2018, the skill gap is almost as concentrated for men as for women: comparing the bottom 20 (20th percentile) to the top 20 (80th percentile), the gap spans less than 0.5. This wage compression is a likely a reflection of the labour market institutions of Italy – in particular, wage bargaining. Precisely because these rules can constrain firms' wage setting decisions, they have an incentive to 'soft discriminate' and place capable women in worse jobs.

Finally, capital deepening is measured as log assets per employee. In figure 2.3 we report the empirical CDFs for capital deepening, also distinguishing by tangible and intangible asset class. Notice that capital deepening has increased over time as lines move rightward. Total and tangible capital deepening has increased particularly for more capital intensive firms. On the other hand, intangible capital deepening has increased across the whole firm distribution. We attribute the difference in the evolution of capital deepening for tangible and intangible assets to the fast expansion of more knowledge-based tasks in the production base Acemoglu and Autor (2011).

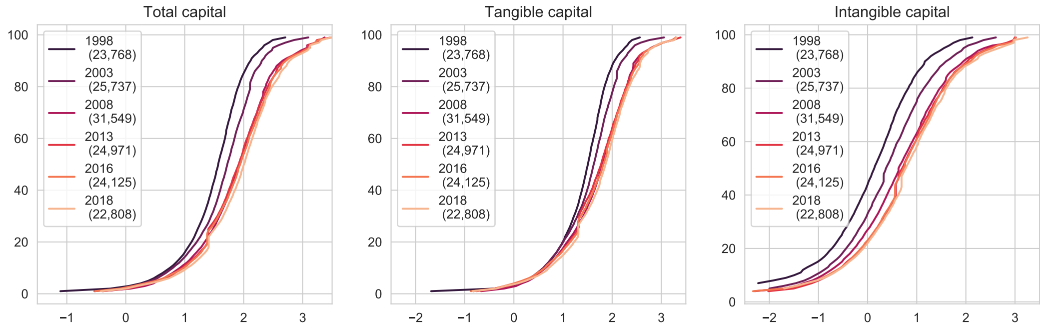


Figure 2.3: Capital deepening is defined as log assets per employee. All samples are restricted to firms employing at least 20 full-time employees. On the vertical axis there are percentiles of the capital deepening distribution.

Taking stocks of the overall evidence presented in this section, we advance the thesis that part of the reduction in the gender pay gap is attributable to skill-biased technological change. The increased availability of capital in firms has raised the value of white-collar workers since they are more complementary to capital than blue-collars. The increased demand for white-collar workers could not just be matched by the supply of white-collar men, rising the demand for white-collar women as well. This argument provides a combined explanation for the average changes in the CDFs of the wage gap, skill gap and capital deepening variables. However this argument alone falls short of explaining the differences in the evolution of the CDFs at various points of the curve. For instance, the gender gap has reduced in the average firm, but less so for firms starting with very high gaps. In the next section, we propose to recur to a sorting model of the labor market in order to explain this higher-order changes in the CDFs.

2.3 Theoretical framework

We propose a competitive matching model of the labor market to discipline our empirical findings. Our model is introduced in a constructive way. First, we describe a baseline version of the model with two genders but without occupations. This version resembles Eeckhout and Kircher (2018) as it only features heterogeneity in the *type* of workers and firms that match in equilibrium to form production units. We use this baseline model to discuss what is the effect of discrimination on the equilibrium matching and wages. Secondly, we distinguish blue and white-collar occupations. Introducing occupations permits to jointly determine equilibrium gender and skill gaps. Finally, we build on Eeckhout et al. (2021) and introduce capital in the model. The way capital complements differently with occupation alters firms demand for types and quantities in equilibrium. Higher complementarity of capital with white collar workers inflates their wage. In equilibrium, increased capital deepening makes discrimination more costly and consequently raises women's white collar wages more than men's. As wages are based on the marginal product of workers in their production unit, the effect of introducing capital in the model is not identical across the type distribution as it also depends on the equilibrium matching. This last feature lets the gender and wage gap react with different intensity across worker types and has the potential to explain the higher-order changes in the CDFs documented in the previous section.

2.3.1 A competitive matching economy with discrimination

The economy is static and populated by firms, female and male workers. Firms are indexed by a continuous variable y distributed according to the generalized CDF $\mathcal{F}(y)$ for a compact set Y.[27] Men and women are indexed in x_M and x_F, respectively distributed in $\mathcal{M}$ and $\mathcal{W}$ for compact sets M and W.

Firm of type y matches with a group of x_F workers and group of x_M workers, deciding how many to hire of each kind, l_F and l_M. They produce, sell and split the profits. The production function has five arguments:

$$f(y, x_F, l_F, x_M, l_M) \tag{2.1}$$

where f displays decreasing returns to scale both in l_F and l_M.[28] We assume that man and women are equally productive, that is f is such that $f(y, x_F, l_F, x_M, l_M) = f(y, x_M, l_M, x_F, l_F)$. This is a five-input model and firm y solves:

$$\max_{\{x_M, x_F, l_F, l_M\}} f(x_F, x_M, y, l_F, l_M) - w_F(x_F)l_F - w_M(x_M)l_M. \tag{2.2}$$

$w_F(x_F)$ and $w_M(x_M)$ are wages taken as given by the firm. Wages are determined in equilibrium so that labor markets clear:

$$\int_{\mu_M(x_M)}^{\bar{y}} l_M(s)d\mathcal{F}(s) = \int_{x_M}^{\bar{x}_M} d\mathcal{M}(s), \tag{2.3}$$

$$\int_{\mu_F(x_F)}^{\bar{y}} l_F(s)d\mathcal{F}(s) = \int_{x_F}^{\bar{x}_F} d\mathcal{W}(s), \tag{2.4}$$

where $\{\mu_F, \mu_M\}$ denote the optimal hiring policy in worker quality for a given firm of type y, and the optimal policy in numbers of workers is given by $\{l_F, l_M\}$.

We introduce taste-based discrimination (Becker, 1957) in this economy by modifying the firm problem (2.2) and including a wage wedge when hiring women. We have:

$$\max_{x_F, l_F, x_M, l_M} \quad f(y, x_F, l_F, x_M, l_M) - (\tau_a + \tau_m w_F(x_F))\, l_F - w_M(x_M)l_M. \tag{2.5}$$

Note a wage wedge for women, given by the terms $\tau_a > 0$ and $\tau_b > 1$, is included in equation 2.5. This is a very flexible specification that encompasses multiplicative and additive types of wedges. The additive term τ_a is a cost wedge on hiring a certain quantity of women regardless of their type, e.g. for each employed women forgoes home-production. The multiplicative term τ_m reflects the idea that highly skilled jobs are more demanding in terms of time, meaning more of a problem for women to participate.

We include both $\tau_a \geq 0$ and $\tau_m \geq 1$ in the model because they have strikingly different implications for wages in equilibrium. While τ_m is a larger cost for highly paid women, τ_a is

[27] A generalized CDFs in the sense that $\mathcal{F}, \mathcal{W}, \mathcal{M}$ fulfill all the properties of cumulative distribution function but do not need to integrate up to 1.

[28] In Eeckhout and Kircher (2018)'s terminology, we are directly presenting the production function in intensity units of r, a latent variable denoting how much quantity of y there is in a given firm. When this extra argument is included in the production function, there are jointly constant returns to scale in (r, l_F, l_M). However, because of constant returns to scale, factor r can be normalized 1 without any loss of generality.

flat across women's wage distribution and so it disproportionally affects the cost structure of firms employing low wage women. We can formalize this intuition by making a few predictions for the wage gap by solving the first order conditions of problem (2.5) in x_F and l_F:

$$F_{x_F} - \tau_m \cdot w'_F(x_F) l_F = 0$$
$$F_{l_F} - \tau_a - \tau_m \cdot w_F(x_F) = 0.$$

Dividing the second by the counterpart for men we get

$$\frac{w_M(x_M)}{w_F(x_F)} = \tau_m \cdot \frac{F_{l_M}(x_M)}{F_{l_F}(x_F) - \tau_a}.$$

If $\tau_a > 0$, this means that compared to the case where there is no wedge ($\tau_a = 0$ and $\tau_m = 1$) the ratio between the wage of men and women is larger: men are better paid in the firm. If the wedge is multiplicative ($\tau_m > 1$) then the wedge is directly proportional to wage inequality. Considering the first order condition with respect to x_F we get

$$\frac{w'_M(x_M)}{w'_F(x_F)} = \tau_m \cdot \frac{l_F}{l_M} \frac{F_{x_M}(x_M)}{F_{x_F}(x_F)},$$

which means that if $\tau_m > 1$ men's wage function is steeper than women's — in other words, the gap is larger at the top than a the bottom of the wage distribution. The larger τ_m, the larger is the difference between the wage slopes. An additive constant would not be in this equation, reflecting that if $\tau_a > 0$ and $\tau_m = 1$, the wage function just shifts downward for women — the wage gap is the same at the top and the bottom in absolute terms, yet decreasing in relative terms.

Equilibrium

The equilibrium of the model is characterized by a system of 4 differential equations (potentially 6 if we count the wage) and 4 boundary conditions: 2 for $\mu'_F(\cdot)$ and $\mu'_M(\cdot)$ and 2 for $l'_F(\cdot)$ and $l'_M(\cdot)$. The boundary conditions are given by the market clearing constraints in equation (2.3) and (2.4). The required differential equations are in Eeckhout and Kircher (2018) proposition 3.

In spite of the considerations just made on the role of τ_a and τ_m in affecting wages, the role of taste-based discrimination is actually quite limited in distorting equilibrium allocation of workers across firms.

Proposition 2.1. *If in equilibrium all male and female workers are employed, then τ_a and τ_m do not affect the equilibrium μ_i and l_i for $i \in \{F, M\}$.*

The proof of this proposition is in appendix B.1. This proposition states that taste-based discrimination does not affect the matching of workers as long as there are no unemployed left. Intuitively, if all workers are employed the only role for discrimination is shifting downward and make flatter the wage profile of women, however this does not change their outside option to form certain production units because there is no possible re-shuffle of workers across firms able to generate more surplus. However, this proposition also provides a limit for this logic. If τ_a is particularly high, lower types of x_F might remain unemployed as their contribution to firm surplus would be negative for any (positive) wage level. In this case, the effects of changes in discrimination propagate on the allocation of the rest of the employed workforce.

2.3.2 Introducing occupations in the model

The theory to extend the model for two types of occupations is relatively straightforward given the generality of proposition 3 in Eeckhout and Kircher (2018). In particular, we assume workers are indexed in x_G^O with $G \in \{M, F\}$ denoting gender and $O \in \{U, S\}$ denoting if they are blue or white collars. Generalized CDFs adjust accordingly and the firm production problem becomes:

$$\begin{aligned}\max_{x_F^U, l_F^U, x_M^U, l_M^U, x_F^S, l_F^S, x_M^S, l_M^S} \quad & f(y, x_F^U, l_F^U, x_M^U, l_M^U, x_F^S, l_F^S, x_M^S, l_M^S) \\ & - w_M^U(x_F^U) l_M^U - \left(\tau_a^U + \tau_m^U w_F^S(x_F)\right) l_F^U - w_M^S(x_M) l_M^S - \left(\tau_a^S + \tau_m^S w_F^S(x_F)\right) l_F^S.\end{aligned} \tag{2.6}$$

In this case, the production function is symmetric in gender at the occupation level. In this sense, notice that it is only possible to rank workers of different genders within a certain occupation but there is no natural ordering across occupation. For instance, the comparison between x_F^S and x_F^U bears no intrinsic economic meaning because those workers are involved in distinct occupations and they enter the production function differently.

In this richer economic environment, it is now possible to calculate both the gender and skill pay gap and trace them back to the discrimination wedges of the τ terms. Once again, as for the previous section, discrimination does not affect the equilibrium matching as long as all male and female worker in both occupations are employed.

2.3.3 Introducing capital in the model

Following Eeckhout et al. (2021), the introduction of one or two types of capital in the model comes at no additional cost, provided the production function has constant returns to scale in labor and capital quantities. The evidence from the empirical analysis suggests that capital deepening has followed a different path for tangible and intangible assets. Further, as capital equipment displays different elasticities of substitution by occupation (Krusell et al., 2000), we propose to include multiple varieties of capital. Let us denote with $\vec{k}$ the quantity of capital in each of its varieties, the firm production problem becomes:

$$\begin{aligned}\max_{x_F^U, l_F^U, x_M^U, l_M^U, x_F^S, l_F^S, x_M^S, l_M^S, \vec{k}} \quad & f(y, x_F^U, l_F^U, x_M^U, l_M^U, x_F^S, l_F^S, x_M^S, l_M^S, \vec{k}) - \vec{r_k}^T \vec{k} \\ & - w_M^U(x_F^U) l_M^U - \left(\tau_a^U + \tau_m^U w_F^S(x_F)\right) l_F^U - w_M^S(x_M) l_M^S - \left(\tau_a^S + \tau_m^S w_F^S(x_F)\right) l_F^S,\end{aligned} \tag{2.7}$$

where $\vec{r_k}$ is the vector with the rental cost of capitals and it is exogenous to the model. Depending on how capital enters the production function, it can be more complement to blue or white-collar workers. The sole addition of capital in the model does not interact with how the matching of workers and firms occur in equilibrium. However, it is debated in the literature (Sánchez et al., 2020; Marjit and Oladi, 2022) whether capital is equally complement to men and women. For instance, if capital is more complement to those workers that can expand more flexibly their labor supply, it would be the case that the gender pay gap is magnified by the introduction of capital in the model.[29] However, assuming gender-specific complementarities breaks the gender symmetry in the production function and mechanically implies the matching function between firms and workers becomes asymmetric by gender. In alternative to this view, we propose in the next session that the introduction of capital in the model has the potential to affect women labor market participation and, via this channel, propagate throughout the economy.

2.4 Discussion for a quantitative exercise

In the previous section we proposed a rich matching theory that has the potential for explaining how the gender pay gap has evolved across and within firms and how to relate the changes in the distribution of the wage gap to occupations, capital intensity and taste-based discrimination. In this section we discuss more in detail what is the potential for the complete version of the model of capturing the patterns we documented in our data analysis.

Let us call "marginal type of women" the lowest x_F type assigned to a firm. If τ_a^U or τ_a^S are sufficiently high that proposition 1 does not apply, there are women at the bottom of the x_F^S or x_F^U distribution that remain unemployed because their surplus when assigned to a firm would be too low. As a consequence, the marginal type of women employed in production is strictly higher than bottom type. This type of women are matched, in the case the allocation is PAM, with bottom male workers because men are not affected by the same distortions in τ_a as their female counterpart.

Suppose to compare two economies identical in their primitives but for the cost of capital in $\vec{r}$. The economy with lower $\vec{r}$ is more capital intensive. In the more capital intensive economy, jobs are more productive and the marginal type of women x_F^S or x_F^U is closer to the bottom of the respective type distribution compared to the less capital intensive economy. Under PAM, the marginal type of women always pairs with bottom men. Therefore, lowering the marginal type of women being employed has a cascade positive effect on the marginal product of all the other women by freeing up better men types to be matched to. In this scenario, women benefit more than men from an increasingly capital intensive economy: in the more capital intensive economy men are paired to worse type of women, while it is the contrary for women. In summary, the gender pay gap reduces across firms.

However, to determine how much this cascade effect propagates across the women's type distribution one needs to look at τ_m^S and τ_m^U. The higher is their value the less sensitive are top women to relocation of workers at the bottom of the distribution. At the same time, the lower is the value for τ_m, the larger it becomes the women skill pay gap when production becomes more capital intensive.

The scenario just described about lowered rental cost of capital is realistic in the sense that it is a common policy-instrument for governments to subsidize capital investment. Subsidies incentivize firms to become more capital intensive and this has the potential to trigger the series of events that our model helps to predict. Based on such predictions,

[29] As an additional layer of complexity, the labor complementarity to capital, by gender, might have also changed over time. For instance, as a consequence of the increased possibilities of working from home and the more flexible work arrangements since Covid, it might be the case that women labor supply elasticity has increased more than for men (Alon et al., 2020). We ignore this in our static model, yet post-pandemic (not considered in our work), women might have become relatively more complementary to capital than men.

we argue that our theory is well-suited in explaining the empirical evidence we reported in section 2.2 and, more generally, in framing what is the relationship between the gender pay gap and capital intensity at the firm-level and in the economy.

2.5 Conclusion

In this paper we documented how the gender wage gap has evolved across firms in Italy between 1998 and 2018. We linked this evolution to those of the skill pay gap and of capital intensity. We proposed a theory of this relationship that leverages on the notions of skill-biased technological change and increased women participation to the labor market. In the presence of discrimination, some women do not find an employment. However, their participation is revised when firms become more capital intensive and need more labor, especially among white-collars. A larger base of working women has positive spillovers all along the women distribution via equilibrium effects. As a consequence, the gender pay reduces yet the skill pay gap may increase.

We acknowledge that simulating and estimating the model is an important future step to advance our thesis. In spite of the qualitatively promising predictions, little can be said on the magnitude of these predictions from using theoretical speculation only. Moreover, quantifying the effects of increased capital deepening on the reduction of the gender pay gap has important policy applications.

Chapter 3

Coworker Learning with Firm Dynamics

Gabriele Macci

Abstract

In this paper I use a firm dynamics model to study coworker learning opportunities after accounting for endogenous workforce composition. The theory is a general equilibrium model with human capital spillovers among colleagues. The allocation of human capital across firms is efficient as it optimally trades-off learning and production complementarities by firm productivity level. Additionally, firms continuously hire and separate from workers of different human capital types so to respond to productivity shocks. The estimation strategy uses employer-employee data where I derive novel stylized facts on coworker spillovers conditional on firm growth and size. Ultimately, I argue that a large pool of small growing firms is essential to generate opportunities for coworker learning.

3.1 Introduction

None of us would think our colleagues were assigned to them randomly. Depending on the occupation and line of business that one is in, one will interact with a diverse network from within and outside of their organization. The quality of these interactions significantly affects the potential for on-the-job learning and shapes individual workers' career trajectories. However, the overall supply of opportunities to learn from coworkers for the economy as a whole is constrained by the equilibrium production structure across firms.

This paper examines these constraints based on a production structure that arises from a standard firm dynamics model with labor market frictions (Hopenhayn, 1992; Schaal, 2017), augmented with human capital spillovers occurring *within* the firm. Firms make personnel decisions based on a trade-off between production (static) and learning (dynamic) complementarities when forming teams through new hires or separations. As a result, learning opportunities and their accessibility vary across firms. One might conjecture, for example, that unemployed workers are less likely to join firms that provide more learning opportunities compared to those who are already employed, which can negatively affect their future career prospects. At a more general level, the allocation of workers across firms affects aggregate human capital, and I argue that studying the joint firm size and growth distributions can shed light on this phenomenon.

The endogenous allocation of workers across firms is the main challenge to overcome when studying coworker learning. To tackle this, I follow a structural approach as mobility decisions of workers and firm hiring and firing decisions are derived as solutions of constrained optimization problems (Jarosch et al., 2021; Herkenhoff et al., 2018). This approach carefully models how coworkers interact over production and learning activities so that workers are willing to forgo higher paying positions and move to high-learning firms if convenient for their career. Compared to the rest of the literature on coworker learning, the innovative aspect of my work is accounting for firm labor demand when studying the dynamics of workers allocation. Firms actively engage in keeping certain types of workers while letting go of others so to respond to productivity shocks. In other words, worker mobility decisions are the joint product of interactions with coworkers and with their employer.

The next challenge when studying coworker learning is comparing wages against peers and across time. These comparisons are widespread practice when studying coworker learning since wage levels and wage growth reflect, respectively, human capital and learning processes. However, interpreting the comparisons is difficult at best because wages are only a noisy measure for human capital. I deal with this challenge indirectly by replicating in the model a series of forces that shape wages in the data. I follow Schaal (2017)'s recursive contracts approach so that wages not only reflect labor productivity, but they also factor in the probability of workers and firm future events such as human capital changes, unemployment, job changes and firm productivity shocks. In addition, contrary to Jarosch et al. (2021), the labor market in Schaal (2017) is frictional so new hire wages also include hiring bonuses. Such a rich protocol for wage determination proves handy in the estimation of key parameters of the model.

The model estimation strategy is via simulated method of moments.[30] The focus of the estimation is on salient data moments about worker selection into firms and wage dynamics based on coworkers' characteristics. I follow the tradition started by Jarosch et al. (2021) and Herkenhoff et al. (2018) of regressing worker future wage and job-to-job transition probabilities on coworker average wage. The coworker coefficients in these regressions are generally positive and stronger for lower-than-average paid workers. Notably, the regression *coefficients* are stable across firm size and growth classes, suggesting the coworker spillovers are the same across firms conditional on colleagues. However, because the characteristics of coworkers vary across firm size and growth classes, the regression *predictions* are different. In particular, compared to the universe of firms, predictions of coworker spillovers are the strongest for small growing firms. Therefore, the coefficients and the predictions of these regressions offer an important set of moments to match in the estimation process.

Ultimately, I discuss how to use the estimated model to further understand the role of firm heterogeneity in creating coworker learning opportunities for the aggregate economy. I describe a series of counterfactuals shocking the incentive of firms to become bigger, to grow faster and to be a more homogeneous pool. Consistent with the forces present in the model, I argue for the importance of firm size and growth heterogeneity in sustaining a higher level

[30]Notice a full-fledged estimation of the model is undergoing. Consequently the parameter values and quantitative results from the model are preliminary. Despite parameter choices are informed by the data, the simulation fails to replicate some salient characteristics of the data.

of aggregate human capital.

This paper contributes to the literature on coworker learning by highlighting the role of firms in the team formation process. Jarosch et al. (2021) seminal work propose a competitive labor market framework where any difference in wages or team composition is directly attributable to differences in knowledge. Herkenhoff et al. (2018) relax this simplification by introducing a frictional labor market and look at job mobility events as key for the diffusion of knowledge in the economy. However, they follow a combinatorial approach to model job mobility decisions which proves cumbersome and limits the scope of their modelling to teams employing not more than two workers. By proposing that firms hire a continuum of workers, I can overcome this limitation and account for firm size. As firms grow larger, the marginal product of labor decreases and this impacts on the firm workforce composition motive.

My work also differs from other frictional labor markets studies on workplace learning[31] (Gregory, 2021; Lee, 2021; Hong, 2022; Ma et al., 2023) as I am the first one using directed search to determine which firm-worker matches occur in equilibrium.[32] This choice comes naturally as firm recruitment is generally targeted towards specific human capital types and workers rarely search for positions that they are under- or over-qualified for (Bagger et al., 2020; Carrillo-Tudela et al., 2020). On top of that, contrary to random search (Herkenhoff et al., 2018; Nix, 2020; Gregory, 2021), directed search allows for the equilibrium being constrained efficient despite the presence of a coworker externality. By assuming that information is perfect and wage contracts are complete and state-contingent, firms correctly internalize learning spillovers and there is no improvement left for the use of subsidies. One can therefore view the outcome as the optimal choice of learning given production considerations.

There is also a large macroeconomic literature that investigates the role of bilateral meetings as a key mechanism for economic growth. Lucas (2009); Perla and Tonetti (2014); Lucas and Moll (2014); Jovanovic (2014); Luttmer (2015); Benhabib et al. (2021); Herkenhoff et al. (2018); Akcigit et al. (2018); Caicedo et al. (2019) and, Ma et al. (2023) developed models on the determinants of these meetings and their aggregate consequences. While my work is developed for a stationary economy, it talks to this literature by proposing and quantifying the role for firms as ultimate arrangers of these meetings.

My work also relates to the literature on firm dynamics and frictional labor markets (Kaas and Kircher 2015; Schaal 2017; Gavazza et al. 2018; Bilal et al. 2022, among others). I follow Kaas and Kircher (2015), Schaal (2017), Shi (2018), Kaas (2020) and Carrillo-Tudela et al. (2020) in assuming the labor market operates under directed search. I deviate from the canonical model by introducing different human capital types. In-house human capital accumulation is a new dimension of firm growth. When assembling their workforce, firms can decide to reach their optimal size and composition either by jumping to it at once, or over a sequence of periods through in-house human capital accumulation. Extending Schaal (2017) for multiple human capital types also allows the model to capture the qualitative

[31] The works of Gregory (2021) and Ma et al. (2023) document firm-level differences in human capital accumulation. However, they attribute part of these differences directly to the employer characteristics and not to coworkers. In this sense, I refer to workplace learning as a broader category than coworker learning.

[32] Previous work by Moen and Rosén (2004) also focuses on endogenous training and poaching choices using directed search and arguing the economy reaches constrained efficiency. However my work extends theirs by making the training technology endogenous to the firm and possibly evolving over the firm life-cycle. In my work every firm at any time can specialize into training or production, in the language of Moen and Rosén (2004), since learning intensity depends on workforce composition.

feature of churning, that is excess gross job flows, which is empirically relevant as discussed by Elsby et al. (2021) for the US and by Grinza (2021) for the Italian data.

I also contribute to the literature on labor market sorting between workers and firms (Abowd et al. 1999 and Eeckhout and Kircher 2011, among many others) and between workers and coworkers (Anderson 2015 and Lopes de Melo 2018). Like in Gulyas (2020), the optimal matching between firm and worker types varies along the firm life cycle. Freund (2022) roots the success of superstar firms in the composition of their workforce. While my work does not enter in the details of the production function at a similar level, I share the result that, all things equal, firm future trajectory is shaped by the composition of human capital types in the team.

Finally, my results are complementary to other works studying coworker learning on the same Italian data source, Veneto Worker Histories.[33] Arellano-Bover and Saltiel (2022) and Hong and Lattanzio (2022) adopt a reduced-form approach and find firm characteristics only mildly predict coworker learning opportunities. Despite not directly comparable, my results point towards a more significant role for firms once that team formation motives are factored in. Hong (2022) adopts a structural approach like Herkenhoff et al. (2018) and finds that the joint effect of firms and coworker complementarities explain about 60% of workers' lifetime income variation.

The rest of the paper is structured as follows. Section 3.2 introduces the model environment and its equilibrium properties. Section 3.3 presents the data, the estimation strategy and describes some quantitative properties for a model simulation. Section 3.4 is devoted to discuss comparisons with counterfactual economies that vary in average firm size, average firm growth rate and the degree of homogeneity across firms. Section 3.5 concludes.

3.2 Model

In order to study coworker learning after accounting for endogenous workforce composition, I build a firm dynamics model with labor-search frictions and within firm human capital spillovers. The fact that production occurs in firms, not just in pairs of workers (Herkenhoff et al., 2018), is crucial to get the right dynamics for labor market transitions. Otherwise high productive pairs never hire and never share their knowledge. Similarly, low human capital pairs can only get to work with high human capital individuals after sending one of the incumbent workers to unemployment. Therefore, firm size is crucial to study the dynamics of coworker selection without mechanically constraining the inflows to the outflows of workers.

The second feature of the model compared to other theories of coworker learning are firm productivity shocks. Firm productivity shocks imply that the fit of workforce with a certain employer and of workers with other workers is time-varying Gulyas (2020). For instance, a very productive team might become more open to hiring in certain periods than in others, generating a correlation between learning opportunities and firm growth.

Below I describe the different blocks of the model. The model backbone is inspired by Schaal (2017).

[33] Other related works using Veneto Worker Histories are Battisti (2017) on contemporaneous peer effects, Parrotta and Pozzoli (2012), Serafinelli (2019) and Poggi and Natale (2020) on firms' learning by hiring and Grinza (2021) on gross and net worker flows and firm performance.

Population and Technology Time is discreet and infinite. The economy is populated by risk-neutral workers and firms that discount the future at rate β.

Each worker has a human capital level $h \in \{1, \dots, H\}$ and H is finite. Human capital level is time-varying as it can accumulate or depreciate over time (more afterwards). Workers are subject to stochastic death and survive with probability s_0, independently from the human capital or employment status. For each worker dying, a new one enters the labor force as unemployed and with initial human capital level drawn from distribution $\pi_0(h)$. The total mass of workers is stationary and normalized to one unit, so that every period a new cohort of mass $1 - s_0$ enters the economy.

Firms produce all an identical good using the production function $e^z F(\vec{n}\,)$. Productivity $z \in \mathcal{Z}$ is idiosyncratic and evolves over time independently across firms according to the Markov process $\pi_z(z', z)$. The workforce is described by vector $\vec{n} = [n_1, \dots, n_H]$, where $n_h \in \mathbb{R}_+$ is the *mass* of workers of type h. The function F is increasing and concave in each of its components and $F(\vec{0}\,) = 0$.

In every period, incumbent firms can exit and new firms can enter the market. At entry, firms pay a sunk cost k_e and draw their first productivity draw from distribution $\pi_e(z)$. Each active firm pays an operational cost k_f every period. At any time, including at entry, active firms can decide to exit the market and get a normalized payoff of zero.

Human capital accumulation A worker can either be unemployed or employed at a certain firm $(z, \vec{n}\,)$. Human capital accumulates off-the-job and on-the-job, respectively, according to Markov processes $\pi_u(h'|h)$ and $\pi_w(h'|h; \vec{n}\,)$. When unemployed, human capital tends to depreciate $\mathbb{E}\pi_u(h'|h) \leq h$. On the contrary, human capital accumulates while employed, $\mathbb{E}\pi_w(h'|h; \vec{n}\,) \geq h$. Intuitively, given a worker of type h, a set of coworkers $\vec{n}$ provides better learning opportunities than coworkers $\hat{\vec{n}}$ if $\pi_w(h'|h; \vec{n}\,)$ first order stochastic dominates $\pi_w(h'|h; \hat{\vec{n}})$.

Since we assumed firms employ a continuum of workers, despite the evolution of an individual human capital is stochastic, the firm incumbent workforce follows a deterministic path. For a given workforce $\vec{n}$, the next-period mass of incumbent workers of type h is:

$$n'_h = s_0 \sum_{m=1}^{H} n_m \pi_w(h|m, \vec{n}\,).$$

The expression above only describes in-house workforce evolution. Notice, however, that the firm workforce follows a non-deterministic path as productivity realizations are stochastic and they are reflected into hirings, firings and quits.

Labor market Search is directed on the worker and firm side, with firms opening vacancies that are human capital specific. Opening a vacancy for a worker of type h means posting a contract promising utility x_h to workers upon matching. Firms offering the same contract (x_h, h) compete on the same submarket. The labor market for workers of type h is described as a continuum of submarkets $x_h \in [\underline{x}_h, \bar{x}_h]$. A vacancy catering intended for worker of type h costs c_h units of output and attracts $\theta(x_h, h)$ searchers. The matching function displays constant return to scale so that the job finding probability can be represented as $p(\theta(x_h, h))$, with p strictly increasing and concave. The job filling probability is $q(\theta) = p(\theta)/\theta$, with q is strictly decreasing and convex.[34] We assume there is on-the-job search. Employed worker search efficiency, though, is lower than unemployed by a factor of $\lambda < 1$. The equilibrium tightness can be written as $\theta(x_h, h) = \nu/\mu$, where ν stands for the number of vacancies posted on submarket (x_h, h) and μ is the corresponding efficiency-weighted number of searching workers. The number of vacancies posted per firm does not need to be a discreet number so that a firm posting ν vacancies on a submarket with tightness θ exactly hires $\nu q(\theta)$ workers of that type.

I assume workers search in one market per period and that there is no search in the period of transition from employment to unemployment. On the other hand, for each worker type h, firms can post in multiple markets. Which firms end up posting in which submarket is determined in equilibrium by ranking firms separately for each human capital market. Firms are ranked based on the relative and absolute number of workers they want to hire. The higher a firm rank is, the higher is the promised utility and, therefore, the higher is the yield per vacancy. This criteria is an assumption of the model as with linear vacancy costs firms would be indifferent between opening more vacancies or having a better yield per vacancy. However, this assumption can be rationalized by thinking of taking the limit of an underlying *convex* vacancy cost function as it becomes linear. Kaas and Kircher (2015) showed that a convex vacancy cost implies firms hiring more workers find optimal to post vacancies with a higher yield.

Contracting Contracts regulate several aspects of the job relationship like workers' remuneration and firing probabilities. As contracts are expressed in a recursive formulation, they also need to record the lifelong amount of utility promised for delivery by the firm to the worker. Every contract is described by a tuple

$$\omega = \{w, \tau, d, W'\},$$

where w is the wage paid by the firm after matching and before production take place (see stage C in figure 3.1). The rest of the variables refer to the subsequent period realization of $(z', \vec{n}\,', h')$, as we may say that contracts are complete and state-contingent. In particular, τ denotes the firing probability, d indicates whether firm $(z', \vec{n}\,')$ will exit the market or not and it is identical for all workers in the same firm, finally W' is the next-period promised utility. In every period, the firm in $(z, \vec{n}\,)$ is committed to deliver to each worker of type h at least $W(z, \vec{n}\,, h)$ as it was agreed one period in advance (firm-side commitment).[35] For new hires, the first period promised utility corresponds to the market utility posted on the market during the hiring phase, x_i.

Every period, after matching and before production take place, contracts are re-written so to account for the changes in the firm and worker state space $(z', \vec{n}\,', h')$ and to discount future utility promises by the present period paid wage. As it will be clearer from the firm problem below, notice that next-period employment $\vec{n}\,'$ is a deterministic function of z', as we can write $\vec{n}\,' = \mathbf{\vec{n}}'(z', \vec{n}\,)$. Therefore, it is equivalent to regard contracts as contingent on the tuple $(z', \vec{n}\,, h')$.

[34] See footnote and/or assumption 2 from Schaal 2017's online appendix for more details.

[35] Footnote on how it compares to Schaal (2017). Discuss both the logical step from sequential contracts to recursive and the fact that I assume right away they are homogeneous for certain group of workers. Further, I immediately assume that workers do not commit.

Beginning of period t.
STAGE A.

- **Learning** individual human capital evolves from h^- to h. It follows the Markov process π_u for the unemployed and π_w for the employed;
- **Birth and Death**: survival probability is s_0, newborn workers enter the market as unemployed with HC distribution π_0.

STAGE B.

- **Entry decision**: no (payoff is 0)/ yes (pay entry cost k_e);
- **Productivity** idiosyncratic draw z is realized for every firm (incumbents and entrants);
- **Exit decision**: yes (payoff is 0)/ stay;
- **Layoff decision**: individual worker is fired with probability τ;
- **Search and matching**:
  - firms choose employment $\vec{n}$ and hire accordingly;
  - employed workers of type h in firm $(\vec{n}, z)$, search in (x_w, h);
  - unemployed workers of type h search on market (x_u, h).

STAGE C.

- **Contracts**: renegotiation and payment of wages;
- **Production**
  - firms produce according to $F(z, \vec{n})$ and pay operating cost k_f;
  - unemployed workers produce b_h.

End of period t.

Figure 3.1: Timing

Worker's problem Value functions are expressed, by convention, when production takes place (see stage C, figure 3.1).

The value of unemployment for a type h worker is:

$$\mathbf{U}(h) = \max_{x_u(h')} b(h) + s_0\beta\mathbb{E}_h\left[p(\theta(x_u(h'), h'))x_u(h') + (1 - p(\theta(x_u(h'), h')))\,\mathbf{U}(h')\right], \quad (3.1)$$

where the right hand side is the sum of the current unemployment benefit possibly varying in h and the future discounted value of finding a job promising utility $x_u(h')$ or remaining unemployed. Notice that next period human capital is h' and so the expectation operator is with respect to the Markov process π_u. Also notice that the market where search occurs is contingent on the future human capital level and that the search problem of an agent does not depend on h but only h'. In other words, regardless of current h, every unemployed with realization h' in the next period will search in the exact same submarket.

The value of employment at firm $(z, \vec{n}\,)$ for a worker of type h with contingent contract $\omega = \{w, \tau(z', \vec{n}\,', h'), W'(z', \vec{n}\,', h'), d(z', \vec{n}\,')\}$ is obtained after solving for worker's optimal contingent search strategy $x_w(z', \vec{n}\,', h')$ and equals (omit contract's and x_w arguments):

$$\begin{aligned}\mathbf{W}(z, \vec{n}\,, h, \omega) = \max_{x_w} \quad & w + s_0 \beta \mathbb{E}_{z,h} \Big[\big(d + (1-d)\tau\big) \mathbf{U}(h') \\ & + (1-d)(1-\tau) \left[\lambda p(\theta(x_w, h')) x_w + \big(1 - \lambda p(\theta(x_w, h'))\big) W' \right] \Big],\end{aligned} \tag{3.2}$$

where $x_w(z', \vec{n}\,', h')$ is the promised utility attached in the submarket where employed workers direct their search if next-period human capital is h' and firm state space becomes $(z', \vec{n}\,') = (z', \vec{\mathbf{n}}\,'(z', \vec{n}\,))$. The expectation operator is over the product of π_z and π_w. The value function is the sum of the current period wage and the expected continuation value. The continuation value is adjusted for the survival probability s_0 and discount rate β and is the probability weighted sum of the value of unemployment ($\mathbf{U}$), job-to-job transition (x_w) and staying in the firm (W').

Notice that the decision on which submarket to search is completely forward looking both in the case of unemployed and employed workers. In particular, employed workers whose next period state is $(z', \vec{n}\,', h')$ will search on

$$x_w(z', \vec{n}\,', h') = \arg\max_{x} \quad p(\theta(x, h'))(x - W'(z', \vec{n}\,', h')) \tag{3.3}$$

where $W'(z', \vec{n}\,', h')$ is determined in current period contract ω.

Firm's problem Firm $(z, \vec{n}\,)$ problem consists in managing its employment relationships with incumbent workers and hiring new workers.

Incumbent workers. The firm manages incumbent workers in stage C by renegotiating contracts. In order to simplify the exposition, I will describe incumbent workers in state $(z, \vec{n}\,, h)$ as homogeneous in their inherited contract ω^- and in particular in their current-period promised utility level $W(h)$ and (stage B) firing probability $\tau(h)$. Once τ is assumed constant per human capital class, a constant level for W follows as a result.[36] Assuming a constant τ across incumbent workers $(z, \vec{n}\,, h)$ is without loss of generality regarding the total number of hires, quits and layoffs that occur in a firm by human capital class in equilibrium.[37] The firm renegotiates ω by deciding on how much to pay workers and what contingencies will occur in the next period depending on the realization of z'. In renegotiating, we assumed the firm must respect the promise keeping constraint:

$$\forall h: \quad W(h) \leq \mathbf{W}(z, \vec{n}\,, h, \omega) \tag{3.4}$$

[36] Further intuition on why τ constant implies W constant is discussed in the wage-determination section below.

[37] To be precise, it is only the firm total number of worker inflows and outflows per human capital class that it is unaffected as discussed by Schaal (2017) in Appendix E. However, the exact firm-to-firm flows would be altered if there is firm heterogeneity in τ within human capital class. A different τ affects worker search x_w as derived in (3.2) which, in turn, reflects into visiting different submarkets that are potentially populated by different firms. The firm-to-firm flows remain unaltered only when all firms post vacancies in all submarkets proportionally to the number of searchers per submarket.

where $\mathbf{W}$ is the value obtained from the workers maximization problem (3.2) after they solved for their optimal contingent search strategy $x_w(z', \vec{n}', h')$ in (3.3).

Entrant workers. In stage B the firm decides how many workers to hire and on which submarkets.[38] As firms can recruit workers of type h across a multitude of submarkets, the firm recruiting problem is described by the function $\tilde{n}_i(x_i, h; z, \vec{n})$ that indicates for each submarket (x_i, h) the mass of workers recruited by firm $(z, \vec{n})$. This function is well defined as x_i takes values in a closed interval $[\underline{x}_h, \bar{x}_h]$ that will be described in the free-entry section below. In stage C, the firm also renegotiates contracts ω with new hires and must keep the following promise:

$$\forall (x_i, h) \quad \text{s.t. } \tilde{n}_i(x_i, h; z, \vec{n}) > 0 : \quad x_i \leq \mathbf{W}(z, \vec{n}, h, \omega), \tag{3.5}$$

that is the firm will choose ω (including the first wage) so that new hires lifelong value $\mathbf{W}$ is no lower than what accepted in the recruiting phase. Notice that new hires become incumbent workers in the next period and, therefore, $W'(z', \vec{n}', h')$ is the same in (3.4) and (3.5).[39] It follows (see the wage equation in 3.14 below) that current wages of new hires do not need to equal those of incumbent workers. For notation purposes, we will refer to wage of incumbents and new entrants, respectively, as w_h and $\tilde{w}_h(x_i)$. Finally, we will refer to $\tilde{\omega}(h)$ as all the elements of a contract other than the wage. It follows $\tilde{\omega}(h)$ is identical for workers of type h regardless if they are incumbent or new hires.

After the matching stage has occurred and the firm recruited $\overline{\tilde{n}_i(x_i, h)}$ workers, the overall firm problem in stage C can be described as follows:

$$\begin{aligned}
\mathbf{J}\left(z, \vec{n}, \overline{\tilde{n}_i(x_i, h)}, \{W(h)\}_{h=1}^H\right) = \max_{\tilde{n}_i(x_i, h; z', \vec{n}), \tilde{w}(h, x_i; z, \vec{n}), w(h; z, \vec{n}), \tilde{\omega}(h; z, \vec{n})} & F(z, \vec{n}) - k_f \\
& - \sum_{h=1}^H \left[w(h; z, \vec{n}) n_h + \int_{\underline{x}_h}^{\bar{x}_h} \tilde{w}(h, x_i; z, \vec{n}) \overline{\tilde{n}_i(x_i, h)} \quad \mathrm{d}x_i \right] \\
& + \beta \mathbb{E}_z \left\{ - \sum_{h=1}^H \int_{\underline{x}_h}^{\bar{x}_h} \tilde{n}_i(x_i; z', \vec{n}') \frac{c_h}{q(\theta(x_i, h))} \quad \mathrm{d}x_i \right. \\
& \left. + \mathbf{J}\left(z', \vec{n}', \tilde{n}_i(x_i, h; z', \vec{n}), \{W(h; z', \vec{n}')\}_{h=1}^H\right) \right\}^+,
\end{aligned} \tag{3.6}$$

subject to constraints (3.4) and (3.5) and law of motion:

$$n'_h(z', \vec{n}) = \int_{\underline{x}_h}^{\bar{x}_h} \tilde{n}_i(x_i, h; z', \vec{n}) \mathrm{d}x_i + s_0(1 - \tau(z', \vec{n}, h))(1 - p(\theta(x_w(z', \vec{n}, h), h))) \sum_{m=1}^H \pi_w(h|m, \vec{n}) n_m. \tag{3.7}$$

The firm maximizes its value by choosing on how many workers to recruit in the next period contingent on the realization of z', and on how to compensate its current workforce in $(z, \vec{n})$, so to keep its promises as summarized in $(\overline{\tilde{n}_i(x_i, h)}, \{W(h)\}_{h=1}^H)$ for new hires and incumbent workers, respectively. The first line of (3.6) represents the value of production

[38] There is no need to separately track the number of opened vacancies. Given the number of hires in a specific submarket (x_i, h), the number of opened vacancies will directly follow by inverting for the job filling probability function q.

[39] In other words, new hires are as likely to get fired as incumbent workers, conditional on h'.

net of the fix operating cost k_f, the second line details the wage of incumbent and new hires. Notice that the number of new hires in stage C is given while their wage is chosen by the firm. The last two lines are the non-negative continuation value of the firm, contingent on the realization of z' according to $\pi_z(z'|z)$ and on not exiting the market. The third line describes the cost of hiring $\tilde{n}_i(x_i; z', \vec{n}\,')$ units of labor in submarket (x_i, h). The fourth line is the firm problem in stage C of the next period. Notice that in this problem $W'(z', \vec{n}\,')$ is implicitly pinned down by the choice of $\tilde{\omega}(h; z, \vec{n}\,)$ as it prescribes what value is assigned to W' in the case z' realizes. Finally, notice that the state space W has been omitted from the policy functions arguments. This is a small abuse of notation, but as it will soon be clear from the next section has no practical repercussions.

Joint surplus maximization Optimal policies maximize the joint surplus of a firm and its incumbent workers. This is the case because utility is transferable, contracts are complete and state-contingent, and firms meet their promises. This can be formally established by building on Schaal (2017, appendix G.1) as it readily extends to the case of multiple human capital types and stochastic worker death.

In stage C, the joint surplus maximization problem for a firm and its workers (incumbent and just hired in stage B) is:

$$\begin{aligned}\mathbf{V}(z, \vec{n}\,) = \max_{d(z',\vec{n}\,),\tilde{n}_i(x_i,h;z',\vec{n}\,),\tau(z',\vec{n}\,,h),x_w(z',\vec{n}\,,h)} \quad & F(z, \vec{n}\,) - k_f \\ + \beta \mathbb{E}_z \Bigg[s_0 \sum_{h=1}^{H} \sum_{m=1}^{H} \pi_w(h|m; z, \vec{n}\,) n_m & \Bigg(\mathbf{U}(h) \big(d(z', \vec{n}\,) + (1 - d(z', \vec{n}\,)) \tau(z', \vec{n}\,, h) \big) \\ + x_{w,h}(z', \vec{n}\,, h; m, j)(1 - d(z', \vec{n}\,))(1 - \tau(z', \vec{n}\,, h)) \lambda p(\theta(x_w(z', \vec{n}\,, h), h)) \Bigg) \\ + (1 - d(z', \vec{n}\,)) \Bigg(V(z', \vec{n}\,(z', \vec{n}\,) + \sum_{h=1}^{H} \kappa(h) \int_{\underline{x}_h}^{\bar{x}_h} \tilde{n}_i(x_i, h; z', \vec{n}\,) \mathrm{d}x_i \Bigg) \Bigg], \end{aligned} \tag{3.8}$$

subject to the low of motion for $\vec{n}$ described in (3.7). The surplus function is obtained after accounting for the next period optimal contingent decisions in terms of shutting down, recruiting, firing, and job-to-job transitions. As all decisions regard the next period, the problem is purely forward looking. The surplus value equals the flow value of stage C and the expected continuation value. Notice that terms in the second and third line are adjusted by the survival probability s_0 and learning process π_w as they realize in next period stage A. These two lines describe, respectively, incumbent workers future gains from becoming unemployed and changing employer. The last line first component describes the continuation value for the firm and next period employed workers as in stage C. The second component captures the cost of next period hiring where the integral in $\tilde{n}_i$ gives the number of hires of type h and the $\kappa(h)$ is the cost of hiring, including both the vacancy posting component and the relevant submarket utility cost. The reason $\kappa(h)$ is not indexed in x_i will be presented in the free entry section below.

There are more properties about problem (3.8). First, the state space is considerably reduced compared to the firm problem (3.6). This is the principal gain from using the surplus formulation as introduced by Schaal (2017). The gain is that we do not need to keep track explicitly of all the promised contracts of incumbent workers in order to determine with what probability workers are going to leave or how many workers will be recruited. Another difference with the firm problem is that the surplus is maximized over x_w as well. Firms cannot contract with workers where they are going to direct their search. However, firms can indirectly induce a particular search behavior by adjusting W' as it results from (3.3) above. It turns out that in equilibrium firms can always adjust W' so that it maximizes their own profits and, yet, keep their utility promises with the workers (Schaal, 2017, appendix G.3).

Free entry At the beginning of every period, perspective entrants can pay k_e. Upon payment of the entry cost, they draw their first productivity value z from distribution π_e. Depending on the realization of z, firms may decide to stay or exit immediately, respectively denoted as $d_e(z) = 0$ or 1. If staying, they start hiring and producing in stage B and C, respectively. Differently from any incumbent firm, however, we assume that entrant firms can only hire a single human capital type of worker, $h \in \{1, \dots, H\}$. Entrant firms can post an unlimited number of contracts across multiple submarkets within their chosen h. Different entrants can opt for different h types. Despite restrictive, this assumption only constraints hiring behavior in the first period and firm can hire every human capital type from next period onward. On the other hand, this assumption guarantees the model is tractable and an equilibrium exists. As long as across entrant firms every type h is hired, the economy is block-recursive (Menzio and Shi 2010; Kaas 2020) and market tightness can be computed in each possible submarket (x, h) without that agents keep track of the aggregate distribution in search and posting behavior.

The problem facing an entering firm of type z in stage B (notice k_e was already paid and z is drawn) is:

$$\mathbf{J}_e(z) = \max_{\hat{h}, \tilde{n}_e(x_e, \hat{h}; z, \bar{0}), d_e(z)} \quad d_e(z) \left[\mathbf{V}(z, \vec{n}_e) - \int_{\underline{x}_{\hat{h}}}^{\bar{x}_{\hat{h}}} \tilde{n}_e(x_e, \hat{h}; z, \bar{0}) \left(x_e + \frac{c_h}{q(\theta(x_e, \hat{h}))} \right) \mathrm{d}x_e \right], \tag{3.9}$$

where $\vec{n}_e$ is an H-dimensional zero vector except for component $\hat{h}$ that equals $\int_{\underline{x}_{\hat{h}}}^{\bar{x}_{\hat{h}}} \tilde{n}_e(x_e, \hat{h}; z, \bar{0}) \mathrm{d}x_e$. The entry value in (3.9) is equal to the stage C value $\mathbf{V}$ adjusted for the current stage B cost of hiring the initial workforce. Notice that the cost of hiring can be further decomposed. The submarket x_e only appears through the term $x_e + \frac{c_h}{q(\theta(x_e, \hat{h}))}$ which describes a per-unit hiring cost common to both entering and incumbent firms. Thus, entering firm hiring decision can be split in two stages. First, choose in which submarket $(x, \hat{h})$ to search. Second, decide on the number of workers to recruit. The minimal hiring cost for a given $\hat{h}$, across x is:

$$\kappa(\hat{h}) = \min_{\underline{x}_{\hat{h}} \leq x \leq \bar{x}_{\hat{h}}} x_h + \frac{c_h}{q(\theta(x_h, h))}. \tag{3.10}$$

Let us Optimal entry requires only submarkets minimizing hiring costs are open in equilibrium. So the following complementary slackness condition must hold:

$$\forall (x, h): \quad \theta(x, h) \left[x + \frac{c_h}{q(\theta(x, h))} - \kappa(h) \right] = 0. \tag{3.11}$$

So equilibrium market tightness for every open submarket is:

$$\theta(x,h) = q^{-1}\left[\frac{c_h}{\kappa(h) - x}\right]. \tag{3.12}$$

It is clear from (3.12) that no optimal hiring occurs in submarkets (x,h) with $x > \bar{x}_h = \kappa(h) - c_h$ as job finding probabilities would be zero. Similarly, there is no worker willing to search in a submarket that is dominated in value by unemployment so we can write that optimal search implies $x > \underline{x}_h = \min_h\{\mathbf{U}(h)\}$. Further notice that incumbent firms are equally able to split their hiring problems in the same two steps, i.e., choosing the submarkets (x_i,h) deciding on quantities.[40]

Furthermore, the free-entry condition drives expected profits of entrant firms to the entry cost:

$$k_e = \sum_{z\in\mathcal{Z}} \mathbf{J}_e(z)\pi_e(z). \tag{3.13}$$

Notice that when combining equations from (3.9) to (3.13), we are able to pin down the value of $\kappa(\hat{h})$, conditional on $\hat{h}$. The vector $\vec{m}_e$ records how many entrant firms of each type $\hat{h}$ are needed in equilibrium to drive the ex-ante value of entry to zero. Notice that the equilibrium vector $\vec{m}_e$ is unique as proved in appendix C.1 through the use of Walras law.

Equilibrium I follow Schaal (2017) in defining the equilibrium in a constructive way. First, a candidate equilibrium is a solution to both the workers problems (3.1)-(3.2) and firms' problem (3.6) that further satisfies the free entry condition (3.13). A candidate equilibrium is an equilibrium for the economy if the vector $\vec{m}_e$ of entrant firm masses as obtained from (C.12) is strictly positive in every component. Strictly positive entry in each human capital market h guarantees relative hiring costs $\kappa(h)$ are pinned down by the free entry condition. In this case the equilibrium is block-recursive because hiring costs do not depend on the infinite-dimensional distribution of firms.[41] In practice, focusing on equilibria with positive firm entry for each human capital market does not seem to be too unrealistic in the VWH data given how dispersed are the entrant firms hires in terms of pay, past experience and occupations.

Definition 3.1. *Define the following concepts:*

1. *A candidate equilibrium is (i) a set of value functions* $\mathbf{U}(h)$, $\mathbf{W}(z,\vec{n},h,\omega)$, $\mathbf{J}(z,\vec{n},\overline{\tilde{n}_i(x_i,h)},\{W(h)\}_{h=1}^{H})$, $\mathbf{V}(z,\vec{n})$ *and* $\mathbf{J}_e(z)$*; (ii) a decision rule for unemployed workers* $\{x_u(h')\}$*, employed workers* $\{x_w(z',\vec{n},h')\}$*, entering firms* $\{\hat{h},\tilde{n}_e(x_e,\hat{h};z,\bar{0}),d_e(z)\}$*, and incumbent firms* $\{\tilde{n}_i(x_i,h;z',\vec{n}),\tilde{w}(h,x_i;z,\vec{n}),w(h;z,\vec{n}),\tilde{\omega}(h;z,\vec{n})\}$*; (iii) a set of hiring costs* $\{\kappa(h)\}_{h=1}^{H}$ *and corresponding labor market tightness* $\theta(x,h)$ *such that equations (3.1)-(3.13) are satisfied.*

2. *A block-recursive equilibrium is a candidate equilibrium such that* $\vec{m}_e$ *is positive in each entry.*

[40]Compare with (3.6). The only difference for incumbent firms is that the utility promised x does not show up in the value function but in the promise keeping constraint. Also notice this justifies our surplus function formulation in $\kappa(h)$.

[41]A candidate equilibrium violating strict positive entry can still be an equilibrium, yet the hiring costs κ will the depend on the whole distribution of firms.

Proposition 3.1. *Under weak regularity conditions, a candidate equilibrium always exists. (ii) An equilibrium fulfilling block-recursivity, when it exists, is efficient.*

The statement in proposition 3.1 is proved in Schaal (2017, appendix G.2).[42] The proof uses Shrauder's fixed point theorem to show the existence of a joint-solution to the unemployed, surplus and free-entry problem. A surplus solution can always be implemented as a solution to the firm and employed workers problem (see surplus section above) so a candidate equilibrium always exist. Unfortunately, guaranteeing the existence of a block-recursive equilibrium is not as easy as $\vec{m}_e$ depends on the aggregate distribution of firms. In practice, it is convenient to solve for the candidate equilibrium and to check ex-post for positive entry.

Proposition 3.2. *A block-recursive equilibirum, when it exists, is efficient.*

Proposition 3.2 states the economy maximizes total welfare. This property is also inherited from Schaal (2017) and proved therein in appendix G.2. In particular, the human capital accumulation of workers is constrained efficient as in Jarosch et al. (2021). Firms and workers correctly price the value of learning. They write contracts that internalize the current and future value of human capital spillovers on the rest of the workforce and on the firm itself. The market generally does not achieve efficiency in other models of coworker learning (Herkenhoff et al., 2018; Nix, 2020; Ma et al., 2023) as a result of random search or information frictions. Proposition 3.2 establishes that welfare in this economy cannot be improved by a policy intervention. In this sense, the model offers an efficient benchmark in which there is no human capital accumulation mispricing, nor inefficient separations or hirings.

A remarkable difference from Schaal (2017) is that the model can generate labor market *churning* since firms hire and separate from workers at the same time. This is typical of labor markets (Elsby et al., 2021) but it is often a missing feature of theoretical firm dynamics models. The current framework is able to reproduce this feature as firms are contemporaneously active on multiple human capital markets. For instance, it can be the case that some workers have just upgraded their human capital level but there is no demand for them in the firm. Consequently, they will separate – most likely by performing a job-to-job transition — and the firm will replace them with new lower level workers.

Wage determination In equilibrium, contracts will induce workers to search in specific submarkets x_w by fine-tuning the next-period promised utility W' as described above in (3.3). In the *current* period, fixed a value for x_w, expression (3.3) pins down the unique value for W' that makes x_w incentive compatible for the worker.

The current period wage is obtained after inverting the worker value function (3.2). The wage is the difference between the current end-of-period promised utility W' and the beginning-of-next-period expected utility:

$$w(z, \vec{n}, h) = W'(z, \vec{n}, h) - s_0 \beta \mathbb{E}_{z,h} \left[\boldsymbol{W}^B(z', \vec{n}'(z, \vec{n}), h') \right], \tag{3.14}$$

[42]Regarding existence, assumption 1 in Schaal (2017, appendix G.2) needs to further qualified. As production function F takes labor factor as a vector in our model, one needs to assume F is bi-Lipschitz continuous for each of its labor input dimensions (holding fixed the others labor components). Assumption 4 also modifies accordingly by assuming $\bar{n} > \max_h \left\{ \underline{\alpha}_{hV}^{-1}(k_e + k_f) \right\}$.

where the beginning of period utility is (set for brevity $\gamma = \{z, \vec{n}, h\}$):

$$\boldsymbol{W}^B(\gamma) = d(\gamma)\boldsymbol{U}(h) + (1-d(\gamma))\left[\tau(\gamma)\boldsymbol{U}(h) + (1-\tau(\gamma))\left(\lambda p(\theta(x_w(\gamma), h))x_w(\gamma) + (1-\lambda p(\theta(x_w(\gamma), h)))W'(\gamma)\right)\right], \tag{3.15}$$

and the expectation operator is over $\{z', h'\}$ according to $\pi_z(z'|z)\pi_w(h'|h; \vec{n})$.

First of all, notice that wages display *compensating differentials* as workers will receive lower wages if their opportunities to learn are higher. To see this, suppose $h' > h$ implies higher future utilities and that team $\hat{\vec{n}}$ offers better learning chances of team $\vec{n}$. In formulas, $\boldsymbol{W}^B(z, \vec{n}, h') > \boldsymbol{W}^B(z, \vec{n}, h)$ and $\forall h : \pi_w(h'|h, \hat{\vec{n}})$ stochastic dominates $\pi_w(h'|h, \vec{n})$. Therefore, fixed W', wages will be lower in team $\hat{\vec{n}}$ than in team $\vec{n}$ because the last term in expression (3.14) is larger in the case of $\hat{\vec{n}}$. As a corollary, changes in the workforce composition affect workforce payroll through the direct effect on continuation utility $\boldsymbol{W}^B$ and also indirectly via the expectation operator on the future human capital level.

As in Schaal (2017, online appendix F), wages also have other properties. First there is *wage dispersion* as workers in the same firm with the same human capital level will be paid differently if W' differs. This is the case when comparing incumbent workers and new hires. New hires are promised $x_i(h)$ that generally differs from W' and, since $x'_w(z, \vec{n}, h)$ is identical for new hires and incumbent workers, workers with identical human capital h are paid differently according to (3.14) depending on their seniority level.[43] Second, there is a *firm-size premium* because larger firms are more likely —in a statistical sense— to have undergone a growth period recently. As growing firms must be retaining their incumbent workers, they have directed their workers towards high-end submarkets at the cost of promising high level of W', inflating wages as described in (3.14).

3.3 Data and measurement

In this section, I spell out the model functional specification and discuss data moments that are relevant for properly measuring the value of its parameters. In particular, I use the data to document that small growing firms are associated with stronger coworker learning predictions. However, since coworker groups are endogenous with respect to worker and firm characteristics, this data analysis does not directly measure learning function parameters. Therefore, I use the firm dynamics model to account for the endogenous selection of workers across firms and recover the value of underlying parameters. Finally, I discuss some of the equilibrium properties of the simulated model.

3.3.1 Functional Forms, Parameters and Stochastic Processes

Human capital is discrete and indexed with h_k with $k \in \{1, \dots, H\}$ and $h_k = 1 + (k-1)\Delta$ with $\Delta > 0$. Human capital stochastically appreciates of one level if employed:

$$\Pr(h_{k+1}|h_k, \vec{n}) = \phi_0 + \phi_1 \frac{\sum_{j=k+1}^{H}(h_j - h_k)n_j}{\sum_{j=1}^{H} n_j}, \tag{3.16}$$

[43] Notice, however, that the model predicts that the wage is identical across all incumbent workers that have the same level of human capital. To see this, it is sufficient to observe that policy functions for layoffs and quits do not vary within a group of colleagues that have the same human capital level. Consequently, new hires get a wage differential that exactly matches the difference in promised utilities between them and incumbent workers and they become indistinguishable from any other incumbent worker ever after.

with the complementary probability being human capital remains unchanged. This learning function specification builds on the work of Lucas (2009) and Herkenhoff et al. (2018). It assumes human capital does not depreciate while employed and that individuals only learn from more knowledgeable coworkers. The higher is the share of more knowledgeable coworkers —or the human capital difference—, the more likely human capital will appreciate. At the same time, expression (3.16) is free from firm-size effects as the coefficients are normalized by total workforce. This is in line with the empirical analysis below about stable regression coefficients and the finding of Jarosch et al. (2021).[44] Finally, human capital depreciates of one level if unemployed with probability ϕ_u[45]

Firms produce using own productivity z and labor inputs $\vec{n} \in \mathbb{R}^H$.

$$F(z, \vec{n}) = Ae^z \left(\sum_{k=1}^{H} h_k \left(n_k\right)^{\frac{1}{\rho}} \right)^{\rho\alpha} \tag{3.17}$$

with $\alpha \in (0, 1)$. The CES structure allows for coworkers of different human capital levels to be substitutes $\rho < 1$ or complements $\rho > 1$. Depending on the value of ρ, firms have a production motive to assemble a more homogeneous or diverse workforce. The decreasing returns to scale parameter α guarantees that there is a non-degenerate firm-size distribution in the economy.

Productivity draws for incumbents follow an AR(1) process:

$$z_t = \rho_z z_{t-1} + \varepsilon_{z,t} \qquad \varepsilon_{z,t} \sim \mathcal{N}(0, \sigma_\varepsilon^2) \tag{3.18}$$

made discrete for a finite grid $z \in \mathcal{Z}$. The first productivity draw, $z_0 \in \mathcal{Z}$, realizes according to (3.18) with $z_{-1} = 0$.

I follow Menzio and Shi (2010) and Schaal (2017) in picking the CES matching function $p(\theta) = \theta(1 + \theta^\gamma)^{-1/\gamma}$. Home production and vacancy posting cost are set constant across human capital types: $b(h) = b_0$ and $c_k = c_0$ for all h. Finally, the first human capital level at birth is drawn from π_0 that is distributed as a truncated geometric distribution with average $\bar{\pi}_0$.

3.3.2 Data and Empirical Evidence

Compared to a standard model of firm dynamics with frictional labor market, this model also features workers with different human capital levels that interact through production and learning. These interactions are captured by the parameters $\{\phi_0, \phi_1\}$ in learning and ρ in production. To measure the strength of complementarities in production and learning I do not leverage on exogenous variation from case studies (Hong and Lattanzio, 2022) or from an experimental setting (Ichino and Maggi, 2000). Instead, I consider a series of regressions between an individual worker's outcome variables and his coworkers' characteristics. These regressions do not directly measure complementarities in learning or production, because coworkers' characteristics are not a random with respect to individuals. Therefore, I follow

[44] If anything, Jarosch et al. (2021) show that the learning function displays mild decreasing returns to scale. However, I prefer to use a constant returns to scale learning function because I argue it is the *composition of workers* in small firms that makes coworker learning more effective and not directly a stronger learning process.

[45] As the human capital grid is finite, there is no depreciation for $h = 1$ and no further appreciation when $h = H$.

Herkenhoff et al. (2018) and consider these same regressions in the model as well. Since the model explicitly accounts for endogenous sorting of workers into firms and coworkers, I am able to generate the same source of endogeneity present in the data. Consequently, I can use the model to measure indirectly the values for $\{\phi_0, \phi_1\}$ and ρ.

The two regression specifications that I will present below are performed on the whole sample and then repeated by classes of firm size and growth. Repeating the regression on sub-samples of the data allows to verify that the regression coefficient for the coworkers' characteristics are stable across firms. I interpret the stability in the regression coefficients as an indication that the learning and production processes are transversal across firm size and growth classes. Despite the regression coefficient are stable, I document that the outcome predictions differ across firm size and growth sub-samples. I interpret the difference in the predicted outcome variable as an indication that workforce composition differs across firms' classes. This finding is novel in the literature and points towards the importance of considering firm size and growth characteristics when measuring the values of $\{\phi_0, \phi_1\}$ and ρ.

I estimate the model using Italian data from Veneto Worker Histories (VWH) 1982-2001 for the provinces of Vicenza and Treviso. VWH are an employer-employee administrative dataset and report information on every employment spell, regardless of its length, but for the agricultural and public sectors. For each spell, it is reported the total yearly compensation, number of days worker, and a broad occupational variable (trainee, blue-collar, white-collar or manager). Each observation comes with an employer identifiers and, consequently, it is possible to recover granular information on coworkers' characteristics.

Below I present two regression specification borrowed from Herkenhoff et al. (2018) that are indicative of complementarities in learning and production, respectively. Note the regression coefficient do not measure directly the value of parameters in the model, but they can be used as guide when to chose parameter values after running the same regressions in the simulated model as well.

Complementarities in learning For this regression specification I only consider workers continuously employed at same employer f throughout year t, that go through at least one quarter of unemployment in year $t+1$ and return to be continuously employed in year $t+2$ at a different employer from year t.[46] For this sample of workers, I consider the following specification:

$$w_{i,t+2} = \alpha_0 + \alpha_1 w_{i,t} + \alpha_2 w_{-i,t} + AX_{i,t} + \varepsilon_{i,t}, \tag{3.19}$$

where $w_{i,t+2}$ and $w_{i,t}$ are worker wages i wages and $w_{-i,t}$ is the average coworkers' wage in year t. Finally $X_{i,t}$ controls for year, worker gender, age, age squared, nationality, tenure in the firm, occupation, firm age, and a 10 groups k-means dummy based on firm wage mean and standard deviation.

In this specification, wages are a proxy for (unobservable) human capital. The claim is the higher the average wage $w_{-i,t}$ in a firm, the higher is the human capital level of one's coworkers. Since we are controlling for individual i wage, all we need to suppose is that wages are a good proxy in ranking human capital *within* a given firm. In this sense, the

[46] When workers do not appear in VWH, it is not possible to distinguish whether they went through unemployment, moved abroad or got employed in the agricultural or public sector. Notice, however, that VWH follows workers also outside of Veneto as long as they remain in Italy. Finally notice that requiring workers to be employed both in t and $t+2$ in Veneto maximizes the chances that a quarterly missing spell in year $t+1$ actually occurred because of unemployment.

Complementarities in learning - (3.19)

	All firms (1)	Small shrinking firms (2)	Small growing firms (3)	Big shrinking firms (4)	Big growing firms (5)
Panel A					
$w_{i,t}$	0.704***	0.653***	0.646***	0.723***	0.745***
	(0.0046)	(0.0086)	(0.0065)	(0.0103)	(0.0073)
$w_{-i,t}$	0.070***	0.061***	0.081***	0.077***	0.077***
	(0.0074)	(0.015)	(0.0113)	(0.0202)	(0.0154)
Controls	yes	yes	yes	yes	yes
R^2	0.66	0.58	0.59	0.68	0.71
# of th. obs.	259	49	76	48	77
# of th. firms	35	18	23	4	5
Panel B					
mean($\hat{w}_{i,t+2} - w_{i,t}$)		0.076	0.083	0.047	0.045
		(0.0004)	(0.0003)	(0.0004)	(0.0004)

Table 3.1: Estimation of specification (3.19) and average prediction of $\hat{w}_{i,t+2} - w_{i,t}$ by firm class. Column (1) includes all firms while columns (2) to (5) only include firms by their class. Small (large) firms employ at most (more than) 20 full-year equivalent employees during year t. Growing (shrinking) firms are defined with respect to their employment size in $t-1$. Notice that total number of firms from (2) to (5) is larger than in (1) because the same firm can change its class over time. **Panel A**: firm clustered standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05. **Panel B**: the prediction $\hat{w}_{i,t+2}$ is based on regression coefficients from column (1).

coefficient of interest for measuring complementarities in learning is α_2. Finally notice, by restricting our focus on workers going through unemployment in year $t+1$, we have purged out the effect of heterogeneity in outside options when workers were still employed in year t.

In table 3.1, panel A column (1), are reported the estimated coefficients for (3.19). The estimate of α_2 is positive and is about a tenth of α_1. These estimates imply that a 10% increase in coworkers average wage increases their salary in $t+2$ of 0.7%. In columns (2) to (5) the estimation is repeated by four different firm size and growth classes. Remarkably, the estimates of α_2 are similar across the different firm class samples.

In table 3.1, panel B, I use the estimated coefficients from column (1) to compute the average predicted wage growth from year t to $t+2$ by firm size class. The four means are statistically different from one another, and the highest predicted growth, about 8%, occurs for workers leaving small growing firms. Workers leaving bigger firms, regardless of the firm's growth status, only experience an average predicted growth of 4%.

To summarize, regression estimates for α_2 are similar across firm classes but predicted wage growth is not. For smaller firms, predicted wage growth nearly doubles the prediction of larger ones. I interpret this finding as suggestive that the learning function parameters are constant across firms (i.e. no decreasing or increasing returns to scale), yet the composition of the workforce in small growing firms is better suited to generate beneficial human capital spillovers. Despite these results, one should be cautious in making any causal statement at this stage because the selection of workers into firms and into coworkers is not random.

Complementarities in production - (3.20)					
	All firms	Small shrinking firms	Small growing firms	Big shrinking firms	Big growing firms
	(1)	(2)	(3)	(4)	(5)
Panel A					
$w_{i,t}$	-0.009***	-0.011***	-0.011***	-0.003	-0.01***
	(0.0009)	(0.0015)	(0.0012)	(0.0018)	(0.0019)
$w_{-i,t}$	0.007	0.009	0.011**	0.004	0.014
	(0.0039)	(0.0049)	(0.004)	(0.0084)	(0.009)
Controls	yes	yes	yes	yes	yes
R^2	0.004	0.004	0.003	0.013	0.01
# of th. obs.	2738	422	615	632	974
# of th. firms	42	30	34	4	6
Panel B					
mean($\hat{EE}_{i,t+1}$)		0.014	0.015	0.013	0.014
		(0.0036)	(0.0036)	(0.0041)	(0.0041)

Table 3.2: Estimation of specification (3.20) and average prediction of $\hat{EE}_{i,t+1}$ by firm class. Column (1) includes all firms while columns (2) to (5) only include firms by their class. Small (large) firms employ at most (more than) 20 full-year equivalent employees during year t. Growing (shrinking) firms are defined with respect to their employment size in $t-1$. Notice that total number of firms from (2) to (5) is larger than in (1) because the same firm can change its class over time. **Panel A**: firm clustered standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05. **Panel B**: the prediction $\hat{EE}_{i,t+1}$ is based on regression coefficients from column (1).

Complementarities in production For this regression specification I only consider workers continuously employed at same employer f throughout year t, and continuously employed in year $t+1$ as well. The outcome variable I consider is $EE_{i,t+1}$ which is a dummy taking value 1 if worker i changes employer by the first quarter of year $t+1$. For this sample of workers, I consider the following specification:

$$EE_{i,t+1} = \beta_0 + \beta_1 w_{i,t} + \beta_2 |w_{i,t} - w_{-i,t}| + BX_{i,t} + \varepsilon_{i,t}, \tag{3.20}$$

where $w_{i,t}$ is worker i wage and $w_{-i,t}$ is the average coworkers' wage in year t. Control variables in $X_{i,t}$ are year, worker gender, age, age squared, nationality, tenure in the firm, occupation, firm age, and a 10 groups k-means dummy based on firm wage mean and standard deviation.

The term $|w_{i,t} - w_{-i,t}|$ is the absolute distance between average pay in the firm and individual's wage. Supposing wages are proxy for human capital, then $|w_{i,t} - w_{-i,t}|$ is a proxy for human capital differences. The closer the value is to zero, the more likely individual i has an average human capital level in the firm. If the production function favors similarities, in particular $\rho > 1$, than workers that differ the most from the average human capital level will be more likely to leave the firm.

We test for this hypothesis by estimating (3.20), results are reported in table 3.2. The point estimate for β_2 is positive, yet not significant. Notice that when we separately estimate regression 3.20 by firm class, point estimates do not vary significantly. Differently from specification (3.19), the predicted value for $EE_{i,t+1}$ also does not vary across firms' classes. Overall, I interpret these estimates as evidence for only mild complementarities in

	Parameter	Description	Value
Set exogenously:			
	H	# HC types	2
	β	discount factor	0.988
	A	Aggregate productivity	1
	α	DRS	0.65
	ρ_z	Productivity persistence	0.99
	σ_ε	Productivity innovation s.d.	0.22155
	γ	Matching function elasticity	1.599
	s_0	Survival probability	0.99
Estimate			
	ρ	Substitution across HC types	2.5
	Δ	HC difference	0.4
	ϕ_0	Individual learning	0.025
	ϕ_1	Coworker learning base	0.0083
	ϕ_u	Unemployed HC depreciation	0.125
	δ_0	Exogenous exit probability	0.015
	k_f	Operational cost	5
	k_{e0}	Entry cost baseline	7.105
	b_0	Home production baseline	0.14
	c_o	Vacancy creation baseline cost	0.358
	λ	employed search efficiency	0.55
	$\bar{\pi}_0$	average HC at birth	1.1

Table 3.3: Free parameters of the model and their values.

production.

3.3.3 Estimation strategy and Equilibrium Properties

A full-fledged estimation of the model via the simulated method of moments requires further work. On this regard, the current simulated economy is for illustrative purposes and works as a proof of concept about the mechanisms and channels that the model is able to produce. Below I discuss which are the salient moments from the data that need to be matched by the model. Further, I present a preliminary simulation informed on some of these moments.

Table 3.3 summarizes the free parameters of the model and distinguishes between those internally estimated and externally set. The frequency of the model is quarterly and it is compared to VWH data on time-span 1982-2001. The discount factor is set to .988 or about 5% yearly. I set the decreasing returns to scale parameter α in between of Schaal (2017) that uses .85 for the US economy and Cooper et al. (2022) that estimate a value of .5122 using indirect inference on Italian data and assuming a production function that takes labor as the only factor of production. Cooper et al. (2022) also provide estimates for the yearly auto-regressive process of firm productivity, I convert their estimates to a quarterly frequency.

The strategy to estimate the labor market parameters γ, b_0, c_0 and λ follows Schaal (2017) in matching UE, EU and EE flows. The estimation of the exogenous exit probability δ_0 relies on matching average firm age in the data and the hazard rate estimate from an exponential survival distribution. There is no clear benchmark to set the values for the

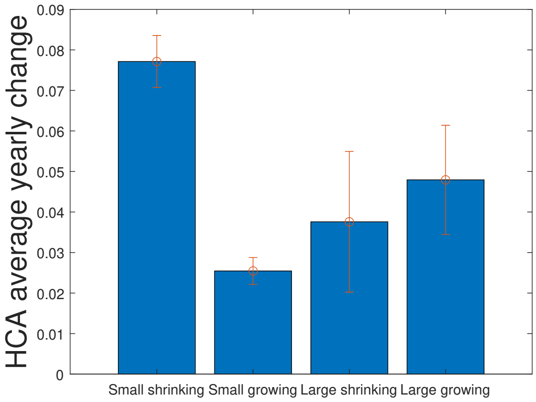


Figure 3.2: Predicted human capital change after full year t continuously employed in same firm f and conditional on remaining in f during all $t+1$. Predictions by firm size and growth class. Large firms employ an yearly average of at least 20 workers in t, otherwise they are small. Growing (shrinking) firms are firms with a larger (smaller) employment size in t than in $t-1$.

entry cost k_e and operational cost k_f. However, the higher is the cost of k_e the smaller is the mass of entrant firms so one can discipline k_e by matching the fraction of hiring done by entrant firms over total hiring in the economy. The operational cost k_f must be paid every period by active firms. The higher is k_f the larger firms must be to meet this recurrent cost. Consequently, k_f is related to matching the average firm size from the data.

Human capital parameters are $H, \Delta, \phi_0, \phi_1, \phi_u, \bar{\pi}_0$ and ρ. The parameters H and Δ govern the overall amount of human capital heterogeneity present in the economy. Since H determines the number of components in vector $\vec{n}$, it is not feasible to let this parameter to be estimated internally but it is fixed ex-ante. Herkenhoff et al. (2018) set $H = 7$ but keep maximum firm size to two workers. In the simulated economy I present in here, I set $H = 2$ and leave firm size unrestricted.[47] The value of Δ is chosen so to match the ratio between top and bottom 20th percentile of the wage distribution. Parameter ϕ_0 captures learning-by-doing and consequently is associated to experience. Therefore, the value of this parameter is informed by the average wage growth for employed workers that do not change their employer. The average human capital type when entering the labor market, $\bar{\pi}_0$ is disciplined by looking at the ratio between average wage at 54 and at 24, assuming workers enter the labor market at 21. The rate of depreciation of human capital, ϕ_u is adjusted so to match the correlation in the data between average unemployment duration and first wage upon leaving unemployment. Finally, the parameters ϕ_1 and ρ are related to the degree of complementarity among workers in learning and production. The estimated coefficients α_2 and β_2 from regression (3.19) and (3.20) and the related predictions by firm size and growth class are informative for these parameters. In addition, the share of wage variance between and within firms is also a relevant moment in measuring ρ. The higher is ρ, the more likely similar workers are employed together, therefore, the between component of wage dispersion is higher (Herkenhoff et al., 2018).

In table 3.4, I report the data and model generated moments. Despite the estimation strategy outlined above, the model simulated model still performs relatively poor. There

[47]Operationally, firms can only choose up to $N = 30$ workers for each human capital type and so maximum firm size is $N * H = 60$. However, because of decreasing returns to labor the upper point N is not binding in equilibrium.

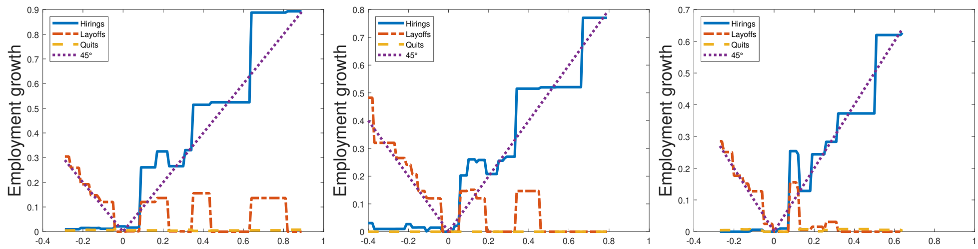
Moment name	Data value	Model value
Within / between firm wage dispersion	0.298	0.730
p80/p20 wage ratio	1.649	1.131
Mean wage at 54 / Mean wage at 24	1.375	1.022
Mean annual wage growth	0.057	0.003
Annual wage growth s.d.	0.038	0.041
α_2 in (3.19)	0.070	0.845
β_2 in (3.20)	0.007	-0.021
corr.(unemp. duration, earning at UE)	-0.033	-0.047
Firm average size	13.604	2.887
Firm average growth rate	0.071	0.014
Firm average age	10.060	8.219
Hazard rate estimate	0.064	0.034
Share of recruiting by entrant firms	0.155	0.452
EU rate	0.048	0.048
EE rate	0.029	0.005

Table 3.4: Matching moment from the data and the model. As the economy is stationary the unemployment inflows and outflows are the same: the data moment is the average between EU and UE.

is excessive wage compression in the simulated data and most of wage dispersion, contrary to the data, occurs between firms. Also the age-wage profile is too flat compared to the data, with the exception of the standard deviation in wage growth rate across workers. Unfortunately, the estimates for α_2 and β_2 are also off. Instead of directly comparing the predicted outcome variable of regression (3.19) and (3.20) against their model counterpart, I plotted in figure 3.2 the expected change in human capital for workers continuously employed over two years at the same firm, by firm size and growth class. The values in the graph are off compared to the regression analysis of the previous section, because small growing firms are reported as the least effective category for human capital accumulation. Nonetheless, this graph is an instance for the kind of analysis that the model allows. Finally, the correlation between unemployment duration and first wage when back to employment is relatively good.

Moments related to firms are also not well-aligned to the data. While the average age and hazard rate estimate are not too unrealistic, the average firm size and growth rate are about one magnitude too small. The difficulty in raising average firm size and increase dynamism in growth rates originates from the firm productivity process. The AR(1) process I am using is persistent and innovations are relatively big. These two properties, combined with the fact that firm productivity enters exponentially in the production function, tend to produce an excessive number of small firms compared to the data. At the same time, since the productivity process is so persistent, few firms experience growth and when they do they change their size very drastically. As a corollary of the lack of dynamism of incumbent firms, the share of recruiting done by entrant firms is excessively high.

While the model does a good job in matching flows between employment and unemployment, the simulated EE rate is about five times smaller than observed in the data. Beyond not being a good match with the data counterpart, a small fraction of EE flows helps explaining why the wage dispersion within firms is so small compared to data. To see this, recall that wage determination in the model follows equation (3.14) and it is based on



(a) All worker flows. (b) Only flows for $h = 1$. (c) Only flows for $h = 2$.

Figure 3.3: Hirings, layoffs and quits by yearly percentage of firm growth (x-axis).

workers' promised utility levels and corresponding threats to change job. Workers performing job-to-job transition search on higher promised-utility submarkerts than unemployed workers of the same human capital level. An higher submarket also generally translates into an "hiring bonus" when the worker joins a new firm from employment. More EE transitions increase the frequency of these bonuses and therefore increase the share of wage dispersion attributable to the within firm component.

Finally, the model generates a labor share of approximately .48, arguably a bit too low. To get a sense of the magnitudes of k_f and k_e, we can compare them to the value of production. In particular the operational costs weight for about 18% of the average revenues generated by active firms in the economy. On the other hand, the entry cost is about 8 times a average firm's first year revenue.

In figure 3.3, I report the inflow and outflow of workers by firm growth level. This is a typical exercise in firm dynamics (Kaas and Kircher, 2015; Schaal, 2017) as it shows what is the average gross composition of hires, quits and layoff that are associated to a certain net job creation/destruction level (Davis et al., 2006). There is a reason in showing this figure even though the model has not been properly estimated and simulation is coarse. Since in the model are present different human capital levels, we can condition the inflows and outflows by them. As firms grow, the composition of their human capital demand also changes. For instance, in the simulated model when firms grow they virtually never fire $h = 2$ workers but recur to quits sometimes. However, layoffs of $h = 1$ workers are as frequent in shrinking firms as in growing ones.

Finally, as discussed in the theoretical wage determination section, there are compensating differentials. Since high human capital individuals are better-off in this economy, workers are willing to give up part of their earnings and be compensated in learning opportunities (Jarosch et al., 2021). Accordingly, for low human capital workers, plot in figure 3.4 shows that the probability of becoming high types is inversely correlated with their wages.

3.4 Counterfactuals

The purpose of this section is to develope counterfactuals to compare against the estimated model. Of particular interest in the comparison are the aggregate level of human capital and the dynamics of its accumulation across firms and workers. Three counterfactuals are of interest for this analysis: *i)* a rise in the operational cost k_f, *ii)* an increased firm productivity persistence ρ_z, and *iii)* a fall in the dispersion of productivity innovations σ_ε. Unfortunately, without a complete estimation of the model, this analysis cannot deliver

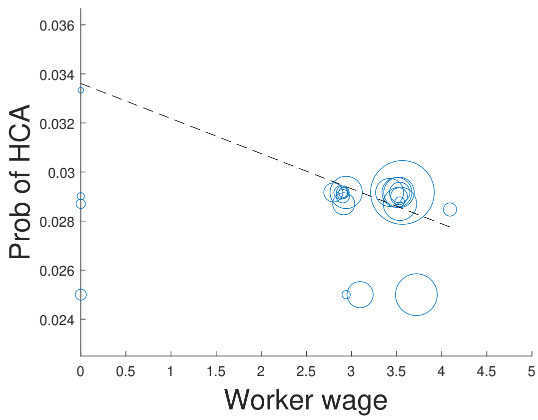


Figure 3.4: Evidence for compensating differentials. Scatter plot with regression line for sample of workers with human capital $h = 1$. On the y-axis is reported the probability the worker is going to turn into $h = 2$ human capital level by the next quarter.

quantitative insights. However, I still discuss below some of the expected implications of these changes.

A rise in k_f implies fewer but larger firms in equilibrium. As a consequence, workers stay longer in their employment and have more time to accumulate human capital. At the same time, firms' size become stickier and labor reallocation diminishes. This implies less chances for unemployed workers to return to employment and accumulate human capital. It is a quantitative question whether in the end aggregate human capital will be higher in this counterfactual economy than in the baseline one.

When firm productivity persistence is higher, firms' optimal size become more predictable over longer time horizons. Consequently, firms' option value of building their workforce gradually reduces and they tend to reach more quickly their optimal size (Schaal, 2017). Thus, the trade-off between raising human capital in house and recruiting on the market is affected. On one side, a more predictable size implies firms can better plan their future human capital needs and assemble current teams in that view. On the other side, if labor market frictions are not particularly high, firms will find advantageous of neglecting the in-house human capital accumulation process and jump faster to their optimal size and composition via recruiting on the market. Again, which of the two forces will prevail in equilibrium is a quantitative matter.

Finally, when σ_ε reduces there is less uncertainty about future productivity shocks and the same arguments from the previous paragraph about increased persistence apply. In addition, the firm size distribution will be less dispersed and workforce composition demand will be more similar across firms. As a result of a more similar workforce across firms, the expectation is that aggregate human capital becomes lower. Intuitively, when workers upgrade their human capital they are more likely to change employer. However, a less diverse firms' landscape slows down job-to-job dynamism and ultimately reduces the incentive for workers of accumulating human capital. Notice that the effects of reduced dynamism will be partially mitigated in equilibrium as compensating differentials (like those in figure 3.4) also become flatter.

3.5 Conclusions

In this work, I propose a novel link between coworker learning and firm dynamics. To study coworker learning, one needs to account for endogenous selection of workers into coworkers' characteristics and I do so by developing a theory of firm dynamic labor demand. I argue that this theory is better suited in accounting for endogenous selection into coworkers than previous structural work that only allows for workers to match into pairs (Herkenhoff et al., 2018). By letting workers form larger teams, the model avoids to impose a mechanical trade-off between producing with similar individuals or fostering future human capital accumulation.

A second reason to link coworker learning with firm dynamics comes from the documented heterogeneity in predicted coworker spillovers by firm size and growth class. Not only firm dynamics helps to account for endogenous workforce selection, but in doing so it also highlights the relevance of firm heterogeneity in providing learning opportunities.

I acknowledge that the observations made so far have an important caveat. Since the model is not properly estimated, the magnitudes of my analysis are likely different from those in the data. A more serious estimation exercise via simulated method of moments is required to gain more confidence in quantifying the various aspects of the model. Moreover, a full-fledged estimation would also allow to quantitatively assess the counterfactual predictions from section 3.4 and, eventually, derive policy recommendations to foster human capital accumulation in the economy.

Bibliography

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

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

Acemoglu, D., and Restrepo, P. (2022) "Tasks, Automation, and the Rise in U.S. Wage Inequality." *Econometrica*, 90(5): 1973–2016.

Akcigit, U., Caicedo, S., Miguelez, E., Stantcheva, S., and Sterzi, V. (2018) "Dancing with the Stars: Innovation Through Interactions."Technical Report w24466, National Bureau of Economic Research, Cambridge, MA.

Alon, T., Doepke, M., Olmstead-Rumsey, J., and Tertilt, M. (2020) "The Impact of COVID-19 on Gender Equality." *Covid Economics: Vetted and Real-Time Papers*(4): 62–85.

Anderson, A. (2015) "A dynamic generalization of Becker's assortative matching result." *Journal of Economic Theory* 290–310.

Arellano-Bover, J., and Saltiel, F. (2022) "Differences in On-the-Job Learning Across Firms." June, working paper.

Bagger, J., Fontaine, F., Galenianos, M., and Trapeznikova, I. (2020) "Vacancies, Employment Outcomes and Firm Growth: Evidence from Denmark." IZA Discussion Paper 14436, IZA.

Bartolucci, C., Devicienti, F., and Monzón, I. (2018) "Identifying Sorting in Practice." *American Economic Journal: Applied Economics*, 10(4): 408–438.

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

Becker, G. S. (1957) *The Economics of Discrimination*. Economic Research Studies, Chicagi: University of Chicago Press.

Benhabib, J., Perla, J., and Tonetti, C. (2021) "Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier." *Econometrica*, 89(5): 2261–2301.

Bhalotra, S., Fernandez Sierra, M., and Wang, F. (2022) "The Distribution of the Gender Wage Gap: An Equilibrium Model." IZA Discussion Paper 15258, Institute of Labor Economics (IZA).

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

Blau, F. D., and Kahn, L. M. (2017) "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature*, 55(3): 789–865.

Briskar, J., Di Porto, E., Rodriguez Mora, J. V., and Tealdi, C. (2022) "Decomposition of Italian Inequality." working paper 53, INPS.

Caicedo, S., Lucas Jr., R. E., and Rossi-Hansberg, E. (2019) "Learning, Career Paths, and the Distribution of Wages." *American Economic Journal: Macroeconomics*, 11(1): 49–88.

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

Card, D., Devicienti, F., and Maida, A. (2014) "Rent-sharing, Holdup, and Wages: Evidence from Matched Panel Data." *The Review of Economic Studies*, 81(1): 84–111.

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

Carrillo-Tudela, C., Gartner, H., and Kaas, L. (2020) "Recruitment Policies, Job-Filling Rates and Matching Efficiency." Discussion Paper 13240, IZA.

Casarico, A., and Lattanzio, S. (2022) "What Firms Do: Gender Inequality in Linked Employer-Employee Data."Technical Report 24, INPS.

Contini, B., and Trivellato, U. eds. (2005) *Eppur si muove: dinamiche e persistenze nel mercato del lavoro italiano*. Percorsi, Bologna: Il mulino.

Cooper, R., Horn, C.-W., and Indraccolo, L. (2022) "Covid and Productivity in Europe: A Responsiveness Perspective." mimeo, EUI, mimeo.

Costinot, A., and Vogel, J. (2010) "Matching and Inequality in the World Economy." *Journal of Political Economy*, 118(4): , p. 40.

Cuberes, D., Mora, J. V. R., Teignier, M., and Visschers, L. (2021) "Family Labor Market Decisions and Statistical Gender Discrimination." Mimeo.

Davis, S. J., Faberman, R. J., and Haltiwanger, J. (2006) "The Flow Approach to Labor Markets: New Data Sources and Micro-Macro Links." *Journal of Economic Perspectives*, 20(3): 3–26.

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

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

Eeckhout, J., Kircher, P., Lafuente, C., and Macci, G. (2021) "Corrigendum to Capital Investment in "Assortative Matching With Large Firms"." *Econometrica*, 89(4): o11–o14.

Elsby, M. W. L., Gottfries, A., Michaels, R., and Ratner, D. (2021) "Vacancy Chains." October, working paper.

Freund, L. B. (2022) "Superstar Teams: The Micro Origins and Macro Implications of Coworker Complementarities." December, working paper.

Garicano, L. (2000) "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economy*, 108(5): , p. 31.

Gavazza, A., Mongey, S., and Violante, G. L. (2018) "Aggregate Recruiting Intensity." *American Economic Review*, 108(8): 2088–2127.

Gregory, V. (2021) "Firms as Learning Environments: Implications for Earnings Dynamics and Job Search." Working Paper 2020-036B, Federal Reserve Bank of St. Louis.

Grinza, E. (2021) "Worker flows, reallocation dynamics, and firm productivity: new evidence from longitudinal matched employer–employee data." *Industrial and Corporate Change*, 30(1): 75–108.

Grossman, G. M., Helpman, E., and Kircher, P. (2017) "Matching, Sorting, and the Distributional Effects of International Trade." *Journal of Political Economy*, 125(1): 224–264.

Gulyas, A. (2020) "Firm Dynamics with Labor Market Sorting." Discussion Paper 181, University of Bonn and University of Mannheim.

Haltiwanger, J. C., Hyatt, H. R., and Spletzer, J. (2022) "Industries, Mega Firms, and Increasing Inequality."Technical Report 29920, National Bureau of Economic Research, Inc, Publication Title: NBER Working Papers.

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

Hong, L. (2022) "Coworker Sorting, Learning, and Inequality." December, working paper.

Hong, L., and Lattanzio, S. (2022) "The Peer Effect on Future Wages in the Workplace." March, working paper.

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

Ichino, A., and Maggi, G. (2000) "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm." *The Quarterly Journal of Economics*, 115(3): 1057–1090.

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

Jovanovic, B. (2014) "Misallocation and Growth." *American Economic Review*, 104(4): 1149–1171.

Kaas, L. (2020) "Block-Recursive Equilibria in Heterogenous-Agent Models."Technical Report 8737, CESifo, Publication Title: CESifo Working Paper Series.

Kaas, L., and Kircher, P. (2015) "Efficient Firm Dynamics in a Frictional Labor Market." *American Economic Review*, 105(10): 3030–3060.

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

Lee, E. (2021) "Different Learning Opportunities across Firms: Does Knowledge Come from Firms or Coworkers?" September, working paper.

Lucas, R. E. (2009) "Ideas and Growth." *Economica*, 76(301): 1–19.

Lucas, R. E., and Moll, B. (2014) "Knowledge Growth and the Allocation of Time." *Journal of Political Economy*, 122(1): 1 – 51.

Luttmer, E. G. J. (2015) "An Assignment Model of Knowledge Diffusion and Income Inequality." Research Department Staff Report 509, FRB Minneapolis.

Ma, X., Nakab, A., and Vidart, D. (2023) "How do Workers Learn? Theory and Evidence on the Roots of Lifecycle Human Capital Accumulation." January, working paper.

Marjit, S., and Oladi, R. (2022) "Gender Discrimination in Competitive Markets." CESifo Working Paper Series 9705, CESifo.

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

Menzio, G., and Shi, S. (2010) "Block recursive equilibria for stochastic models of search on the job." *Journal of Economic Theory*, 145(4): 1453–1494.

Moen, E. R., and Rosén, A. (2004) "Does Poaching Distort Training?" *Review of Economic Studies*, 71(4): 1143–1162.

Nix, E. (2020) "Learning Spillovers in the Firm." Working Paper 14, IFAU.

Parrotta, P., and Pozzoli, D. (2012) "The effect of learning by hiring on productivity." *The RAND Journal of Economics*, 43(1): 167–185.

Perla, J., and Tonetti, C. (2014) "Equilibrium Imitation and Growth." *Journal of Political Economy*, 122(1): 52–76.

Poggi, A., and Natale, P. (2020) "Learning by hiring, network centrality and within-firm wage dispersion." *Labour Economics*, 67, p. 101922.

Rubery, J., and Koukiadaki, A. (2016) "Closing the Gender Pay Gap: A Review of the Issues, Policy Mechanisms and International Evidence."Technical report, ILO.

Sampson, T. (2016) "Assignment reversals: Trade, skill allocation and wage inequality." *Journal of Economic Theory*, 163 365–409.

Schaal, E. (2017) "Uncertainty and Unemployment." *Econometrica*, 85(6): 1675–1721.

Serafinelli, M. (2019) ""Good" Firms, Worker Flows, and Local Productivity." *Journal of Labor Economics*, 37(3): 747–792.

Shi, S. (2018) "Efficient Investments, Job Dynamics and Frictional Dispersion." May.

Skans, O. N., Edin, P.-A., and Holmlund, B. (2007) "Wage dispersion between and within plants: Sweden 1985-2000."Technical Report w13021, National Bureau of Economic Research, Cambridge, MA.

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

Sorkin, I. (2017) "The Role of Firms in Gender Earnings Inequality: Evidence from the United States." *American Economic Review*, 107(5): 384–387.

Sánchez, R., Finot, J., and Villena, M. G. (2020) "Gender Wage Gap and Firm Market Power: Evidence from Chile." Working Paper 13856, IZA.

Tattara, G., and Valentini, M. (2007) "The Cyclical Behaviour of Job and Worker Flows."Technical Report 016, University Ca' Foscari, Venice.

Appendix

A Chapter 1 appendix

A.1 PAM & NAM

If not interested in the derivation of the general formulation, directly jump to equation (A.11) to see the specific formulation for the model in section 1.2.

Conditions for PAM with uni-dimensional capital investment Consider the production function:

$$\tilde{F}(x, y, u, s) = \max_k F(x, y, u, s, k) - ik \tag{A.1}$$

and assume that F is twice differentiable; convex in u, s, and k; and displays CRS in u, s, and k (so that $\tilde{F}$ has CRS in s and u). Problem (A.1) can be rewritten by explicitly solving the maximization problem, that is:

$$\tilde{F}(x, y, u, s) = F(x, y, u, s, k^*) - ik^* \tag{A.2}$$

where k^* depends on (x, y, u, s) and solves the FOC[48]:

$$F_k(x, y, u, s, k^*(x, y, u, s)) = i \tag{A.3}$$

Calculating the gradient of k^* will be useful for further simplifications below, by applying the implicit function theorem, we have:

$$\left[\frac{\partial k^*}{\partial x}, \frac{\partial k^*}{\partial y}, \frac{\partial k^*}{\partial u}, \frac{\partial k^*}{\partial s}\right] = \left[-\frac{F_{xk}}{F_{kk}}, -\frac{F_{yk}}{F_{kk}}, -\frac{F_{uk}}{F_{kk}}, -\frac{F_{sk}}{F_{kk}}\right] \tag{A.4}$$

We know from Eeckhout and Kircher (2018) that PAM arises if and only if:

$$\tilde{F}_{xy}\tilde{F}_{us} > \tilde{F}_{xs}\tilde{F}_{yu} \tag{A.5}$$

[48]The SOC is automatically fulfilled when assuming F in concave in k, *i.e.* $F_{kk} < 0$; where the lower index denotes a derivative, for instance $F_k = \frac{\partial F}{\partial k}$.

using the formulation in (A.2), the second derivatives in (A.5) can be rewritten as:

$$[F_{xy}, F_{us}, F_{xs}, F_{yu}] = \left[F_{xy} + F_{xk}\frac{\partial k^*}{\partial y}, F_{us} + F_{uk}\frac{\partial k^*}{\partial s}, F_{xs} + F_{sk}\frac{\partial k^*}{\partial x}, F_{yu} + F_{yk}\frac{\partial k^*}{\partial u}\right] \tag{A.6}$$

substituting (A.4) into (A.6) and then replacing the terms in (A.5), the PAM condition becomes

$$\left(F_{xy} - \frac{F_{xk}F_{yk}}{F_{kk}}\right)\left(F_{us} - \frac{F_{uk}F_{sk}}{F_{kk}}\right) > \left(F_{xs} - \frac{F_{xk}F_{sk}}{F_{kk}}\right)\left(F_{yu} - \frac{F_{yk}F_{uk}}{F_{kk}}\right) \tag{A.7}$$

Multiply both sides of (A.7) by F_{kk} (notice that one needs to revert the sign of the inequality because $F_{kk} < 0$) and simplify to get:

$$\begin{aligned} &F_{xy}F_{us}F_{kk} - F_{xy}F_{uk}F_{sk} - F_{xk}F_{yk}F_{us} \\ &\quad < F_{xs}F_{yu}F_{kk} - F_{xs}F_{yk}F_{uk} - F_{xk}F_{yu}F_{sk} \end{aligned} \tag{A.8}$$

Conditions for PAM with n-dimensional capital investment Following the same steps as before, one can show PAM will occur only if[49]

$$\begin{aligned} &\left(F_{xy} - \sum_{j=1}^{n}\frac{F_{xk_j}F_{yk_j}}{F_{k_jk_j}}\right)\left(F_{us} - \sum_{j=1}^{n}\frac{F_{uk_j}F_{sk_j}}{F_{k_jk_j}}\right) > \\ &\quad \left(F_{xs} - \sum_{j=1}^{n}\frac{F_{xk_j}F_{sk_j}}{F_{k_jk_j}}\right)\left(F_{yu} - \sum_{j=1}^{n}\frac{F_{yk_j}F_{uk_j}}{F_{k_jk_j}}\right) \end{aligned} \tag{A.9}$$

or equivalently, after simplifying and multiplying both sides by $\prod_{j=1}^{n} F_{kk}$:

$$\begin{aligned} &F_{xy}F_{us}\left(\prod_{j=1}^{n} F_{k_jk_j}\right) - F_{xy}\left[\sum_{j=1}^{n}\left(\prod_{m\neq j} F_{k_mk_m}\right)F_{uk_j}F_{sk_j}\right] \\ &\quad - F_{us}\left[\sum_{j=1}^{n}\left(\prod_{m\neq j} F_{k_mk_m}\right)F_{xk_j}F_{yk_j}\right] \\ &\quad + \sum_{j=1}^{n}\left[\sum_{m\neq j}\left[\left(\prod_{p\neq j,m} F_{k_pk_p}\right)F_{uk_m}F_{sk_m}\right]F_{xk_j}F_{yk_j}\right] \gtrless \\ &F_{xs}F_{yu}\left(\prod_{j=1}^{n} F_{k_jk_j}\right) - F_{xs}\left[\sum_{j=1}^{n}\left(\prod_{m\neq j} F_{k_mk_m}\right)F_{yk_j}F_{uk_j}\right] \\ &\quad - F_{yu}\left[\sum_{j=1}^{n}\left(\prod_{m\neq j} F_{k_mk_m}\right)F_{xk_j}F_{sk_j}\right] \\ &\quad + \sum_{j=1}^{n}\left[\sum_{m\neq j}\left[\left(\prod_{p\neq j,m} F_{k_pk_p}\right)F_{yk_m}F_{uk_m}\right]F_{xk_j}F_{sk_j}\right] \end{aligned} \tag{A.10}$$

[49] Under the assumption that the $n \times n$ matrix of second derivatives with respect to $k_1, \ldots, k_n$ of $F(x, y, u, s, k_1, \ldots, k_n)$ is negative definite.

where the inequality sign $\gtrless$ is "smaller than" if $\prod_{j=1}^{n} F_{kk} < 0$, and "greater than" if $\prod_{j=1}^{n} F_{kk} > 0$.

Conditions for PAM in the model of section 1.2 Using the same notation as in section 1.2, formulation (A.10) in case of PAM becomes:

$$F_{xy}F_{bw}F_{kk}F_{ee} - F_{xy}F_{bk}F_{wk}F_{ee} - F_{xy}F_{be}F_{we}F_{kk} - F_{xk}F_{yk}F_{bw}F_{ee} - F_{xe}F_{ye}F_{bw}F_{kk} > F_{xw}F_{yb}F_{kk}F_{ee} - F_{xw}F_{yk}F_{bk}F_{ee} - F_{xw}F_{ye}F_{be}F_{kk} - F_{xk}F_{yb}F_{wk}F_{ee} - F_{xe}F_{yb}F_{we}F_{kk} \tag{A.11}$$

and will have the reversed sign for the case of NAM.[50]

A.2 Other results on Wage Dispersion

Let us compare economy A with economy B. Let us resume with the case described in main text section 1.2.3 in assuming the only difference between the economies is $i_{q,B} < i_{e,A}$. This implies the *direct effect* on total wage dispersion is absent. We briefly discuss how to relax this at the end of this appendix. Let us now focus on providing analytical expressions for the *compositional effect* and for the two types of indirect effects, namely *worker-capital complementarities* and *worker-teammates complementarity*. Each of these effects will be describe below. Please refer to table 1.1 as a reference scheme.

The variation in wage dispersion from economy A to economy B is $\sigma_B^2 - \sigma_A^2$ which can be further decomposed into changes in dispersion within and between teams.

$$\sigma_B^2 - \sigma_A^2 = W_B - W_A + B_B - B_A. \tag{A.12}$$

1. Within component decomposition Exploiting the decomposition in (1.10), we can write:

$$W_B - W_A = \int_{\underline{x}}^{\bar{x}} \sigma_B^2(x) - \sigma_A^2(x) \mathrm{d}\mathcal{SW}(\mu(x)) = \int_{\underline{x}}^{\bar{x}} W_{BA}(x) \mathrm{d}\mathcal{SW}(\mu(x)) \tag{A.13}$$

where $W_{BA}(x)$ represents the change in within team inequality for a blue-collar of type x as defined in (1.7). Clearly, the team composition of an x-type in economy A and B does not need to be the same . Denote with $\{\mu_i(x), \theta_i(x), k_i(x), q_i(x)\}$ the team of x in economy $i \in \{A, B\}$. Similarly, $\omega_i^u(x)$ and $\omega_i^s(x)$ denote the wage of the blue-collar and his white-collar teammate in the two economies. Let us define $x' = \mu_B^{-1}(\mu_A(x))$ as the blue-collar worker from economy B who is teamed up with $y = \mu_A(x)$ in economy A, and $x'' = \mu_A^{-1}(\mu_B(x))$ as the blue-collar worker from economy B who is teamed up with

[50] Notice that the last term on both the LHS and RHS of (A.10) does not appear when there are less than three capital inputs.

$y' = \mu_B(x)$ in economy A. We can write:[51]

$$
\begin{aligned}
W_{BA}(x) = &\left[\omega^s_B(x) - \omega^u_B(x)\right]^2 \frac{\theta_B(x)}{\theta_B(x)+1} - \left[\omega^s_A(x) - \omega^u_A(x)\right]^2 \frac{\theta_A(x)}{\theta_A(x)+1} \\
= &\left[\left[\omega^s_B(x) - \omega^u_B(x)\right]^2 \frac{\theta_B(x)}{\theta_B(x)+1} - \left[\omega^s_B(x') - \omega^u_B(x)\right]^2 \frac{\theta_A(x)}{\theta_A(x)+1}\right]/2 \quad \Big\} \text{ variation due to changing allocation (at economy B's wages)} \\
&+ \left[\left[\omega^s_A(x'') - \omega^u_A(x)\right]^2 \frac{\theta_B(x)}{\theta_B(x)+1} - \left[\omega^s_A(x) - \omega^u_A(x)\right]^2 \frac{\theta_A(x)}{\theta_A(x)+1}\right]/2 \quad \Big\} \text{ variation due to changing allocation (at economy A's wages)} \\
&+ \left[\left[\omega^s_B(x) - \omega^u_B(x)\right]^2 \frac{\theta_B(x)}{\theta_B(x)+1} - \left[\omega^s_A(x'') - \omega^u_A(x)\right]^2 \frac{\theta_B(x)}{\theta_B(x)+1}\right]/2 \quad \Big\} \text{ variation due to changing wages (at economy B's allocation)} \\
&+ \left[\left[\omega^s_B(x') - \omega^u_B(x)\right]^2 \frac{\theta_A(x)}{\theta_A(x)+1} - \left[\omega^s_A(x) - \omega^u_A(x)\right]^2 \frac{\theta_A(x)}{\theta_A(x)+1}\right]/2, \quad \Big\} \text{ variation due to changing wages (at economy A's allocation)}
\end{aligned}
\tag{A.14}
$$

where the first two lines are changes attributable to team allocation (i.e. *compositional effect*) and the last two lines are changes attributable to wages (i.e. *indirect effect*, mixed for capital deepening and teammates complementarities). In particular, the compositional effect is the average between the effect of changes in team allocation at wages of economy B (first line) and wages in economy A (second line). Notice that if a particular x is assigned to the exact same $\mu(x)$ and $\theta(x)$ in both economies, then the first two lines will be equal to zero. The indirect effect is the average of changes in wage dispersion at economy B team allocation (third line), and economy A (last line). To prove this decomposition —and for all the next ones others— is sufficient to add and subtract terms with the same color.

2. Between component decomposition Exploiting the decomposition in (1.10), we can write:

$$
B_B - B_A = \int_{\underline{x}}^{\bar{x}} \left[\omega_B(x) - \omega_B\right]^2 (\theta_B(x)+1) - \left[\omega_A(x) - \omega_A\right]^2 (\theta_A(x)+1) \mathrm{d}\mathcal{SW}(\mu(x)) = \int_{\underline{x}}^{\bar{x}} B_{BA}(x) \mathrm{d}\mathcal{SW}(\mu(x)) \tag{A.15}
$$

where ω_i and $\omega_i(x)$ are the economy-wide and x-team average wage in economy $i = A, B$ as defined in (1.6) and (1.8), respectively. Let us define again $x' = \mu_B^{-1}(\mu_A(x))$ as the blue-collar worker from economy B who is teamed up with $y = \mu_A(x)$ in economy A, and $x'' = \mu_A^{-1}(\mu_B(x))$ as the blue-collar worker from economy B who is teamed up with $y' = \mu_B(x)$ in economy A. For instance, given x, we have $\omega^s_B(x') = f_s(x', \mu_A(x), \theta_B(x'), k_B(x'), q_B(x'))$.

[51]Denote with μ^{-1} the inverse of μ this inverse is a.e. defined (Eeckhout and Kircher, 2018).

We can write:

$$
\begin{aligned}
B_{BA}(x) &= \left[\frac{\theta_B(x)\omega_B^u(x)+\omega_B^s(x)}{\theta_B(x)+1}-\omega_B\right]^2(\theta_B(x)+1)-\left[\frac{\theta_A(x)\omega_A^u(x)+\omega_A^s(x)}{\theta_A(x)+1}-\omega_A\right]^2(\theta_A(x)+1)\\
&= \left[\left[\frac{\theta_B(x)\omega_B^u(x)+\omega_B^s(x)}{\theta_B(x)+1}-\omega_B\right]^2(\theta_B(x)+1)-\left[\frac{\theta_A(x)\omega_B^u(x)+\omega_B^s(x')}{\theta_A(x)+1}-\omega_B\right]^2(\theta_A(x)+1)\right]/2 \quad \left.\right\}\text{variation due to changing allocation (at economy B's wages)}\\
&+ \left[\left[\frac{\theta_B(x)\omega_A^u(x)+\omega_A^s(x'')}{\theta_B(x)+1}-\omega_A\right]^2(\theta_B(x)+1)-\left[\frac{\theta_A(x)\omega_A^u(x)+\omega_A^s(x)}{\theta_A(x)+1}-\omega_A\right]^2(\theta_A(x)+1)\right]/2 \left.\right\}\text{variation due to changing allocation (at economy A's wages)}\\
&+ \left[\left[\frac{\theta_B(x)\omega_B^u(x)+\omega_B^s(x)}{\theta_B(x)+1}-\omega_B\right]^2(\theta_B(x)+1)-\left[\frac{\theta_B(x)\omega_A^u(x)+\omega_A^s(x'')}{\theta_B(x)+1}-\omega_A\right]^2(\theta_B(x)+1)\right]/2 \left.\right\}\text{variation due to changing wages (at economy B's allocation)}\\
&+ \left[\left[\frac{\theta_A(x)\omega_B^u(x)+\omega_B^s(x')}{\theta_A(x)+1}-\omega_B\right]^2(\theta_A(x)+1)-\left[\frac{\theta_A(x)\omega_A^u(x)+\omega_A^s(x)}{\theta_A(x)+1}-\omega_A\right]^2(\theta_A(x)+1)\right]/2, \left.\right\}\text{variation due to changing wages (at economy B's allocation)}
\end{aligned}
\tag{A.16}
$$

where the first two lines are changes attributable to team allocation (i.e. *compositional effect*) and the last two lines are changes attributable to wages (i.e. *indirect effect*, mixed for capital deepening and teammates complementarities).

By definition, the sum of compositional effect within and between firms is zero. In formulas, integrating in x the first two components in (A.14) summed up with the first two lines in (A.16) gives zero. This is the case because just varying the composition of teams from economy A to B, holding fixed wages, does not affect total wage dispersion. Therefore, the variation due to changing allocations at constant wages only affects the relative contribution of between and within team dispersion, but not total dispersion. This is the *compositional effect*. We now turn our attention in further decomposing the indirect effect in the worker-capital complementarity and worker-teammates complementarity components.

3. Decompose the indirect effect for the *within* component Changes in wages reflect changes in the marginal product of workers. There is no change in the primitives of the production function, so the marginal product variation from economy A to B can be decomposed into two components only: the change in the complementarities with capital (how much k and q there are per worker) and the change in the complementarities with teammates (how many colleagues θ and of which quality μ there are).

To formally derive how to decompose the indirect effect, recall that wages in equilibrium are equal to marginal products, i.e. $w^j(x) = f_j(x, \mu(x), \theta(x), k(x), q(x))$ for $j = u, s$, respectively, for a blue and white-collar worker of type $\{x, \mu(x)\}$. Consider the third line in (A.14) and notice that the colored two terms only differ on the square difference of wages (it can be done an equivalent decomposition also for the fourth line of A.14). Thus, we can write:

$$
\begin{aligned}
&[\omega^s_B(x) - \omega^u_B]^2 - [\omega^s_A(x'') - \omega^u_A]^2 = \\
&\quad [f_s(x, \mu_B(x), \theta_B(x), k_B(x), q_B(x)) - f_u(x, \mu_B(x), \theta_B(x), k_B(x), q_B(x))]^2 \\
&\quad - [f_s(x'', \mu_B(x), \theta_A(x''), k_A(x''), e_A(x'')) - f_u(x, \mu_A(x), \theta_A(x), k_A(x), q_A(x))]^2 = \\
&\left.\begin{aligned}&[f_s(x, \mu_B(x), \theta_B(x), k_B(x), q_B(x)) - f_u(x, \mu_B(x), \theta_B(x), k_B(x), q_B(x))]^2 \\ &-[f_s(x, \mu_B(x), \theta_B(x), k_A(x), q_A(x))) - f_u(x, \mu_B(x), \theta_B(x), k_A(x), q_A(x))]^2\end{aligned}\middle/ 2\right\} \text{variation due to changing capital (at economy B"s teammates)} \\
&\left.\begin{aligned}&+[f_s(x'', \mu_B(x), \theta_A(x''), k_B(x''), e_B(x'')) - f_u(x, \mu_A(x), \theta_A(x), k_B(x), q_B(x))]^2 \\ &-[f_s(x'', \mu_B(x), \theta_A(x''), k_A(x''), e_A(x'')) - f_u(x, \mu_A(x), \theta_A(x), k_A(x), q_A(x))]^2\end{aligned}\middle/ 2\right\} \text{variation due to changing capital (at economy A"s teammates)} \\
&\left.\begin{aligned}&+[f_s(x, \mu_B(x), \theta_B(x), k_B(x), q_B(x)) - f_u(x, \mu_B(x), \theta_B(x), k_B(x), q_B(x))]^2 \\ &-[f_s(x'', \mu_B(x), \theta_A(x''), k_B(x''), e_B(x'')) - f_u(x, \mu_A(x), \theta_A(x), k_B(x), q_B(x))]^2\end{aligned}\middle/ 2\right\} \text{variation due to changing teammates (at economy B"s capital)} \\
&\left.\begin{aligned}&+[f_s(x, \mu_B(x), \theta_B(x), k_A(x), q_A(x))) - f_u(x, \mu_B(x), \theta_B(x), k_A(x), q_A(x))]^2 \\ &-[f_s(x'', \mu_B(x), \theta_A(x''), k_A(x''), e_A(x'')) - f_u(x, \mu_A(x), \theta_A(x), k_A(x), q_A(x))]^2\end{aligned}\middle/ 2\right\} \text{variation due to changing teammates (at economy A"s capital)}
\end{aligned}
\tag{A.17}
$$

The expression in (A.17) provides a way to decompose the difference from economy A to economy B of the squared wage difference between a white-collar $\mu_B(x)$ and a blue-collar x. This difference can partly be attributed to changes in the marginal returns to labor due to worker-capital complementarities (the first two blocks in A.17, *capital deepening complementarities*). For the remaining part, this difference is explained by changes in the marginal returns to labor due to worker-teammates complementarities (the last two block in A.17, *teammates complementarities*).

Let us compute the equivalent decomposition also for the last line in (A.14). In other words, let us provide a decomposition for the difference from economy A to economy B of the squared wage difference between a white-collar $\mu_A(x)$ and a blue-collar x. We have:

$$
\begin{aligned}
&[\omega^s_B(x') - \omega^u_B(x)]^2 - [\omega^s_A(x) - \omega^u_A(x)]^2 = \\
&\quad [f_s(x', \mu_A(x), \theta_B(x'), k_B(x'), e_B(x')) - f_u(x, \mu_B(x), \theta_B(x), k_B(x), q_B(x))]^2 \\
&\quad - [f_s(x, \mu_A(x), \theta_A(x), k_A(x), q_B(x)) - f_u(x, \mu_A(x), \theta_A(x), k_A(x), q_B(x))]^2 \\
&\left.\begin{aligned}&[f_s(x', \mu_A(x), \theta_B(x'), k_B(x'), q_B(x')) - f_u(x, \mu_B(x), \theta_B(x), k_B(x), q_B(x))]^2 \\ &-[f_s(x', \mu_A(x), \theta_B(x'), k_A(x'), q_A(x')) - f_u(x, \mu_B(x), \theta_B(x), k_A(x), q_A(x))]^2\end{aligned}\middle/ 2\right\} \text{variation due to changing capital (at economy B's teammates)} \\
&\left.\begin{aligned}&+[f_s(x, \mu_A(x), \theta_A(x), k_B(x), q_B(x)) - f_u(x, \mu_A(x), \theta_A(x), k_B(x), q_B(x))]^2 \\ &-[f_s(x, \mu_A(x), \theta_A(x), k_A(x), q_A(x)) - f_u(x, \mu_A(x), \theta_A(x), k_A(x), q_A(x))]^2\end{aligned}\middle/ 2\right\} \text{variation due to changing capital (at economy A's teammates)} \\
&\left.\begin{aligned}&+[f_s(x', \mu_A(x), \theta_B(x'), k_B(x'), q_B(x')) - f_u(x, \mu_B(x), \theta_B(x), k_B(x), q_B(x))]^2 \\ &-[f_s(x, \mu_A(x), \theta_A(x), k_B(x), q_B(x)) - f_u(x, \mu_A(x), \theta_A(x), k_B(x), q_B(x))]^2\end{aligned}\middle/ 2\right\} \text{variation due to changing teammates (at economy B's capital)} \\
&\left.\begin{aligned}&+[f_s(x', \mu_A(x), \theta_B(x'), k_A(x'), q_A(x')) - f_u(x, \mu_B(x), \theta_B(x), k_A(x), q_A(x))]^2 \\ &-[f_s(x, \mu_A(x), \theta_A(x), k_A(x), q_A(x)) - f_u(x, \mu_A(x), \theta_A(x), k_A(x), q_A(x))]^2\end{aligned}\middle/ 2\right\} \text{variation due to changing teammates (at economy A's capital)}
\end{aligned}
\tag{A.18}
$$

The expression in (A.18) provides a way to decompose the difference from economy A to economy B of the squared wage difference between a white-collar $\mu_A(x)$ and a blue-collar

x. This difference can partly be attributed to changes in the marginal returns to labor due to worker-capital complementarities (the first two blocks in A.18, *capital deepening complementarities*). For the remaining part, this difference is explained by changes in the marginal returns to labor due to worker-teammates complementarities (the last two block in A.18, *teammates complementarities*).

4. Decompose the indirect effect for the *between* component Just as in the previous paragraph, the marginal product change from economy A to B can be further decomposed into two components: the change in the complementarities with capital (how much k and q there are per worker) and the change in the complementarities with teammates (how many colleagues θ and of which quality μ there are).

To formally derive such a decomposition, define ω_{i,q_j,k_j} as the economy-wide average wage prevailing in an economy with allocation as in economy $i \in \{A, B\}$ but with capital investment decision as in economy $j \neq i$. Consider the third line in (A.16) and notice that the two terms only differ on the square difference of wages (it can be done an equivalent decomposition also for the fourth line of A.16). So we can write as follows:

$$
\begin{aligned}
&\left[\frac{\theta_B(x)\omega_B^u(x)+\omega_B^s(x)}{\theta_B(x)+1}-\omega_B\right]^2-\left[\frac{\theta_B(x)\omega_A^u(x)+\omega_A^s(x'')}{\theta_B(x)+1}-\omega_A\right]^2 \\
&=\left[\frac{\theta_B(x)f_b(x,\mu_B(x),\theta_B(x),k_B(x),q_B(x))+f_s(\mu_B(x),\theta_B(x),k_B(x),q_B(x))}{\theta_B(x)+1}-\omega_B\right]^2 \\
&\quad-\left[\frac{\theta_B(x)f_b(x,\mu_A(x),\theta_A(x),k_A(x),q_A(x))+f_s(x'',\mu_B(x),\theta_A(x''),k_A(x''),q_A(x''))}{\theta_B(x)+1}-\omega_A\right]^2 \\
&=\left\{\left[\frac{\theta_B(x)f_b(x,\mu_B(x),\theta_B(x),k_B(x),q_B(x))+f_s(\mu_B(x),\theta_B(x),k_B(x),q_B(x))}{\theta_B(x)+1}-\omega_B\right]^2\right. \\
&\quad\left.-\left[\frac{\theta_B(x)f_b(x,\mu_B(x),\theta_B(x),k_A(x),q_A(x))+f_s(\mu_B(x),\theta_B(x),k_A(x),q_A(x))}{\theta_B(x)+1}-\omega_{B,k_A,q_A}\right]^2\right\}\Big/2 \quad \text{variation due to changing capital (at economy B's teammates)} \\
&\quad+\left\{\left[\frac{\theta_B(x)f_b(x,\mu_A(x),\theta_A(x),k_B(x),q_B(x))+f_s(x'',\mu_B(x),\theta_A(x''),k_B(x''),q_B(x''))}{\theta_B(x)+1}-\omega_{A,k_B,q_B}\right]^2\right. \\
&\quad\left.-\left[\frac{\theta_B(x)f_b(x,\mu_A(x),\theta_A(x),k_A(x),q_A(x))+f_s(x'',\mu_B(x),\theta_A(x''),k_A(x''),q_A(x''))}{\theta_B(x)+1}-\omega_A\right]^2\right\}\Big/2 \quad \text{variation due to changing capital (at economy A's teammates)} \\
&\quad+\left\{\left[\frac{\theta_B(x)f_b(x,\mu_B(x),\theta_B(x),k_B(x),q_B(x))+f_s(\mu_B(x),\theta_B(x),k_B(x),q_B(x))}{\theta_B(x)+1}-\omega_B\right]^2\right. \\
&\quad\left.-\left[\frac{\theta_B(x)f_b(x,\mu_A(x),\theta_A(x),k_B(x),q_B(x))+f_s(x'',\mu_B(x),\theta_A(x''),k_B(x''),q_B(x''))}{\theta_B(x)+1}-\omega_{A,k_B,q_B}\right]^2\right\}\Big/2 \quad \text{variation due to changing teammates (at economy B's capital)} \\
&\quad+\left\{\left[\frac{\theta_B(x)f_b(x,\mu_B(x),\theta_B(x),k_A(x),q_A(x))+f_s(\mu_B(x),\theta_B(x),k_A(x),q_A(x))}{\theta_B(x)+1}-\omega_{B,k_A,q_A}\right]^2\right. \\
&\quad\left.-\left[\frac{\theta_B(x)f_b(x,\mu_A(x),\theta_A(x),k_A(x),q_A(x))+f_s(x'',\mu_B(x),\theta_A(x''),k_A(x''),q_A(x''))}{\theta_B(x)+1}-\omega_A\right]^2\right\}\Big/2 \quad \text{variation due to changing teammates (at economy A's capital)}
\end{aligned}
\tag{A.19}
$$

so the first two components in (A.19) account for changes in worker-capital complementarities in affecting the difference between economy B and economy A in the squared deviation of average wage of team $(x, \mu_B(x))$ with respect to the average wage of the economy (*capital deepening complementarities*). The last two components account for the effect due to changing worker-teammate complementarities (*teammates complementarities*).

We can perform a similar decomposition also for the difference in the fourth line of

(A.16):

$$
\begin{aligned}
&\left[\frac{\theta_A(x)\omega_B^u(x)+\omega_B^s(x')}{\theta_A(x)+1}-\omega_B\right]^2-\left[\frac{\theta_A(x)\omega_A^u(x)+\omega_A^s(x)}{\theta_A(x)+1}-\omega_A\right]^2\\
&=\left[\frac{\theta_A(x)f_u(x,\mu_B(x),\theta_B(x),k_B(x),q_B(x))+f_s(x',\mu_B(x),\theta_B(x'),k_B(x'),q_B(x'))}{\theta_A(x)+1}-\omega_B\right]^2\\
&\quad-\left[\frac{\theta_A(x)f_u(x,\mu_A(x),\theta_A(x),k_A(x),q_A(x))+f_s(x,\mu_A(x),\theta_A(x),k_A(x),q_A(x))}{\theta_A(x)+1}-\omega_A\right]^2\\
&=\Bigg(\left[\frac{\theta_A(x)f_u(x,\mu_B(x),\theta_B(x),k_B(x),q_B(x))+f_s(x',\mu_B(x),\theta_B(x'),k_B(x'),q_B(x'))}{\theta_A(x)+1}-\omega_B\right]^2\\
&\quad-\left[\frac{\theta_A(x)f_u(x,\mu_B(x),\theta_B(x),k_A(x),q_A(x))+f_s(x',\mu_B(x),\theta_B(x'),k_A(x'),q_A(x'))}{\theta_A(x)+1}-\omega_{B,k_A,q_A}\right]^2\Bigg)\Big/2 \quad \text{variation due to changing capital (at economy B's teammates)}\\
&\quad+\Bigg(\left[\frac{\theta_A(x)f_u(x,\mu_A(x),\theta_A(x),k_B(x),q_B(x))+f_s(x,\mu_A(x),\theta_A(x),k_B(x),q_B(x))}{\theta_A(x)+1}-\omega_{A,k_B,q_B}\right]^2\\
&\quad-\left[\frac{\theta_A(x)f_u(x,\mu_A(x),\theta_A(x),k_A(x),q_A(x))+f_s(x,\mu_A(x),\theta_A(x),k_A(x),q_A(x))}{\theta_A(x)+1}-\omega_A\right]^2\Bigg)\Big/2 \quad \text{variation due to changing capital (at economy A's teammates)}\\
&\quad+\Bigg(\left[\frac{\theta_A(x)f_u(x,\mu_B(x),\theta_B(x),k_B(x),q_B(x))+f_s(x',\mu_B(x),\theta_B(x'),k_B(x'),q_B(x'))}{\theta_A(x)+1}-\omega_B\right]^2\\
&\quad-\left[\frac{\theta_A(x)f_u(x,\mu_A(x),\theta_A(x),k_B(x),q_B(x))+f_s(x,\mu_A(x),\theta_A(x),k_B(x),q_B(x))}{\theta_A(x)+1}-\omega_{A,k_B,q_B}\right]^2\Bigg)\Big/2 \quad \text{variation due to changing teammates (at economy B's capital)}\\
&\quad+\Bigg(\left[\frac{\theta_A(x)f_u(x,\mu_B(x),\theta_B(x),k_A(x),q_A(x))+f_s(x',\mu_B(x),\theta_B(x'),k_A(x'),q_A(x'))}{\theta_A(x)+1}-\omega_{B,k_A,q_A}\right]^2\\
&\quad-\left[\frac{\theta_A(x)f_u(x,\mu_A(x),\theta_A(x),k_A(x),q_A(x))+f_s(x,\mu_A(x),\theta_A(x),k_A(x),q_A(x))}{\theta_A(x)+1}-\omega_A\right]^2\Bigg)\Big/2 \quad \text{variation due to changing teammates (at economy A's capital)},
\end{aligned}
\tag{A.20}
$$

so the first two components in (A.20) account for changes in worker-capital complementarities in affecting the difference between economy B and economy A in the squared deviation of average wage of team $(x, \mu_B(x))$ with respect to the average wage of the economy (*capital deepening complementarities*). The last two components account for the effect due to changing worker-teammate complementarities (*teammates complementarities*).

5. A discussion on *direct effect* decomposition In our terminology, there is a direct effect whenever economy A and B differ in terms of the production function parameters or in the generalized distributions of $\mathcal{U}$ and $\mathcal{S}$. More specifically, in the latter case, a difference can occur in the *total mass* of workers, in the *support* of the distributions, or in the *pdf*s.

In the case of a change in the parameters of the production function, the decomposition extension is relatively straightforward. It is sufficient, for instance in last formula (A.20), to divide by 4 instead of 2 and consider twice as many variations. In particular, each variation computed in (A.20) should be evaluated twice, once using the marginal product function with parameters from economy A and one from economy B.

The decomposition for a change in the generalized CDFs is more controversial as there are possible alternative decompositions. One possibility is integrating the formulas above in x only for the CDFs overlap between economy A and B. Any left-out measure of workers would be directly attributable to difference in the labor supply. A second possibility is mapping the two CDFs one into the other through a *ranking measure*. In this case, the integration can be performed using the ranking measure directly and any difference between the decomposition based on this measure and the original CDFs can be attributed to differences in the labor supply. We leave the exploration of these alternative possibilities for future work.

Descriptive statistics for the whole population, the unbalanced and the balanced sample

	1998		2001	
	All	Baseline	All	Baseline
Number of firms	11,530	2,693	11,697	3,580
Number of workers	293,496	114,471	299,124	154,351
Number of equivalent workers[a]	254,784	100,857	258,443	135,483
% blue-collars	69.72	70.41	79.27	68.47
Firm size	22.097	37.452	22.094	37.845
	(83.599)	(84.927)	(74.328)	(109.429)
Blue-white collars ratio	4.440	4.023	4.391	3.972
	(5.932)	(4.752)	(5.692)	(4.480)
ln(Tangible assets per white-collar)[b]		4.468		4.493
		(1.218)		(1.276)
ln(Intangible assets per white-collar)[b]		1.676		1.920
		(1.432)		(1.511)
Average ln(daily wage)	4.606	4.685	4.632	4.704
	(0.246)	(0.203)	(0.201)	(0.177)
Wage std. per firm	0.205	0.239	0.216	0.250
	(0.133)	(0.114)	(0.129)	(0.111)
Total wage variance[c,d]	100.0	100.0	88.5	107.95
% within firms	22.36	45.12	27.45	49.57
% between firms	77.64	54.88	72.55	50.43
Raw total wage variance[c]	100.0	100.0	89.28	100.68
% within firms	55.49	73.48	61.48	78.11
% between firms	44.51	26.52	38.52	21.89

Table A.1: Standard deviation in parenthesis. All monetary values are in 2003 euro. [a]: Total number of days worked divided by 360; all the statistics except for the first two lines are weighted for firm size as measured by equivalent employees. [b]: not available since most of the firms in "All" do not report assets. [c]: Total wage variance in 1998 is normalized to 100 and total wage variance in 2001 is normalized to the labor force number of 1998. Wage variance decomposition is reported after substituting individual wages with the average wage in the respective firm-occupation cell. [d]: Wage variance decomposition is reported after substituting individual wages with the average wage in the respective firm-occupation cell.

A.3 Additional data information

In addition to what reported in section 1.3.1, we also drop as outliers those firms reporting: a) a white-blue-collar ratio (expressed in equivalent workers) below 0.01 or above 100; b) with an average daily wage above 300; c) with tangible assets and intangible assets per white worker, respectively, above 2 million euro or 0.8 million euro.

In table A.1 we report a set of descriptive statistics for the whole dataset of firm (columns "all"), including outliers and those firms not reporting their assets for either 1997 or 2001. The baseline sample is not representative of the universe of active firms as average firm size is almost double. However, average daily wage and wage standard deviation per firm are only marginally larger for the baseline sample. The most remarkable difference is in the total wage variance as in the universe of firms decreases about 11% while for our baseline sample is increasing. While the baseline sample total variance increase is mainly due to changes in the within firm component, the decrease in the firm universe total variance is mainly attributable to a sharp reduction in the dispersion across firms (from .082 to .067). The role of the within firm component of inequality is mechanically larger in the case of

Additional correlations

	1998		2001	
	data	model	data	model
corr(θ, ω)	-0.324	-0.940	-0.403	-0.945
corr(θ, σ)	-0.246	-0.920	-0.335	-0.926
corr(θ, k)	0.282	-0.885	0.250	-0.892
corr(θ, q)	0.065	-0.939	-0.009[a]	-0.951
corr(ω, σ)	0.121	0.998	0.315	0.998
corr(ω, k)	-0.006[a]	0.991	0.056	0.991
corr(ω, q)	0.060	1.000	-0.052	1.000
corr(σ, k)	0.063	0.997	0.046	0.997
corr(σ, q)	0.246	0.999	0.276	0.997
corr(k, q)	0.245	0.991	0.253	0.988

Table A.2: ω is the average log wage per firm; σ is the firm wage dispersion; θ is the firm ratio between blue- and white-collar workers (in equivalent units); k is the log of structures (tangible assets in the data) per equivalent white-collar worker in the firm; q is the log of equipment (intangible assets in the data) per equivalent white-collar worker in the firm. The data coefficients are for the cross-section years reported on the headings of the table. The model counterpart coefficients are calculated for the functions $\{\omega(x), \sigma(x), \theta(x), k(x), e(x)\}$ when solving for the model specifications reported in table 1.3. The data moments are computed weighting firms by their employment equivalent size. [a]: Not significant at the 0.1 level.

the raw total wage variance, because it also includes dispersion within the firm-occupation cell.

A.4 Additional calibration information

We calibrate our parameters trying to match the moments from table 1.4. However, the rich heterogeneity present in the model also allows us to compare to other data moments. Tables A.2 and A.3 compare, respectively, the model correlations and percentiles of $\{\theta, \omega, \sigma, k, e\}$ with the data.

Comparison of data and simulation percentiles

	mean	s.d.	p10	p25	p50	p75	p90
Percentiles for 1998							
ω : Average wage within firm							
data	4.685	0.203	4.477	4.576	4.685	4.794	4.898
model	4.703	0.038	4.652	4.671	4.702	4.735	4.755
σ : Std. of wages within firm							
data	0.239	0.114	0.114	0.161	0.225	0.301	0.375
model	0.235	0.062	0.148	0.182	0.236	0.288	0.319
θ : Blue - White collars ratio							
data	4.023	4.752	0.553	1.304	2.802	4.889	8.589
model	4.000	0.052	3.916	3.965	4.022	4.044	4.048
k : Capital structure intensity[a]							
data	4.468	1.218	2.807	3.634	4.541	5.352	5.989
model	4.463	0.028	4.422	4.441	4.467	4.488	4.498
q : Capital equipment intensity[a]							
data	1.676	1.432	-0.082	0.693	1.631	2.639	3.584
model	3.584	0.010	3.570	3.575	3.584	3.592	3.598
Percentiles for 2001							
ω : Average wage within firm							
data	4.704	0.177	4.489	4.580	4.698	4.814	4.930
model	4.703	0.038	4.652	4.671	4.702	4.795	4.755
σ : Std. of wages within firm							
data	0.250	0.111	0.117	0.168	0.236	0.316	0.399
model	0.235	0.062	0.148	0.182	0.236	0.288	0.319
θ : Blue - White collars ratio							
data	3.972	4.480	0.541	1.268	2.776	5.000	8.565
model	4.000	0.053	3.913	3.637	4.022	4.047	4.050
k : Capital structure intensity[a]							
data	4.493	1.276	2.772	3.625	4.594	5.425	6.081
model	4.464	0.028	4.424	4.442	4.679	4.489	4.499
q : Capital equipment intensity[a]							
data	1.920	1.511	0.000	0.821	1.872	2.944	3.928
model	3.650	0.007	3.640	3.644	3.650	3.656	3.661

Table A.3: Data and simulated percentiles and mean for a set of variables. [a]: intensity is defined as the log of thousand euros of capital per one unit equivalent of white-collar worker.

B Chapter 2 appendix

B.1 Proof of proposition 2.1

We simplify the economy presented in section 2.3.1 in assuming production takes place in *teams* of workers only, without a frim y input. This semplication is for sake of clarity as results go through identically in the presence of y but calculations are more involved. The simplified economy reads as follows. Male workers are indexed in x_M and distributed according to the *pdf* h_M, while female are indexed in x_F and distributed over h_F. A team of type x_M workers matches with a team of x_F workers, they produce, sell and split the profits. The problem of the female worker type x_F is therefore:

$$\max_{\{x_M, l_F, l_M\}} F(x_F, x_M, l_F, l_M) - (\tau_a + \tau_m w_F(x_F))l_F - w_M(x_M)l_M. \tag{B.1}$$

where F display constant returns to scale in $\{l_F, l_M\}$. Therefore the x_M problem can equivalently be re-written using an intensity unit production function $f(x_F, x_M, \theta) = 1/l_M F(x_F, x_M, l_F/l_M, 1)$, with θ denoting the ratio of women to men in the firm. The simplified problem is:

$$\max_{\{x_F, \theta\}} f(x_F, x_M, \theta) - w_M(x_M)\theta. \tag{B.2}$$

In equilibrium, all firms chose policies that solve that problem subject to the feasibility or market clearing constraint:

$$\int_{\mu(x_F)}^{\bar{x}_M} h_M(s)ds = \int_{x_F}^{\bar{x}_F} h_F(s)ds, \tag{B.3}$$

where μ_M denotes the optimal hiring policy in quality of worker type x_M.

Taking the first order conditions of (B.2) we obtain

$$\begin{aligned} f_{x_F} &= \tau_m \cdot w'_F(x_F)\ \theta(x_F) \\ f_\theta &= \tau_a + \tau_m \cdot w_F(x_F). \end{aligned}$$

Total differentiate the second FOC in x_F:

$$f_{x_F\theta} + f_{y\theta}\mu'(x_F) + f_{\theta\theta}\theta'(x_F) = \tau_m \cdot w'_F(x_F),$$

and substituting from the first order condition we get

$$f_{x_F\theta} + f_{y\theta}\mu'(x_F) + f_{\theta\theta}\theta'(x_F) = \tau_m \cdot \frac{f_{x_F}}{\tau_m \cdot \theta(x_F)} = \frac{f_{x_F}}{\theta(x_F)}, \tag{B.4}$$

which does not depend on τ_m and τ_a. Consequently, it is possible to follow Eeckhout and Kircher (2018) in deriving the differential equations for determining equilibrium without need to refer to τ_m or τ_a.

Notice that (B.4) needs to hold for the whole domain of x_F if all women are employed. However, supposing that bottom x_F are not employed because their opportunity cost from home production is too high (a high value for τ_a), than the boundary condition associated to (B.3) also changes (see footnote 13 in Eeckhout and Kircher 2018) and this affect the domain where (B.4) needs to be verified. Thus, the equilibrium μ and θ for paritcipanting types in x_F is also affected.

C Chapter 3 appendix

C.1 Walras' Law

In this section I show that market clearing conditions and stationary unemployment are linearly dependent. In other words, we need to substitute one of these condition with the population feasibility constraint in order to find the equilibrium masses of entrant firms and unemployed workers. This linear dependency implies that the equilibrium mass of entry $\vec{m}_e$ and unemployment by human capital type are unique.

Job accounting per type of entrant firm Denote with $\eta \in \{1, \dots, H\}$ the "dynasty" of firms that in their period of entry only hired workers with human capital of type η. Denote with $a \in \mathbf{N}$ the age of a firm and with $(z, \vec{n})$ its productivity-workforce state space.

The distribution $g_\eta(z, \vec{n}, a)$ can be described either at stage B or stage C according to the following recursive laws of motion:

$$g^B_\eta(z, \vec{n}, 0) = \pi_e(z)\mathbb{1}\{0 == \vec{n}\} \tag{C.1}$$
$$g^C_\eta(z, \vec{n}, 0) = \pi_e(z)(1 - d_e(z))\mathbb{1}\{\vec{n}_e == \vec{n}\} \tag{C.2}$$
$$g^B_\eta(z', \vec{n}, a) = \sum_z \pi_z(z'|z) g^C_\eta(z, \vec{n}, a-1) \tag{C.3}$$
$$g^C_\eta(z', \vec{n}', a) = \sum_{z,\vec{n}} \pi_z(z'|z) g^C_\eta(z, \vec{n}, a-1)(1 - d_i(z', \vec{n})\mathbb{1}\{\vec{n}'(z', \vec{n}) == \vec{n}'\} \tag{C.4}$$

As long as there are no immortal firms (a sufficient condition is $\delta_0 > 0$), it must be the case that $\forall\eta$ the total mass of jobs created is equal to the mass of job separations (quits + layoffs & mass layoffs + deaths). One can formalize the intuition above in formulas and write:

$$JC_\eta(k) = \sum_z \left[g^C_\eta(z, \vec{n}, 0) n_{e,k}(z)\right] + \sum_{a>0}\left[\sum_{z',z,\vec{n}} g^B_\eta(z, \vec{n}, a)\left[n_{i,k}(z', \vec{n})(1 - d_i(z', \vec{n}))\right]\right] \tag{C.5}$$

$$JQ_\eta(k) = s_0 \sum_{a>0}\left[\sum_{z',z,\vec{n}} g^B_\eta(z, \vec{n}, a)\left[\sum_{m=1}^{H} n_m \pi_w(k|m; \vec{n})\right]\left[\lambda p(\theta(x_w(k, z', \vec{n}), k))(1 - \tau_i(k, z', \vec{n}))(1 - d_i(z', \vec{n}))\right]\right] \tag{C.6}$$

$$JL_\eta(k) = s_0 \sum_{a>0}\left[\sum_{z',z,\vec{n}} g^B_\eta(z, \vec{n}, a)\left[\sum_{m=1}^{H} n_m \pi_w(k|m; \vec{n})\right]\left[d_i(z', \vec{n}) + \tau_i(k, z', \vec{n})(1 - d_i(z', \vec{n}))\right]\right] \tag{C.7}$$

$$D_\eta(k) = (1 - s_0) \sum_{a>0}\left[\sum_{z',z,\vec{n}} g^B_\eta(z, \vec{n}, a)\left[\sum_{m=1}^{H} n_m \pi_w(k|m; \vec{n})\right]\left[(1 - \lambda p(\theta(x_w(k, z', \vec{n}), k)))(1 - \tau_i(k, z', \vec{n}))(1 - d_i(z', \vec{n}))\right]\right] \tag{C.8}$$

where the following holds:

$$\sum_k JC_\eta(k) - JQ_\eta(k) - JL_\eta(k) - D_\eta(k) = 0. \tag{C.9}$$

Unemployment and life-cycle accounting Denote with $k \in \{1,\ldots,H\}$ the unemployed workers with human capital in grid point h. For brevity, set the job finding probability of an unemployed worker of type h to $JF_u(k) = p(\theta(x_u(k),k)$. Stationarity requires that the inflows and outflows from unemployment are identical. Denote with m_η the mass of entrant firms of type η. We can write:

$$u(k) = [1 - JF_u(k)] \left((1-s_0)\pi_0(k) + s_0 \sum_{m=1}^{H} \pi_u(k|m)u(m) \right) + \sum_\eta JL_\eta(k)m_\eta \qquad \text{(C.10)}$$

Notice worker population is stationary across human capital types, but it does not need to be balanced by human capital class. In other words, we can write:

$$(1-s_0) = (1-s_0)\sum_k u(k) + \sum_k \sum_\eta m_\eta D_\eta(k). \qquad \text{(C.11)}$$

Market clearing and stationary unemployment Market clearing requires that in every period the total number of jobs created in each aggregate human capital market h is equal to the number of jobs found by unemployed and job switchers. Market clearing can be formulated as $\forall k$:

$$\sum_\eta JC_\eta(k)m_\eta = \sum_\eta JQ_\eta(k)m_\eta + JF_u(k)\left((1-s_0)\pi_o(k) + s_0 \sum_{m=1}^{H} \pi_u(k|m)u(m) \right) \qquad \text{(C.12)}$$

Lemma C.1. ***Walras' law*** *implies that the* $2 * H$ *expression in* (C.10) *and* (C.12) *are linearly dependent.*

Proof. This result can be shown by summing up the H expressions in (C.12) and using (C.9) for each η to substitute out for $\sum_k JC_\eta(k) - JQ_\eta(k)$:

$$\sum_k \sum_\eta \left(JL_\eta(k) + D_\eta(k)\right) m_\eta = \sum_k JF_u(k) \left((1-s_0)\pi_o(k) + s_0 \sum_{m=1}^{H} \pi_u(k|m)u(m) \right)$$

Similarly add up the H expressions in (C.10) and rearrange after using (C.11):

$$\sum_k u(k) = \sum_k [1 - JF_u(k)] \left((1-s_0)\pi_0(k) + s_0 \sum_{m=1}^{H} \pi_u(k|m)u(m) \right) + \sum_k \sum_\eta JL_\eta(k)m_\eta \iff$$

$$\sum_k \sum_\eta \left(JL_\eta(k) + D_\eta(k)\right) m_\eta = \sum_k JF_u(k) \left((1-s_0)\pi_o(k) + s_0 \sum_{m=1}^{H} \pi_u(k|m)u(m) \right).$$

□