

Production and Learning in Teams *

KYLE HERKENHOFF

University of Minnesota and NBER

JEREMY LISE

University of Minnesota and FRB Minneapolis

GUIDO MENZIO

NYU and NBER

GORDON PHILLIPS

Dartmouth College and NBER

September 2023

Abstract

To what extent is a worker's human capital growth affected by the quality of his coworkers? To answer this question, we develop and estimate a model in which the productivity and the human capital growth of an individual depend on the average human capital of his coworkers. The measured production function is supermodular: The marginal product of a more knowledgeable individual is increasing in the human capital of his coworkers. The measured human capital accumulation function is convex: An individual's human capital growth is increasing in coworkers' human capital only when paired with more knowledgeable coworkers, but independent of coworkers' human capital when paired with less knowledgeable coworkers. Learning from coworkers accounts for two thirds of the stock of human capital accumulated on the job. Technological changes that increase production supermodularity lead to labor market segregation and, by reducing the opportunities for low human capital workers to learn from better coworkers, lead to a decline in aggregate human capital and output.

JEL Codes: E24, J24.

Keywords: Human Capital, Knowledge Diffusion, Search Frictions.

*We are grateful to the editor, Gianluca Violante, and three anonymous referees for their thoughtful comments and suggestions. We also wish to thank Axel Anderson, David Autor, Thibault Lamadon, Rasmus Lentz, Ellen McGrattan, Ricardo Lagos, Espen Moen, Tommaso Porzio, Jean-Marc Robin, Coen Teulings, Michela Tincani, and participants at many conferences. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. Herkenhoff and Phillips were supported by NSF grant SES-1824422. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation.

1 Introduction

At least since Becker (1962) and Ben-Porath (1967), learning on the job has been recognized as an important mechanism of human capital accumulation and as the main source of wage growth over the life-cycle. In some formulations, learning on the job has been modeled as an investment process in which a worker’s own time is the only input (e.g., Ben-Porath 1967, Guvenen and Kuruscu 2012). In other formulations, learning on the job has been modeled as a very general investment process in which both the worker’s time and the employer’s assets are used as inputs (e.g., Rosen 1972). In macroeconomics, learning on the job has been typically modeled as a mechanical process of learning-by-doing in which human capital grows at some given rate as long as the worker is employed (e.g., Arrow 1962). The literature, however, has paid little attention to how the quality of coworkers affects the speed at which an individual accumulates human capital on the job. The omission is rather surprising, especially since the schooling literature has documented that the quality of peers is an important determinant of the rate at which an individual student accumulates human capital in the classroom (e.g., Hoxby 2000, Carrell, Sacerdote and West 2013).

In this paper, we measure the extent to which a worker’s human capital growth on the job is affected by the quality of his coworkers. To carry out this measurement, we take a structural approach—we use a model of the labor market to translate empirical evidence on the relation between the wages of the coworkers of an individual and the wage growth of the individual into a causal relation between the human capital of the coworkers of an individual and the human capital growth of the individual. Specifically, using a large matched employer-employee dataset, we first document the relation between the future wage of a worker and the wage of his current coworkers, controlling for the worker’s current wage. The empirical relation does not have a direct causal interpretation because, unless the sorting of workers is random, the wage of the worker and the wage of his coworkers are both noisy measures of the worker’s human capital and, hence, both forecast the worker’s future wage. Using a model of the labor market that is calibrated to reproduce the same pattern of sorting as in the data, we then back out the actual effect of coworkers’ human capital on individual learning from the empirical wage relation.

The main finding of the paper is that the average human capital of an individual’s coworkers has a convex effect on the human capital growth of the individual. More precisely, the human capital growth of an individual who is less knowledgeable than his coworkers depends positively on the average human capital of his coworkers. In contrast, the human capital growth of an individual who is more knowledgeable than his cowork-

ers is essentially independent of the average human capital of his coworkers. In other words, an individual who knows less than his coworkers tends to catch up to them, while an individual who knows more than his coworkers is not dragged-down to their level. Quantitatively, learning from coworkers accounts for about two thirds of the human capital stock that is accumulated on the job, while the remaining third is accounted for by learning-by-doing or other mechanisms that do not depend on the quality of coworkers.

In the first part of the paper, we lay out the structural model of the labor market. In the model, a firm produces output by hiring teams of two workers. The production function depends on the human capital of the two workers. The production function may be supermodular, in the sense that the additional output generated by a worker with more human capital is increasing in the human capital of the other worker, submodular, in the sense that the additional output generated by a worker with more human capital is decreasing in the human capital of the other worker, or modular. In the first case, output is maximized by sorting workers positively. In the second case, output is maximized by sorting workers negatively. In the third case, output is independent of sorting. The worker's human capital accumulation function depends on the worker's own human capital and on the human capital of his coworker. The worker's accumulation function may be concave, convex or linear in the human capital of the coworker. In the first case, learning is maximized by sorting workers negatively. In the second case, learning is maximized by sorting workers positively. In the third case, learning is independent of sorting. The labor market is subject to search frictions. We make this assumption because it implies that wages do not need to reflect changes in workers' human capital instantaneously, but only when a worker changes employer or receives a sufficiently attractive outside offer. This implication turns out to be critical to properly interpret the data.

In the second part of the paper, we access a large matched employer-employee dataset in order to document the relation between a worker's wage growth and the average wage of his past coworkers, which we show is the key moment of the coworkers' wage distribution. We restrict attention to workers who experience a job-to-job transition with an intervening spell of unemployment, so that their wage in the second job reflects their stock of human capital without being affected by the quality of the first job. For workers who earn less than their coworkers on the first job, we find that a 10% increase in the coworkers' average wage forecasts a 1.23% higher wage on the second job. For workers who earned more than their coworkers on the first job, we find a 10% decline in the coworkers' average wage forecasts only a 0.4% lower wage in the second job. These estimates suggest that learning on the job is convex in the human capital of the coworkers. The estimates, however, do not have an immediate structural interpretation because the sorting of workers and

coworkers is not random. In order to measure the pattern of sorting, we document the relation between a worker’s wage and the wage of future coworkers. For workers who earn less than their coworkers in the first job, a 10% increase in their own wage forecasts a 2.5% higher average coworkers’ wage in the next job. For workers who earn more than their coworkers in the first job, a 10% increase in their wage forecasts a 1.7% higher coworkers’ wage in the second job. These estimates reveal that sorting is positive.

Using the structural model, we translate the empirical evidence into parameters of the human capital accumulation and production functions. Specifically, we calibrate the model so that it generates the same relation between a worker’s future wages and the average wage of his past coworkers and the same relation between a worker’s wage and the average wage of his future coworkers as in the data. Additionally, we calibrate the model to reproduce the extent of wage dispersion across workers and across firms, the extent of lifecycle wage growth, and the same frequency of employment transitions that we observe in the data. We find that the human capital accumulation function is indeed convex. In particular, when a worker is less knowledgeable than his coworkers, the coworkers’ human capital has a strong positive effect on the speed at which the worker learns. In contrast, when a worker is more knowledgeable than his coworkers, the coworkers’ human capital has no effect on the worker’s speed of learning. We find that the production function is supermodular in the human capital of the worker and his coworkers. The calibrated model reproduces well all of the empirical targets. Moreover, the calibrated model successfully reproduces the rate at which different types of workers separate from different types of coworkers—both towards unemployment and towards other jobs.

Overall, our structural model is very simple and in some dimensions unrealistic. Indeed, in the model workers have a single coworker rather than many. Once multiple coworkers are collapsed into a representative one, however, the model reproduces quite well the consequences for an individual from being matched with different coworkers (effect of treatments), the frequency with which an individual is matched with different coworkers (frequency of treatments), and how long an individual stays with different coworkers (duration of treatments). That is, the model reproduces quite well the pattern of interactions between an individual and different coworkers. For this reason, we believe that, even though the model is stylized, it does provide credible estimates of the human capital growth and the productivity of an individual when matched with different coworkers.

In the third part of the paper, we use the structural model to carry out some counterfactuals. We first study a counterfactual in which the worker’s learning on the job is unaffected by their coworkers. The counterfactual shows that learning from peers makes the pattern of sorting between workers and coworkers less positive than it would be other-

wise, that learning from peers increases the speed at which individuals accumulate human capital over the lifecycle, and that it increases both the aggregate stock of human capital and the flow of output. Quantitatively, learning from peers accounts for two thirds of the stock of human capital accumulated by workers on the job, while the remaining third is due to learning by doing.

We then study a counterfactual in which technological changes lead to an increase in the supermodularity of the production function, a hypothesis put forward by Kremer (1993). The counterfactual shows that an increase in supermodularity makes the pattern of sorting between workers and coworkers more positive and, in this sense, increases the extent of labor market segregation, a phenomenon documented by Song et al. (2019). Since an increase in labor market segregation implies that low human capital workers have fewer chances to learn from more knowledgeable coworkers, the aggregate stock of human capital declines and so does the flow of output.

Lastly, we compare the equilibrium of the model with the solution to the problem of a utilitarian planner. We prove that the equilibrium is always inefficient. For the benchmark calibration of the labor market, we find that the equilibrium features an inefficiently positive pattern of sorting, an inefficiently low stock of human capital, and an inefficiently low flow of output. All these inefficiencies, however, are rather small.

The paper is motivated by some recent theoretical research that examines the role played by knowledge diffusion—as opposed to knowledge creation—in the aggregate growth of an economy (e.g., Lucas 2009, Lucas and Moll 2014, Perla and Tonetti 2014, Jovanovic 2014, Benhabib, Perla and Tonetti 2021). In these theoretical studies, knowledge diffusion is modeled as the outcome of a random bilateral meeting process between agents. When two agents meet, the one with the lower human capital instantaneously absorbs the stock of knowledge of the one with the higher human capital. Seeking a concrete and quantifiable implementation of these theories, we found it natural to model knowledge diffusion as a time-consuming process that takes place within the workplace. In contemporaneous work, Jarosch, Oberfield, and Rossi-Hansberg (2021) followed the same approach. In follow-up work, Martellini (2020) modeled knowledge diffusion as a process that takes place within a city. Gregory (2020) modeled the knowledge diffusion process in reduced form, by assuming that there are firm-specific features that affect the rate at which their employees accumulate human capital. In related prior work, Lise and Postel-Vinay (2020) considered a version of learning by doing in which workers accumulate the skills that are required by their job.

In studying knowledge diffusion in the workplace, we are forced to confront the issue of sorting on unobservable characteristics (e.g., human capital), which makes it challenging

to interpret any statistical relation between the wage growth of an individual and the wage of his previous coworkers. Sorting on unobservables is also the main challenge in the related and much larger literature on peer effects in schools. In the school context, the challenge is either addressed by looking for exogenous variation in school composition (see, e.g., Hoxby 2000) or by manufacturing exogenous variation through randomized experiments (see, e.g., Carrell, Sacerdote and West 2013 or Booij, Leuven and Oosterbeek 2017). Clearly, such exogenous variation is harder to find, and even harder to manufacture, in the workplace. For this reason, the relatively small literature on knowledge diffusion in the workplace tries to control for unobserved heterogeneity by adding fixed-effects (Cornelissen, Dustmann, and Schoenberg 2018, Nix 2020). We take a different approach. We explicitly model the process of sorting and calibrate it to the data. This allows us to use the model to infer the true causal relation between an individual’s human capital growth and the human capital of his coworkers from the statistical relation between an individual’s wage growth and the wage of his coworkers. We are also the first to explicitly acknowledge the presence of search frictions in the labor market, which forces us to recognize that human capital growth is not immediately priced into wages.

One of the key object of our analysis is the pattern of sorting of workers and coworkers. Key theoretical works on equilibrium sorting in frictionless markets are Becker (1962)—which assumes that agents’ types are fixed—and Anderson and Smith (2010) and Anderson (2015)—which allow for agents’ types to evolve based on whom they match with. A key insight from Anderson and Smith (2010) and Anderson (2015) is that the pattern of sorting depends on both the shape of the production function and the shape of the law of motion for types. Our paper makes extensive use of this insight. There is also a more empirical literature on sorting in labor markets with search frictions (e.g., Lise, Meghir and Robin 2016, Hagedorn, Law and Manovskii 2018, Bagger and Lentz 2019, Lopes de Melo 2018, Lentz, Piyapromdee and Robin 2023, Eeckhout and Kircher 2011, Shimer and Smith 2000). Departing from this literature, we focus on the sorting of workers and coworkers, rather than of workers and firms. Also departing from this literature, we focus on the sorting of workers when not only output, but also learning, depends on who matches with whom.

Our search-theoretic model of the labor market is an extension of Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay and Robin (2006). These models of on-the-job search are a standard framework for studying the employment and wage dynamics of individual workers, the growth of wages over the lifecycle, and the cross-sectional dispersion of wages (see also Burdett and Mortensen 1998, Bagger et al. 2014, Menzio, Telyukova and Visschers 2016, Lise and Postel-Vinay 2020). We contribute to the on-

the-job search framework by modeling—albeit in a simple fashion—the relation between coworkers in both production and human capital accumulation. Technically, modeling the relation between coworkers forces us to introduce multi-worker firms and to extend the concept of gains from trade to situations involving more than one worker and one firm. In contemporaneous work, Bilal, Engbom, Mongey and Violante (2022) proceed in a similar direction—but endogenize firm size and focus on the number of coworkers rather than on the heterogeneity among coworkers. Substantively, modeling the relation between coworkers leads us to discover that luck (i.e., whom an individual worker meets) does not only affect an individual’s current productivity and share of the gains from trade, but—through its impact on human capital growth—also his future productivity.

2 Theory

In this section, we develop a structural model of the labor market in which the quantity of output and the accumulation of human capital of an individual depend on the human capital stock of his coworkers. In Section 2.1, we describe the model and discuss the main assumptions. In Section 2.2, we define a stationary equilibrium of the model.

2.1 Environment

The labor market is populated by a measure 1 of workers. A worker maximizes the present value of labor income discounted at the factor $\beta \in (0, 1)$. At a given point in time, a worker has human capital $h_k \in H$, where $H = \{h_1, h_2, \dots, h_N\}$ and $0 < h_1 < h_2 < \dots < h_N$. We refer to H as the *human capital ladder*. We refer to a worker’s position $k \in K = \{1, 2, \dots, N\}$ along the human capital ladder H as the worker’s *type*. At a given point in time, the worker is in employment state $x \in X$, where $X = \{u, 0, 1, 2, \dots, N\}$. If the worker is in state u , he is unemployed. If the worker is in state 0, he is employed without a coworker. If the worker is in states $\ell = 1, 2, \dots, N$, he is employed with a coworker of type ℓ .

The labor market is also populated by a measure $n > 0$ of firms. A firm maximizes the present value of profits discounted at the factor β . At a given point in time, a firm is in state $y = (\bar{k}, \bar{\ell}) \in Y$, with $Y = \bar{K} \times \bar{K}$ and $\bar{K} = \{0, 1, 2, \dots, N\}$. If the firm is in state $(0, 0)$, it has no employees. If the firm is in state $(k, 0)$ for $k \in K$, it has one employee of type k . If the firm is in state (k, ℓ) for $k, \ell \in K$, it has an employee of type k and an employee of type ℓ . For the sake of simplicity, we assume that a firm can have at most two employees. A firm in state y produces $f(y)$ units of output, where $f : Y \rightarrow \mathbb{R}_+$. The production function f is such that a firm without employees does not produce any output, i.e. $f(0, 0) = 0$. The production function f is symmetric, i.e. $f(\bar{k}, \bar{\ell}) = f(\bar{\ell}, \bar{k})$.

The specification of f allows for the marginal product of an employee to depend not only on his own human capital, but also on the human capital of his coworker.

Every period is divided in four stages: *learning*, *entry-and-exit*, *search-and-matching* and *production*. At the *learning stage*, the human capital of a worker evolves according to a probability distribution function that depends on the worker's type k and on the worker's employment state x . Formally, the worker's human capital evolves from h_k to h_{k+} with probability $g_k(k_+|x)$, with $g_k(k_+|x) \geq 0$ and $\sum_{k_+ \in K} g_k(k_+|x) = 1$. The specification of g allows for the human capital of a worker to evolve differently depending on whether he is employed or unemployed and, conditional on being employed, on the human capital of his coworker.

At the *entry-and-exit stage*, a worker permanently exits the labor market with probability $\sigma \in (0, 1)$. Hence, in aggregate, a measure σ of workers exits the labor market. The workers who exit the labor market are replaced by an equal measure of workers who enter the labor market. The fraction of entering workers who are on the k -th rung of the human capital ladder is π_k , with $\pi_k \geq 0$ and $\sum_{k \in K} \pi_k = 1$. All workers who enter the labor market are unemployed.

At the *search-and-matching stage*, workers and firms meet, match, and separate. An employed worker becomes unemployed for exogenous reasons with probability $\delta \in (0, 1)$. An unemployed worker meets a randomly-selected firm with probability $\lambda_u \in (0, 1]$. An employed worker meets a randomly-selected firm with probability $\lambda_e \in [0, 1]$. Since the meeting process is random, a worker may contact a firm without employees, a firm with one employee, or a firm with two employees. For the same reason, a firm may be contacted by a worker who is unemployed, by a worker who is employed by himself, or by a worker who is employed with someone else. Upon meeting, the worker and the firm match—in the sense that the worker is hired by the firm—if and only if the gains from trade are positive. Since firms can employ at most two workers, a firm with two employees must fire one of them in order to hire a new worker. A firm may also choose to fire an employee without having a replacement for him. A firm does so if and only if the gains from trade have become negative because of either changes in the firm's state or changes in the employee's type.

At the *production stage*, firms in state y generate output according to the production function $f(y)$, they pay some of the output to their employees as wages, and they earn the remaining output as profits. Unemployed workers of type k home-produce $b_k > 0$ units of output.

Hiring decisions are based on the *gains from trade* between a firm and a worker. The gains from trade are defined as the difference between the *value of the firm-worker match*

and the sum of the *outside options of the firm and the worker*. The value of the firm-worker match is defined as the sum of the values to the firm, the firm's current employees, and the worker of being in a production unit that includes the worker. The outside option of the firm is the sum of the values to the firm and its employees of being in a production unit that does not include the worker. The outside option of the worker is the value of unemployment, if the worker is unemployed. If the worker is employed, the outside option of the worker is the *marginal value of the worker* to his current employer—that is, the sum of the values to the worker, the current employer, and the current coworkers of being in a production unit that includes the worker, net of the sum of the values to the employer and the coworkers of being in a production unit without the worker. Given the definition of the value of a firm-worker match and the outside option of a firm, it follows that the gains from trade between a firm and a worker are equal to the marginal value of the worker to the firm net of the worker's outside option.

Firing decisions are also based on the gains from trade. A firm fires a worker if the gains from trade are negative, in the sense that the joint value of the firm-worker match is smaller than the sum between the firm's outside option (i.e. the value to the firm and its remaining employees from not having the worker in their production unit) and the worker's outside option (i.e. the value of unemployment).

The gains from trade determine the transitions of workers across employment states. The division of the gains from trade is determined by bargaining. When a firm hires a worker, the two parties bargain, the worker captures a fraction γ and the firm captures a fraction $1 - \gamma$ of the gains from trade, with $\gamma \in [0, 1]$. The worker's share of the gains from trade is delivered through a wage. We assume that the wage remains constant unless it is such that the value to the worker of being employed at the firm falls below the worker's outside option—either the value of unemployment or the worker's marginal value at a poaching firm—or it is such that the worker's value of being employed at the firm exceeds the marginal value of the worker to the firm. In the first case, we assume that the wage is raised so that the value to the worker of being employed at the firm is equal to his outside option. This guarantees that the worker has a private incentive to remain with the firm whenever the gains from trade are positive. In the second case, the wage is lowered so that the value to the worker of being employed at the firm is equal to his marginal value at the firm. This guarantees that the worker has a private incentive to report all outside offers in which the worker is more valuable to the poacher than to the firm.

A few comments about the environment are in order. We assume that a firm can employ at most two workers. In order to capture the idea that workers may learn from their coworkers, we need to assume that firms have multiple employees and, hence, we need

to depart from the standard search-theoretic models of the labor market (e.g. Pissarides 1985, Mortensen and Pissarides 1994, Postel-Vinay and Robin 2002, etc...). In order to keep the number of state variables of a firm to a minimum, we assume that firms have at most two employees. The assumption is obviously unrealistic. While a worker has only one coworker in the model, he typically has several in the data. Nevertheless, if the multiple coworkers of an individual are properly collapsed into a representative coworker, the model can still capture how often an individual is exposed to different groups of coworkers, how long an individual is exposed to different groups of coworkers, and the consequences for the individual from being exposed to different groups of coworkers. That is, even though the model is stylized, it can still reproduce the pattern of interactions between an individual and different groups of coworkers.

We assume that hiring and firing decisions take into account the value of the decision on all the members of the production units affected by the trade. Specifically, we use a notion of gains from trade between a worker and a firm that includes the effect of the trade on the worker’s current employer and coworkers and the effect of the trade on the firm and its current employees. The view behind this assumption is that members of a production unit find a way—through explicit or implicit contractual clauses—to make sure that individuals fully internalize the effect of their employment decisions on each other and, hence, to maximize the joint value of the partnership.¹ The assumption is commonplace in the labor search literature, where firm-worker pairs separate if and only if it is jointly profitable to do so (see, e.g., Pissarides 1985, Mortensen and Pissarides 1994, Postel-Vinay and Robin 2002, etc...).

We assume that the wage earned by a worker at a firm remains constant over time, unless either the worker’s outside option exceeds the worker’s value of employment at the firm, or the firm’s marginal value of the worker falls below the worker’s value of employment at the firm. The assumption implies that a worker’s wage at a particular firm is changed only when doing so is necessary to induce the worker to make the correct allocative decision—i.e. move to a poacher, move into unemployment, or remain with the firm. The assumption is ubiquitous in search models because—due to frictions—there is no constant competitive pressure pushing the worker’s wage towards the worker’s marginal value (e.g. Postel-Vinay and Robin 2002 and Cahuc, Postel-Vinay and Robin 2006). The assumption is realistic, since wages are known to be sticky for workers who stay with the same employer. The assumption plays a key role in the calibration of the model. Indeed, the wage of a worker does not track instantaneously with his human capital growth.

¹Bilal, Engbom, Mongey and Violante (2022) develop an explicit bargaining game that leads to the same notion of “collective” gains from trade as the one that we have adopted.

Instead, the wage of a worker reflects his human capital growth only when the worker is hired for a new job (or when he receives a sufficiently generous outside offers). If the worker is hired from employment, the wage reflects his human capital, the state of the hiring firm, and the state of the old employer. If the worker is hired from unemployment, the wage will reflect only his human capital and the state of the hiring firm.

Lastly, we motivate our decision of including search frictions in the structural model of the labor market, a decision that differentiates this paper from the work of Anderson (2015), Anderson and Smith (2010), Cornelissen, Dustmann and Schoenberg (2018), Nix (2020), and Jarosch, Oberfield and Rossi-Hansberg (2020). Search frictions are critical to make sense of the data. First, search frictions provide a natural way to interpret the history of individual workers. Without frictions, unemployment spells would have to be interpreted as shocks to the value of leisure and transitions from one employer to another employer would be indeterminate. Second, search frictions provide a natural way to interpret the pattern of sorting of workers and coworkers. Without frictions, it would be difficult to explain why workers earning the same wage are seen with all sorts of coworkers. Third, as discussed in the previous paragraph, search frictions affect the relation between the human capital changes of a worker and his wage changes. Without frictions, human capital changes would be instantaneously priced into the wage. In Section 3, we provide evidence showing that search frictions are indeed critical to properly measure the human capital accumulation function.

2.2 Definition of Equilibrium

In order to define an equilibrium, we need to introduce some notation. We denote as U_k the value of unemployment to a worker of type k at the beginning of the production stage. We denote as \tilde{V}_y the value of state y to a firm and its employees at the beginning of the search-and-matching stage. We denote as \hat{V}_y the value of state y to a firm and its employees after the search and matching process has taken place, but before the firm has had the option of firing some of its employees. We denote as V_y the value of state y to a firm and its employees at the beginning of the production stage. We also need some notation to describe the stationary distribution of workers and firms. We denote as $e_{k,x}$ the measure of workers of type k who are in employment state x . We denote as n_y the measure of firms that are in state y . The distributions $\{e, n\}$ are measured at the beginning of the search-and-matching stage.

The value of unemployment to a worker of type k is such that

$$U_k = b_k + \beta \mathbb{E}_{k+} \left\{ U_{k+} + \sigma [0 - U_{k+}] + (1 - \sigma) \left[\sum_y \lambda_u p_y \gamma (v_{k+}(y) - U_{k+})^+ \right] \right\}. \quad (1)$$

where the operator $(x)^+$ is defined as $\max\{x, 0\}$. In the current period, the income of the worker is b_k . At the learning stage of next period, the worker's type becomes k_+ . At the entry-and-exit stage, the worker exits the labor market with probability σ and his continuation value is 0. At the search-and-matching stage, the worker meets a firm in state y with probability $\lambda_u p_y$, with $p_y = n_y/n$. The gains from trade between the worker and the firm are $v_{k_+}(y) - U_{k_+}$, where $v_{k_+}(y)$ denotes the marginal value of the worker to the firm and U_{k_+} is the worker's outside option. Formally, $v_{k_+}(y)$ equals $\hat{V}_{k_+,0} - \hat{V}_{0,0}$ if $y = (0, 0)$, $\hat{V}_{i,k_+} - \hat{V}_{i,0}$ if $y = (i, 0)$ with $i \in K$, and $\max\{\hat{V}_{i,k_+} + U_j, \hat{V}_{j,k_+} + U_i\} - \hat{V}_{i,j}$ if $y = (i, j)$ with $i, j \in K$ since the firm would have to fire one of its employees in order to hire the worker. If the gains from trade are positive, the worker is hired and his continuation value is given by his outside option plus a fraction γ of the gains from trade. If the gains from trade are negative, the worker remains unemployed and his continuation value is U_{k_+} .

At the beginning of the production stage, the value of state y to a firm and its employees is such that

$$\begin{aligned} V_{0,0} &= 0 + \beta \tilde{V}_{0,0}, \quad V_{k,0} = f(k, 0) + \beta \mathbb{E}_{k_+} \left\{ \sigma \tilde{V}_{0,0} + (1 - \sigma) \tilde{V}_{k_+,0} \right\}, \\ V_{k,\ell} &= f(k, \ell) + \beta \mathbb{E}_{k_+, \ell_+} \left\{ \sigma(1 - \sigma) \left(\tilde{V}_{k_+,0} + \tilde{V}_{\ell_+,0} \right) + \sigma^2 \tilde{V}_{0,0} + (1 - \sigma)^2 \tilde{V}_{k_+, \ell_+} \right\}. \end{aligned} \quad (2)$$

Consider a firm in state $(0, 0)$. In the current period, the firm's profit is equal to its output $f(0, 0) = 0$. In the next period, the firm enters the search-and-matching stage in state $(0, 0)$ and its continuation value is $\tilde{V}_{0,0}$. Next, consider a firm in state $(k, 0)$. In the current period, the sum of the firm's profit and the employee's wage is equal to the output $f(k, 0)$. At the learning stage of next period, the employee's type becomes k_+ . At the entry-and-exit stage, the employee exits the labor market with probability σ , in which case the firm's continuation value is $\tilde{V}_{0,0}$, the employee's continuation value is 0 and, hence, the joint continuation value is $\tilde{V}_{0,0}$. The employee stays in the labor market with probability $1 - \sigma$, in which case the joint continuation value of the firm and the employee is $\tilde{V}_{k_+,0}$. Lastly, consider a firm in state (k, ℓ) . In the current period, the sum of firm's profit and the employees' income is equal to the output $f(k, \ell)$. At the learning stage of next period, the employees' types become k_+ and ℓ_+ . At the exit stage, the employee of type k_+ exits and the employee of type ℓ_+ stays in the labor market with probability $\sigma(1 - \sigma)$, in which case the joint continuation value is $\tilde{V}_{\ell_+,0}$. The employee of type k_+ stays and the employee of type ℓ_+ exits with probability $\sigma(1 - \sigma)$, in which case the joint continuation value is $\tilde{V}_{k_+,0}$. Both employees exit with probability σ^2 , in which case the joint continuation value is $\tilde{V}_{0,0}$, and both employees stay with probability $(1 - \sigma)^2$, in which case the joint continuation value is \tilde{V}_{k_+, ℓ_+} .

At the beginning of the search-and-matching stage, the value of state $(0, 0)$ to a firm is such that

$$\tilde{V}_{0,0} = \hat{V}_{0,0} + \left[\sum_{i,x} q_i(x) (1 - \gamma) (v_i((0, 0)) - z_i(x))^+ \right]. \quad (3)$$

The firm meets a worker of type i in state x with probability $q_i(x)$, where $q_i(u)$ equals $\lambda_u e_{i,u}/n$, $q_i(0)$ equals $\lambda_e e_{i,0}/n$ and $q_i(j)$ equals $\lambda_e e_{i,j}/n$. The gains from trade between the firm and the worker are $v_i((0, 0)) - z_i(x)$, where $v_i((0, 0))$ is the marginal value of the worker to the firm and $z_i(x)$ denotes the outside option of the worker. Formally, $z_i(x)$ equals U_i if $x = u$, $\hat{V}_{i,0} - \hat{V}_{0,0}$ if $x = 0$, and $\hat{V}_{i,j} - \hat{V}_{j,0}$ if $x = j \in K$. If the gains from trade are positive, the firm hires the worker and its continuation value is given by its outside option $\hat{V}_{0,0}$ plus a fraction $1 - \gamma$ of the gains from trade. If the gains from trade are negative, the firm does not hire the worker and its continuation value is $\hat{V}_{0,0}$.

At the beginning of the search-and-matching stage, the value of state $(k, 0)$ to a firm and its employee is such that

$$\begin{aligned} \tilde{V}_{k,0} = & \hat{V}_{k,0} + \delta[\hat{V}_{0,0} + U_k - \hat{V}_{k,0}] \\ & + \sum_{i,x} q_i(x) (1 - \gamma) (v_i((k, 0)) - z_i(x))^+ + \sum_y \lambda_e p_y \gamma (v_k(y) - z_k(0))^+ \end{aligned} \quad (4)$$

The employee moves into unemployment for exogenous reasons with probability δ . In this case, the firm's continuation value is $\hat{V}_{0,0}$, the employee's continuation value is U_k and, hence, the joint continuation value is $\hat{V}_{0,0} + U_k$. The firm meets a worker of type i in state x with probability $q_i(x)$. The gains from trade between the firm and the worker are $v_i((k, 0)) - z_i(x)$, where $v_i((k, 0))$ is the marginal value of the worker to the firm and $z_i(x)$ is the outside option of the worker. If the gains from trade are positive, the firm hires the worker. In this case, the continuation value to the firm and its employee is given by the firm's outside option $\hat{V}_{k,0}$ plus a fraction $1 - \gamma$ of the gains from trade. If the gains from trade are negative, the worker is not hired and the continuation value to the firm and its employee is $\hat{V}_{k,0}$. Similarly, at the search-and-matching stage, the employee meets a poaching firm in state y with probability $\lambda_e p_y$. The gains from trade between the poaching firm and the employee are $v_k(y) - z_k(0)$. If the gains from trade are positive, the employee is poached. In this case, the employee continuation value is equal to his outside option $z_k(0) = \hat{V}_{k,0} - \hat{V}_{0,0}$ plus a fraction γ of the gains from trade and the firm's continuation value is $\hat{V}_{0,0}$. If the gains from trade are negative, the employee is not poached and the joint continuation value is $\hat{V}_{k,0}$.²

At the beginning of the search-and-matching stage, the value of state (k, ℓ) to a firm

²As can be inferred from (4) and (5), we assume that different events at the search-and-matching stage are mutually exclusive for a given production unit. That is, either an employee moves into unemployment, an employee meets a poaching firm, or the firm meets a worker. The assumption would be without loss in generality in a continuous-time version of the model.

and its employees is such that

$$\begin{aligned}\tilde{V}_{k,\ell} = \hat{V}_{k,\ell} & + \delta(\hat{V}_{k,0} + U_\ell - \hat{V}_{k,\ell}) + \delta(\hat{V}_{\ell,0} + U_k - \hat{V}_{k,\ell}) \\ & + \sum_y \lambda_e p_y \gamma (v_k(y) - z_k(\ell))^+ + \sum_y \lambda_e p_y \gamma (v_\ell(y) - z_\ell(k))^+ \\ & + \sum_{i,x} q_i(x) (1 - \gamma) (v_i((k, \ell)) - z_i(x))^+.\end{aligned}\quad (5)$$

Each employee has a probability δ of moving into unemployment. Each employee has a probability $\lambda_e p_y$ of meeting a poaching firm in state y . If, for instance, the employee of type k meets a poaching firm in state y , the gains from trade between the employee and the poaching firm are $v_k(y) - z_k(\ell)$. If the gains from trade are positive, the employee moves to the poaching firm and his continuation value is equal to $z_k(\ell) = \hat{V}_{k,\ell} - \hat{V}_{\ell,0}$ plus a fraction γ of the gains from trade, while the joint continuation value to the firm and the other employee is $\hat{V}_{\ell,0}$. The firm has a probability $q_i(x)$ of meeting a worker of type i in state x . The gains from trade between the firm and the worker are given by the marginal value $v_i((k, \ell))$ of the worker to the firm—which equals $\max\{\hat{V}_{k,i} + U_\ell, \hat{V}_{\ell,i} + U_k\} - \hat{V}_{k,\ell}$ as the firm needs to fire one of its employees to hire the worker—net of the worker's outside option $z_i(x)$. If the gains from trade are positive, the worker is hired and the firm captures a fraction $1 - \gamma$ of the gains from trade.

After the search-and-matching process is complete but before the firm has had the option of firing, the value of state y to a firm and its employees is such that

$$\begin{aligned}\hat{V}_{0,0} &= V_{0,0}, \quad \hat{V}_{k,0} = \max\{V_{0,0} + U_k, V_{k,0}\}, \\ \hat{V}_{k,\ell} &= \max\{V_{k,0} + U_\ell, V_{\ell,0} + U_k, V_{0,0} + U_k + U_\ell, V_{k,\ell}\}.\end{aligned}\quad (6)$$

If the production unit is in state $(0, 0)$, the firm cannot fire any employee and $\hat{V}_{0,0} = V_{0,0}$. If the production unit is in state $(k, 0)$, the firm may keep its employee, in which case their continuation value is $V_{k,0}$, or it may fire its employee, in which case their continuation value is $V_{0,0} + U_k$. If the production unit is in state (k, ℓ) , the firm may keep both employees, fire the employee of type k , fire the employee of type ℓ , or fire both of them. In the first case the joint continuation value is $V_{k,\ell}$, in the second case it is $V_{\ell,0} + U_k$, in the third case it is $V_{k,0} + U_\ell$, and in the last case it is $V_{0,0} + U_k + U_\ell$.

The values U_k , V_y and \hat{V}_y determine the gains from trade between workers and firms and, hence, the firms' hiring and firing policies. The firms' hiring and firing policies determine the laws of motion for workers and for firms and, hence, the stationary distribution of workers and firms $e_{k,x}$ and n_y . A Stationary Equilibrium is such that the values $\{U_k, V_y, \hat{V}_y\}$ satisfy the Bellman equations (1)-(6) given the distributions $\{e_{k,x}, n_y\}$, and the distributions $\{e_{k,x}, n_y\}$ are stationary given the firms' hiring and firing policies implied by the values $\{U_k, V_y, \hat{V}_y\}$. We relegate the stationarity conditions for the distributions

$\{e_{k,x}, n_y\}$ to Appendix A, as they are both straightforward and cumbersome. We relegate the equilibrium conditions for wages to Appendix B, as they are not an integral part of the definition of equilibrium.

3 Data and Measurement

In this section, we calibrate the structural model of the labor market using a large matched employer-employee dataset of US workers and firms. In Section 3.1, we present some new empirical evidence on the relation between an individual’s wage growth and the average wage of his previous coworkers, as well as on the pattern of sorting between workers and coworkers. In Section 3.2, we lay out and discuss the calibration strategy. In Section 3.3, we report and discuss the calibration outcomes.

3.1 Data and Evidence

In order to calibrate a structural model of production and learning in teams, we need information about the human capital growth of a worker of type k when employed with coworkers of type ℓ , and about the gains from trade between a worker of type k and a firm that employs workers of type ℓ . Information about the human capital growth of a particular type of worker when he is employed with different types of coworkers allows us to recover the human capital accumulation function g . Information about the gains from trade between a particular type of worker and firms with different types of employees allows us to recover the production function f . Since human capital is not observable, we use wages as a noisy measure of human capital and use the model to translate wages into human capital. Since individual workers have several coworkers rather than just one, we use averages to collapse multiple coworkers into a representative one.

Our main source of information is the Longitudinal Employer-Household Dynamics (LEHD) dataset. The LEHD is a matched employer-employee dataset that covers 95% of jobs in the US private sector. We have access to the LEHD between 1998 and 2014 across 24 US States, which include California, Illinois, Kansas, Ohio and Pennsylvania and represent about 50% of the total population. For each individual in our data, we observe some of their demographic characteristics (e.g., age and gender), the identity of their employer, some characteristics of their employer (e.g., number of workers, average wage of workers, industry), and their labor earnings in every quarter (which we shall refer to as their wage). We also link the LEHD to the 2000 decennial Census. The decennial Census provides information on education and occupation for about $1/6^{th}$ of the individuals in the LEHD.

We define the *primary employer* of individual i in a given year t as the firm j from which i received the highest labor earnings during year t . We define an individual i as *unemployed* in a given quarter of year t if i earned less than \$1,000 in that quarter, where labor earnings are measured in 2014 dollars. We define an individual i as *fully employed* in year t if i earned at least \$1,000 in every quarter of year t . Similarly, we define an individual i as a *full-year employee* of firm j in year t if i has earned from j at least \$1,000 in every quarter of year t . We define the *stable coworkers* of individual i at firm j in year t as the collection of employees of j who are between 24 and 65 years old, who are full-year employees of j in year t , and who receive some positive earnings from j also in years $t - 1$ and $t + 1$.

Using our extract of the LEHD, we first construct the *E dataset*—a dataset which includes all the individual/year pairs (i, t) that meet the following criteria: (i) individual i in year t is between 24 and 65 years old; (ii) individual i in year t is fully employed; (iii) individual i 's demographic information from the decennial Census is available; (iv) the primary employer of individual i in year t is a single-unit firm j ; (v) individual i has between 1 and 100 stable coworkers at firm j . We impose the last two criteria so as to focus on individuals who are likely to be physically in contact with their coworkers—because the firm has a single unit—and who are likely to interact with most of their coworkers—because there are relatively few of them.

Using the E dataset, we then construct the *EUE sample*—a sample that includes individual/year pairs (i, t) such that the individual i moves from employment at some firm j into unemployment and, then, from unemployment into employment at a different firm j_+ . Formally, in order to be included in the EUE sample, the individual/year pair (i, t) must be such that: (vi) individual i in year t is a full-year employee of firm j ; (vii) individual i is unemployed for at least one quarter in year $t + 1$; (viii) individual i is a full-year employee of a firm $j_+ \neq j$ in year $t + 2$. We impose the additional requirement in order to select transitions of individuals from stable employment at one firm into stable employment at another firm with an intervening spell of unemployment.

We use the EUE sample to construct a measure of the effect that the human capital of coworkers has on the accumulation of human capital of an individual. We measure the effect of coworkers on the human capital accumulation of an individual by looking at the relationship between his wage at the new firm and the wage of his coworkers at the old firm, after controlling for the individual's wage at the old firm. We focus on transitions across firms because, when an individual is hired by a new firm, he must bargain with the new firm over the terms of trade and, hence, his wage must reflect all of the human capital that he has accumulated. In contrast, when an individual remains with the same firm, he

may not be able to bargain over the terms of trade frequently and, hence, his wage need not reflect his recent human capital accumulation. We focus on transitions across firms that have an intervening spell of unemployment because, when an individual is hired out of unemployment, the outcome of the bargain is a wage that is not contaminated by the individual’s marginal value at his old firm.

Specifically, we run the following OLS regression

$$w_{i,t+2} = \phi_0 + \phi_1 w_{i,t} + \phi_2 w_{j,t}^* + \Phi X_{i,j,t} + \epsilon_{i,t}, \quad (7)$$

where $w_{i,t+2}$ denotes the log wage of individual i in year $t + 2$ (i.e., the post-transition wage), $w_{i,t}$ denotes the log wage of worker i in year t (i.e., the pre-transition wage), $w_{j,t}^*$ denotes the average log wage of the stable coworkers of individual i at firm j in year t (i.e., the pre-transition coworkers’ wage), and $X_{i,j,t}$ denotes a dummy for calendar year, a dummy for State, a dummy for the 1-digit Standard Industrial Classification (SIC) code of firm j , and dummies for the race and gender of individual i . We run the regression separately for individuals whose wage $w_{i,t}$ is lower and higher than the average wage $w_{j,t}^*$ of their coworkers, since it is natural to conjecture that the effect of coworkers on an individual’s human capital growth is different depending on whether the coworkers have more or less human capital than the individual.

Table 1 reports the estimates of the OLS coefficients in (7) for the subset of individuals whose wage is lower than the wage of their coworkers (column 1) and for the subset of individuals whose wage is higher than the wage of their coworkers (column 2). The main coefficient of interest is ϕ_2 , the coefficient that measures the relation between the log wage of the coworkers of the individual and the log wage of the individual in his next job. The estimate of ϕ_2 is 0.123 for individuals whose wage is lower than the wage of their coworkers, and 0.045 for individuals whose wage is higher than the wage of their coworkers. The estimates imply that, for an individual who earns less than his coworkers, a 10% increase in the coworker’s wage forecasts a 1.23% higher wage in his next job. For an individual who earns more than his coworkers, a 10% decline in the coworker’s wage forecasts only a 0.45% lower wage in his next job.³ The estimates do not have an immediate structural interpretation because wages are a noisy measure of human capital (more on this issue below). At face value, however, the estimates do suggest that, conditional on the individual being less knowledgeable than his coworkers, an increase in

³In Appendix D, we show that these estimates are robust. We show that the estimates are similar when we use the coworkers’ median wage rather than the average wage. They are similar when we allow the coworkers’ average wage to interact with the size of the firm, when we allow the coworkers’ wage to interact with the growth of the firm, and when we allow the coworkers’ average wage to interact with the duration of the individual’s unemployment spell. The estimates are also similar when we restrict attention to firms with exactly 2 employees.

Table 1: Individual wage regression

	(1)	(2)
Dependent variable	Wage at $t + 2$	Wage at $t + 2$
Sample	$w_{i,t} < w_{j,t}^*$	$w_{i,t} \geq w_{j,t}^*$
Coworker wage at t	0.123 (0.00327)	0.0453 (0.00642)
Individual wage at t	0.513 (0.00337)	0.773 (0.00617)
R-Squared	0.366	0.520
Round N	244000	100000

Notes: SE clustered at SEIN level. Controls include: State, 1-digit SIC, race and gender dummies, and year fixed effects.

the coworkers' human capital speeds up the individual's learning. In contrast, conditional on the individual being more knowledgeable than his coworkers, a decline in the coworkers' human capital does little to slow down the individual's learning.

Exploiting the link between the LEHD and the 2000 decennial Census, we can examine the relationship between an individual's education, occupation, and age and the role played by coworkers in his human capital accumulation process. We find that, for individuals who earn less than their coworkers, an increase in the coworkers' wage has a stronger effect on the wage of the individual in his next job if the individual is more educated (which presumably makes him better at learning), if the individual is in an abstract occupation (which presumably increases the scope for learning), and if the individual is younger (which presumably increases the benefit of learning). These findings corroborate the view that the regression coefficient ϕ_2 , albeit not directly structural, is indeed informative about the role of coworkers in the human capital accumulation process.

We first break down the relation between the wage of an individual in the next job and the wage of his coworkers based on the education of the individual. We classify individuals based on their education: less than high-school (L), high-school (H), some college (S) and college (C). We then run the OLS regression (7) with education fixed-effects and interaction terms between the individual's wage, the coworkers' wage and the education of the individual, using individuals who did not finish high school as the baseline. Table 2 reports the estimates of the coefficients on the interactions between the coworker's wage and the individual's education. For individuals earning less than their coworkers, the coefficient on the coworkers' wage is 0.06 if the individual has a high-school degree or less, 0.08 if the individual has some college education, and 0.16 if the individual has a college degree. That is, the return from having more knowledgeable coworkers is

Table 2: Individual wage regression: Education

	(1)	(2)
Dependent variable	Wage at $t + 2$	Wage at $t + 2$
Sample	$w_{i,t} < w_{j,t}^*$	$w_{i,t} \geq w_{j,t}^*$
Coworker wage at t	0.0585 (0.00913)	0.0442 (0.0254)
Coworker wage at $t \times H$	-0.00280 (0.0109)	-0.00288 (0.0293)
Coworker wage at $t \times S$	0.0257 (0.0109)	-0.0220 (0.0279)
Coworker wage at $t \times C$	0.0992 (0.0139)	0.0387 (0.0276)
R-Squared	0.378	0.529
Round N	244000	100000

Notes: SE clustered at SEIN level. Controls include: State, 1-digit SIC, race and gender dummies, and year fixed effects. Education classifications measured in the 2000 decennial Census. H indicates highschool. S indicates some college. C indicates college or more.

increasing in the education of the individual.

Next, we break down the relation between the wage of an individual in the next job and the wage of his coworkers based on the occupation of the individual. We use the breakdown of tasks in different occupations by Autor and Dorn (2003) to classify occupations into Abstract (A), Manual (M) and Routine (R). Specifically, we classify an occupation as A, M or R as long as its abstract, manual or routine requirements are above the median. We then run the wage regression in (7) with occupation fixed-effects and interaction terms between the individual's wage, the coworkers' wage and the occupation of the individual. Table 3 reports the estimates of the coefficients on the interactions between the coworker's wage and the individual's occupation. For individuals earning less than their coworkers, the coefficient on the coworkers' wage is 0.15 if the individual is an abstract occupation, 0.11 if the individual is in a manual occupation, and 0.08 if the individual is in a routine occupation. That is, the return from having more knowledgeable coworkers is highest for individuals in abstract occupations and lowest for individuals in routine occupations.

Lastly, we examine the role of age. Specifically, we divide individuals into age deciles. We then run the regression (7) with age decile fixed-effects and interaction terms between the individual's wage, the coworkers' wage and the individual's age decile, using the first decile as the baseline. Table 4 reports the estimates of the coefficients on the interactions between the coworker's wage and the individual's age decile. For individuals earning less than their coworkers, the coefficient on the coworkers' wage is strictly decreasing in age.

Table 3: Individual wage regression: Occupation

	(1)	(2)
Dependent variable	Wage at $t + 2$	Wage at $t + 2$
Sample	$w_{i,t} < w_{j,t}^*$	$w_{i,t} \geq w_{j,t}^*$
Coworker wage at t	0.110 (0.00670)	0.0869 (0.0144)
Coworker wage at $t \times A$	0.0420 (0.00709)	0.00749 (0.0139)
Coworker wage at $t \times R$	-0.0259 (0.00667)	-0.0698 (0.0138)
Coworker wage at $t \times M$	0.00243 (0.00688)	-0.0152 (0.0136)
R-Squared	0.372	0.524
Round N	244000	100000

Notes: SE clustered at SEIN level. Controls include: State, 1-digit SIC, race and gender dummies, and year fixed effects. Occupation classifications measured in the 2000 decennial Census. A, R, and M stand for whether the worker's occupation in the 2000 decennial census is above median in terms of abstract, routine, and manual skill requirements.

The coefficient is as high as 0.15 for the youngest workers, and it falls monotonically to 0.10 for the oldest workers. That is, the return from having more knowledgeable coworkers is strictly decreasing in the individual's age.

In all of the above regressions, we followed the literature on peer effects (e.g., Hoxby 2000, Carrell, Sacerdote and West 2013) and assumed that the human capital growth of an individual is affected by the average human capital of his coworkers. For this reason, we used the average wage of coworkers in order to construct the wage of the representative coworker. In reality, though, an individual may interact only with a subset of his peers and, hence, the coworkers' average wage may be a noisy measure of the human capital of the coworkers who affect the individual's human capital growth. In Table 5 we explore this hypothesis. In Columns (1) and (3), we add the standard deviation of the coworkers' wages to the regression in (7). We find that the coefficient on the average of the coworkers' wage increases relative to Table 1, and the coefficient on the standard deviation of coworkers' wages is negative. These findings are consistent with the hypothesis that, when the dispersion in the coworkers' human capital is larger, the coworkers' average wage becomes a noisier measure of the treatment received by the individual, and the coefficient on the coworkers' average wage is attenuated. Similarly, when we add the maximum and the minimum of the coworkers' wages to (7), the coefficient on the coworkers' average wage increases, the coefficient on the maximum wage is negative, and the coefficient on the minimum wage is either positive or small. It should also be noticed that the R^2 of

Table 4: Individual wage regression: Age

Dependent variable	(1) Wage at $t + 2$	(2) Wage at $t + 2$
Sample	$w_{i,t} < w_{j,t}^*$	$w_{i,t} \geq w_{j,t}^*$
Coworker wage at t	0.148 (0.00820)	0.0596 (0.0320)
Coworker wage at $t \times$ Age decile 2	-0.0197 (0.0123)	0.0189 (0.0415)
Coworker wage at $t \times$ Age decile 3	-0.0279 (0.0122)	-0.0154 (0.0378)
Coworker wage at $t \times$ Age decile 4	-0.0259 (0.0133)	-0.0344 (0.0385)
Coworker wage at $t \times$ Age decile 5	-0.0351 (0.0125)	-0.0199 (0.0358)
Coworker wage at $t \times$ Age decile 6	-0.0428 (0.0133)	0.00314 (0.0376)
Coworker wage at $t \times$ Age decile 7	-0.0297 (0.0138)	0.00709 (0.0365)
Coworker wage at $t \times$ Age decile 8	-0.0395 (0.0128)	-0.00877 (0.0361)
Coworker wage at $t \times$ Age decile 9	-0.0397 (0.0131)	-0.0194 (0.0361)
Coworker wage at $t \times$ Age decile 10	-0.0466 (0.0144)	-0.0816 (0.0383)
R-squared	0.372	0.525
Round N	244000	100000

Notes: SE clustered at SEIN level. Controls include: State, 1-digit SIC, race and gender dummies, and year fixed effects.

the regression that include additional moments of the coworkers' wages is essentially the same as the R^2 of the regression that only includes the average of the coworkers' wages. For this reason, we are comfortable using the average wage to summarize an individual's coworkers.

In contrast to the literature on peer effects (e.g., Cornelissen, Dustmann and Schoenberg 2018, Nix 2020, Jarosch, Rossi-Hansberg and Oberfield 2020), our theory explicitly acknowledges the presence of search frictions. Because of search frictions, the wage of a worker who remains with the same employer does not need to immediately reflect changes in the worker's human capital. The wage of a worker who moves to a new employer, however, must reflect changes in the worker's human capital. For this reason, we used the worker's wage in the next job in order to measure the impact of coworkers on human

Table 5: Individual wage regression: Other moments of coworkers' wages

	(1)	(2)	(3)	(4)
Dependent variable	Wage at $t + 2$	Wage at $t + 2$	Wage at $t + 2$	Wage at $t + 2$
Sample	$w_{i,t} < w_{j,t}^*$	$w_{i,t} < w_{j,t}^*$	$w_{i,t} \geq w_{j,t}^*$	$w_{i,t} \geq w_{j,t}^*$
Coworker wage at t	0.191 (0.00481)	0.226 (0.00553)	0.0537 (0.00646)	0.0441 (0.00891)
Individual wage at t	0.496 (0.00351)	0.513 (0.00345)	0.801 (0.00682)	0.789 (0.00667)
Std of coworker wages, t	-0.0410 (0.00206)		-0.0306 (0.00309)	
Min of coworker wage at t		-0.0222 (0.00276)		0.0306 (0.00415)
Max of coworker wage at t		-0.0684 (0.00264)		-0.0287 (0.00530)
R-squared	0.367	0.368	0.521	0.521
Round N	244,000	244,000	100,000	100,000

Notes: SE clustered at SEIN level. Controls include: State, 1-digit SIC, race and gender dummies, and year fixed effects. Minimum, maximum, and standard deviation of coworkers wages are computed among stable coworkers at t .

capital accumulation. Table 6 supports our choice. Table 6 reports the coefficients on a version of the OLS regression (7) in which we focus on workers who remain employed at firm j from year t to year $t + 2$. The coefficients on the coworkers' wage in Table 6 are much lower than in Table 1. For workers earning less than their coworkers, a 10% increase in the coworkers' wage forecasts only a 0.6% increase in the wage of the individual in year $t + 2$.

It is worth stressing that, even when one uses the wage changes of job-switchers, the regression coefficient ϕ_2 still does not have an immediate structural interpretation. The reason is simple. A worker's wage is a noisy measure of his human capital. Unless the pattern of sorting of workers and coworkers is random, the coworkers' wage is also a noisy measure of the worker's human capital and, hence, it forecasts the worker's wage in the next job. If the pattern of sorting is positive, a higher coworkers' wage forecasts a higher wage for the worker in his next job. If the pattern of sorting is negative, a higher coworkers' wage forecasts a lower wage for the worker in his next job. This is a typical problem in measuring peer effects, not only in the workplace but also in the classroom. To address this problem, we do not seek quasi-natural experiments in which sorting of workers and coworkers is presumed to be random. Instead, we use the model to translate the regression coefficient into a structural parameter of the human capital accumulation function.

Table 6: Individual wage regression: Stayers

	(1)	(2)
Dependent variable	Wage at $t + 2$	Wage at $t + 2$
Sample	$w_{i,t} < w_{j,t}^*$	$w_{i,t} > w_{j,t}^*$
Coworker wage at t	0.0600 (0.000418)	0.0301 (0.000423)
Individual wage at t	0.908 (0.000438)	0.950 (0.000391)
R-Squared	0.831	0.872
Round N	6.732e+06	5.392e+06

Notes: SE clustered at SEIN level. Controls include: State, 1-digit SIC, race and gender dummies, and year fixed effects.

The discussion above makes it clear the model needs to reproduce the actual pattern of sorting of workers and coworkers in order to properly translate the regression coefficient into a structural parameter of the human capital accumulation function. Moreover, the model needs to reproduce the actual pattern of sorting of workers and coworkers because the pattern of sorting is the key source of information about the gains from trade between different types of workers and firms with different types of employees and, hence, on the shape of the production function.

We construct two measures of sorting, one based on the EUE sample and another one based on the cross-section of workers and coworkers in the whole economy. Using the EUE sample, we run the following OLS regression

$$w_{j+,t+2}^* = v_0 + v_1 w_{i,t} + v_2 w_{j,t}^* + \Upsilon X_{i,j,t} + \varepsilon_{i,t}, \quad (8)$$

where $w_{j+,t+2}^*$ denotes the average log wage of the stable coworkers of individual i in year $t + 2$ (i.e., the post-transition coworkers' wage), $w_{i,t}$ denotes the log wage of individual i in year t (i.e., the pre-transition wage), $w_{j,t}^*$ denotes the average log wage of the stable coworkers of individual i in year t (i.e., the pre-transition coworkers' wage), and $X_{i,j,t}$ denotes the same set of dummies as in (7). We run the regression separately for individuals whose wage $w_{i,t}$ is lower and higher than the average wage $w_{j,t}^*$ of their coworkers.

Table 7 reports the estimates of the OLS coefficients in (8) for the subset of individuals whose wage is lower than the wage of their coworkers (column 1) and for the subset of individuals whose wage is higher than the wage of their coworkers (column 2). The main coefficient of interest is v_1 , the coefficient that measures the relationship between the log wage of an individual and the log wage of his coworkers in the next job. The estimate of v_1 is 0.25 for individuals whose wage is lower than the wage of their coworkers, and 0.17

Table 7: Coworker wage regression

	(1)	(2)
Dependent variable	Coworker wage at $t+2$	Coworker wage at $t+2$
Sample	$w_{i,t} < w_{j,t}^*$	$w_{i,t} \geq w_{j,t}^*$
Coworker wage at t	0.299 (0.00369)	0.392 (0.00719)
Individual wage at t	0.252 (0.00344)	0.172 (0.00648)
R-squared	0.251	0.295
Round N	244,000	100,000

Notes: SE clustered at SEIN level. Controls include: State, 1-digit SIC, race and gender dummies, and year fixed effects.

for individuals whose wage is higher than the wage of their coworkers. These estimates imply that, for an individual who earns less than his coworkers, a 10% increase in his wage forecasts a 2.5% higher coworkers' wage in the next job. For an individual who earns more than his coworkers, a 10% increase in his wage forecasts a 1.7% higher wage for his coworkers in the next job. These estimates reveal that sorting tends to be weakly positive.

We also employ a measure of sorting based on the whole economy. Song et al. (2019) split the cross-sectional variance of wages into a between-firm component—defined as the variance of the average wage across different firms—and a within-firm component—defined as the average variance of wages within firms. Song et al. (2019) document that the between-firm component accounts for about 40% of the cross-sectional variance of wages, while the within-firm component accounts for the remaining 60%. This variance decomposition is a measure of sorting of workers across firms. Indeed, if sorting was strongly positive, firms would employ workers with similar human capital and wages and most of the cross-sectional variance of wages would be due to the between-firm component. If, on the other hand, sorting was strongly negative, firms would employ workers with different human capital and most of the cross-sectional variance of wages would be due to the within-firm component.

3.2 Calibration strategy

Before discussing our calibration strategy, it is useful to review the fundamentals of the model. The entry-and-exit process is described by the probability σ with which a worker permanently exits the labor market and by the probability distribution π from which a worker who enters the labor market draws his initial human capital. We assume that

the human capital ladder $H = \{h_1, h_2, \dots, h_N\}$ has equally spaced rungs with $h_1 = 1$, $h_N = 5$ and $N = 7$. We assume that the probability distribution π is an exponential with coefficient χ appropriately discretized over the human capital ladder H .

The production process is described by the function f . We assume that

$$f(k, \ell) = A \left(\frac{h_k^\rho + h_\ell^\rho}{2} \right)^{1/\rho}. \quad (9)$$

In words, the output produced by a firm employing a worker of type k and a worker of type ℓ is a CES function of the human capital of the two employees. If the parameter ρ is smaller than 1, the production function is supermodular with respect to the human capital of the two employees. If ρ is equal to 1, the production function is modular in the human capital of the two employees. If ρ is greater than 1, the production function is submodular with respect to the human capital of the two employees. Moreover, we assume that $f(k, 0) = ((k^\rho + 1^\rho)/2)^{1/\rho}$. In words, a firm only employing a worker of type k produces a fraction $1/A$ of the output that it would produce by employing a worker of type k and a worker of type 1.⁴

The human capital accumulation process is described by the probability distribution function g . For an employed worker, we assume that

$$g_k(k+1|\ell) = \alpha_0 + \alpha_1 \frac{(h_\ell - h_k)^+}{(h_N - h_1)}, \quad (10)$$

and $g_k(k|\ell) = 1 - g_k(k+1|\ell)$ for $k = 1, 2, \dots, N-1$. In words, we assume that a worker of type k who is employed with a coworker of type ℓ climbs one rung of the human capital ladder with probability $\alpha_0 + \alpha_1 (h_\ell - h_k)^+ / (h_N - h_1)$ and remains on the same rung with complementary probability. The expression in (10) is easy to understand. If the worker is employed with someone who is more knowledgeable than he is, the worker climbs the human capital ladder with probability $\alpha_0 + \alpha_1 (h_\ell - h_k) / (h_N - h_1)$. If the worker is employed with someone who is less knowledgeable than he is, the worker climbs the human capital ladder with probability α_0 . The parameter α_0 can be interpreted as the contribution of learning by doing to human capital growth. The parameter α_1 can be interpreted as the contribution of learning from a more knowledgeable coworker to human capital growth.⁵ Following this interpretation, we assume that $g_k(k+1|0) = \alpha_0$

⁴We experimented with some alternative specifications of $f(k, 0)$. For example, we considered $f(k, 0) = f(k, k)/2$, according to which the output produced by a firm employing one worker of type k is equal to the output that would be produced by assigning half of the worker's time to one job and the other half to the other job. We found that the main results of the paper are not very sensitive to the exact specification of $f(k, 0)$, as long as $f(k, 0) + f(\ell, 0) < f(k, \ell)$. That is, as long as the production function is such that it is better to create teams than keeping workers alone.

⁵We assume that the law of motion for the human capital of a worker who is employed with a less

and $g_k(k|0) = 1 - \alpha_0$. In words, a worker employed on his own only learns by doing. Naturally, for workers who have reached the highest rung of the human capital ladder, $g_N(N|\ell) = g_N(N|0) = 1$.

For an unemployed worker, we assume

$$g_k(k-1|u) = \alpha_u, \quad (11)$$

and $g_k(k|u) = 1 - g_k(k-1|u)$ for $k = 2, 3, \dots, N$. In words, we assume that a worker who is unemployed descends one rung of the human capital ladder with probability α_u , and remains on the same rung with complementary probability. The parameter α_u captures the notion of human capital depreciation during unemployment. Naturally, for workers who are at the lowest rung of the human capital ladder, $g_1(1|u) = 1$.

The search-and-matching process is described by the probability with which an unemployed worker meets a firm, λ_u , the probability with which an employed worker meets a firm, λ_e , the probability with which an employed worker becomes unemployed for exogenous reasons, δ , and the measure of firms in the labor market, $n = 1$. The bargaining process is described by the worker's bargaining power γ . The worker's preferences are described by the discount factor β and by the flow value of unemployment b_k , which we assume to be equal to a fraction b of the worker's human capital h_k .

There are 13 parameters that need to be pinned down and the length of a period to be chosen. We choose the length of a period to be 1 month. We preset 3 of the 13 parameters. Namely, we set the discount factor β to 0.992, so that workers discount future income at the rate of 10% per year.⁶ We set the exit probability σ to 0.002, so that workers remain in the labor force for an average of 35 years. We set the worker's bargaining power γ to $1/2$, so that a worker and a firm share the gains from trade equally. The remaining 10 parameters are jointly calibrated so as to minimize the distance with respect to 18 empirical targets. Specifically, the empirical targets are: (i-iv) the estimates of coefficients ϕ_1 and ϕ_2 in the regression (7); (v-viii) the estimates of coefficients v_1 and

knowledgeable coworker does not depend on the coworker's human capital. The assumption is essentially without loss in generality. In an earlier version of the paper (Herkenhoff et al. 2018) and in a robustness check for this version, we let $g_k(k+1|\ell) = \alpha_0 + \alpha_1(h_\ell - h_k)^+ / (h_N - h_1) + \alpha_2(h_k - h_\ell)^+ / (h_N - h_1)$, where α_2 denotes the effect of less knowledgeable coworkers on the human capital growth of an individual. We found that the distance between the empirical targets and their model-generated analogues is minimized for a value of α_2 that is very close to zero. This finding is not surprising, since the regression coefficient on the coworkers' wage is close to zero for individuals earning more than their coworkers.

⁶In our model, as in other models in the style of Postel-Vinay and Robin (2002), the discount factor β determines the extent to which workers are willing to take lower wages now in exchange for higher wages in the future. If we allowed for the worker's utility function to be concave, the trade-off between current and future wages would be affected by the risk-aversion parameter. Since, for the sake of tractability, we assume that the worker's utility function is linear, we load the role of risk-aversion on the discount factor and, for this reason, choose a β that is lower than usual.

v_2 in the regression (8); (ix) the average wage change for individuals going through an EUE transition; (x) the 90-10 percentile ratio in the cross-sectional wage distribution; (xi) the within and between firm components of cross-sectional wage dispersion; (xii) the lifecycle growth of average wages; (xiii) the average growth of individual wages; (xiv) the 90-10 percentile ratio in the wage distribution of young workers; (xv-xvii) the average rate at which workers transit between employment, unemployment and across different employers; (xviii) the average ratio between individuals' flow unemployment value and wage.

Even though the parameters are jointly calibrated, there is a clear mapping between each one of the parameters and a particular empirical target. We start by discussing the identification of the parameters that describe the human capital accumulation process. The parameter α_1 is identified by the coefficient on the coworkers' wage in the regression (7) of an individual's wage after an EUE transition, conditional on the individual earning less than his coworkers. Intuitively, the higher is α_1 , the stronger is the effect of having more knowledgeable coworkers on the individual's learning and, hence, the higher the coefficient on the coworkers' wage in the regression (7) should be. The parameter α_0 is identified by the extent of wage growth over the life-cycle, which we measure using the Current Population Survey (CPS) as the ratio between the average wage of workers aged 54 and the average wage of workers aged 24.⁷ Intuitively, the higher is the parameter α_0 , the faster workers accumulate human capital over the life-cycle and, hence, the faster wages should grow over the life-cycle. The parameter α_u is identified by the average wage change for workers who go through an EUE transition. Intuitively, the higher is α_u , the faster workers lose human capital during an unemployment spell and, hence, the lower should their wage be after the unemployment spell. The parameter χ , which controls the shape of the distribution of human capital for workers who enter the labor market, is identified by the 90-10 percentile ratio in the wage distribution for workers aged 24. In Appendix C, we show that, indeed, the parameters α_0 , α_1 , α_u and χ have the expected effect on the model-generated version of the empirical targets.

Next, we discuss the identification of the parameters that describe the production process. The parameter ρ is identified by the empirical pattern of sorting of workers and coworkers. Indeed, the lower is ρ , the stronger is the supermodularity between the human capital of workers and coworkers and, in turn, the more positive the pattern of sorting should be. As discussed in the previous subsection, we use two measures of sorting. The first measure of sorting is the coefficient on the individual's wage in the regression (8) of the wage of his coworkers after an EUE transition. The second measure of sorting is

⁷We assume that the workers are 21 years old when they enter the labor market.

the fraction of the cross-sectional wage variance that is accounted for by differences in the average wage at different firms. The parameter A is identified by the 90-10 percentile ratio in the cross-sectional wage distribution. Intuitively, an increase in A increases the difference between the marginal value of employed workers. In Appendix C, we show that, indeed, the parameters ρ and A have the expected effect on the model-generated version of the empirical targets.

The identification of the parameters that describe the search-and-matching process is simple and standard. The parameter λ_u is identified by the rate at which unemployed workers move into employment (UE rate). The parameter λ_e is identified by the rate at which employed workers move from one employer to another without an intervening spell of unemployment (EE rate). The parameter δ is identified by the rate at which employed workers move into unemployment (EU rate). We measure the UE, EE and EU rates in the CPS rather than in the LEHD. We do so because the LEHD contains quarterly information that is ill-suited for measuring monthly transition rates. The identification of the preference parameter b is also standard. We choose the parameter b so that the flow value of unemployment of a worker is equal, on average, to about 71% of the worker's previous wage—a percentage that Hall and Milgrom (2008) argue is appropriate for the US labor market and that has since been widely adopted as a calibration target.

3.3 Calibration outcomes and validation

The calibrated model matches reasonably well the 18 empirical targets, even though it has only 10 parameters. Table 8 shows that the model closely matches the empirical targets that identify the human capital accumulation process. Indeed, the model closely matches the estimate of the coefficient ϕ_2 on the wage of the individual's coworkers in a regression of the wage of the individual after an EUE transition, the average wage change of an individual after an EUE transition, and the lifecycle growth of average wages. The model also does a good job at matching the empirical targets that describe the sorting of workers and coworkers across firms and, in turn, identify the production function. Indeed, the model matches well the estimate of the coefficient v_1 on the wage of the individual in a regression of the wage of the coworkers of the individual after an EUE transition, and the share of the cross-sectional wage variance that is due to wage differences between firms. Lastly, the model does a good job at replicating the empirical targets that identify the search-and-matching process.

In calibrating the model, we did not use any evidence on the rate at which different types of workers leave different types of firms. We can use such evidence to further validate the model. We construct measures of EU and EE rates for different types of workers at

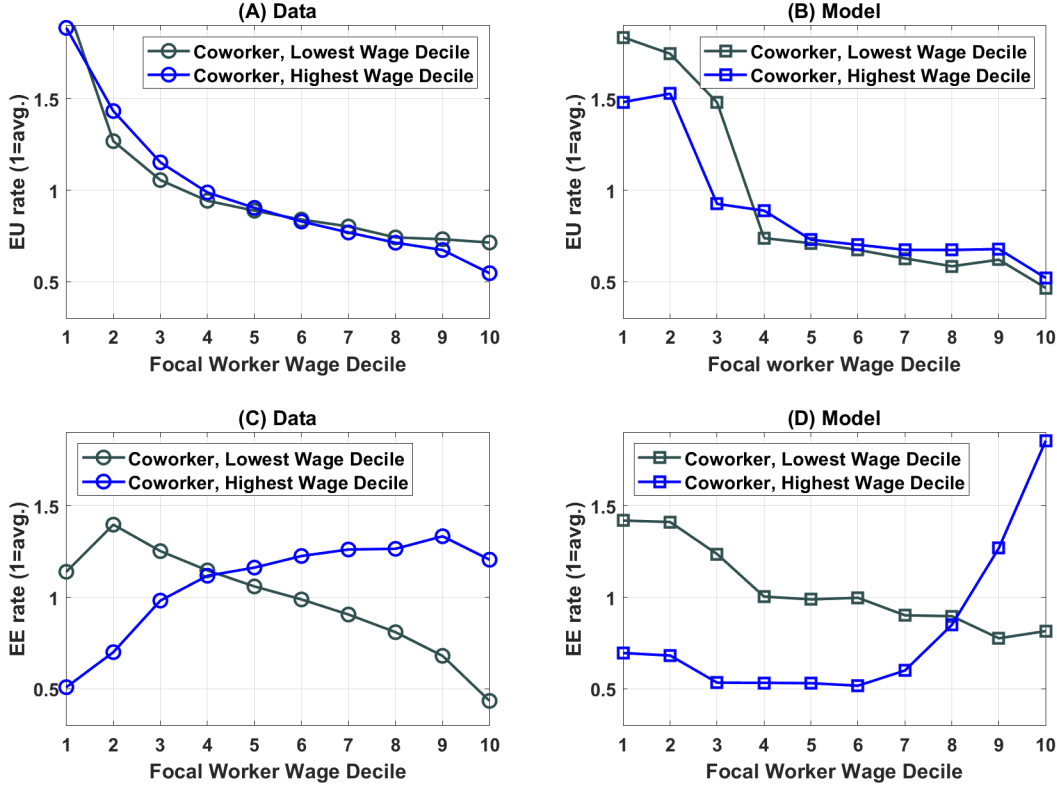
Table 8: Calibration targets and model fit

Target	Source	Data	Model
$\phi_1, w_{it} \leq w_{jt}^*$	LEHD	0.51	0.52
$\phi_1, w_{it} > w_{jt}^*$	LEHD	0.77	0.78
$\phi_2, w_{it} \leq w_{jt}^*$	LEHD	0.12	0.12
$\phi_2, w_{it} > w_{jt}^*$	LEHD	0.05	0.03
Between-firm wage variance	Song et al. (2019)	0.40	0.44
$v_1, w_{it} \leq w_{jt}^*$	LEHD	0.25	0.32
$v_1, w_{it} > w_{jt}^*$	LEHD	0.17	0.22
$v_2, w_{it} \leq w_{jt}^*$	LEHD	0.30	0.12
$v_2, w_{it} > w_{jt}^*$	LEHD	0.39	0.02
EUE average wage loss	LEHD	-0.18	-0.22
54-to-24 y.o. wage ratio	CPS	1.88	1.93
Mean wage growth	CPS	0.02	0.02
p90/p10 wage ratio	CPS	4.23	3.50
p90/p10 wage ratio 24 y.o.	CPS	2.77	2.07
UE Rate	CPS	0.22	0.24
EE Rate	CPS	0.02	0.01
EU Rate	CPS	0.01	0.01
Flow unemployment value	Hall & Milgrom (2008)	0.71	0.73

different types of firms using the LEHD. We consider all the workers in the E dataset. We compute a yearly EU rate as the fraction of these workers who move from firm j to a full-quarter of unemployment during year $t + 1$. We compute a yearly EE rate as the fraction of these workers who move from full-year employment at firm j to some other firm $j_+ \neq j$ in year $t + 1$, without an intervening quarter of unemployment. We then break down the EU and EE rates for workers and coworkers at different wage deciles. Figure 1 shows that the EU rate declines dramatically with the worker's wage decile. Workers at the lowest wage decile have an EU rate that is 2 times higher than the average EU rate. Workers at the highest wage decile have an EU rate that is 50% smaller than the average EU rate. Conditional on the worker's wage, the EU rate is almost invariant to the wage of the coworkers. These findings suggest that low-human capital workers are always more likely to be replaced by a new hire. Figure 1 also shows that the EE rate is decreasing in the worker's wage when the worker has coworkers at the lowest wage decile. In contrast, the EE rate is increasing in the worker's wage when the worker has coworkers at the highest wage decile. These findings suggest that the gains from trade between a worker and a firm are larger when the human capital of the worker and the human capital of his coworkers are different rather than similar.

Figure 1 shows that the calibrated model does a good job at reproducing the pattern

Figure 1: Validation: EU and EE Rates



of transitions of different workers out of different firms. The finding provides validation for the calibrated production and human capital accumulation functions—which together determine the gains from trade between workers and firms and, in turn, the pattern of transitions— as well as for the overall architecture of the structural model. Moreover, the finding implies that, even though the structural model is simple and unrealistic (a worker has a single coworker), it can still reproduce the consequences for an individual from being matched with different coworkers (the effect of treatments), the frequency with which an individual is matched with different coworkers (the frequency of treatments), and how long an individual stays with different coworkers (the duration of treatments). For this reason, the model should provide sensible estimates of the human capital growth and the productivity of an individual when matched with different coworkers.

The calibrated values of the parameters are reported in Table 9. The calibrated value of α_0 is 0.001 and the calibrated value of α_1 is 0.02. The value of α_0 implies that, for a worker of type 1, the growth rate of human capital due to learning-by-doing is about 1% per year. The value of α_1 implies that, for a worker of type 1 who is employed with a coworker of type 2, the growth rate of human capital due to learning-from-coworkers is about 2.5% per year. Hence, learning from coworkers is potentially a more important

Table 9: Calibrated parameter values

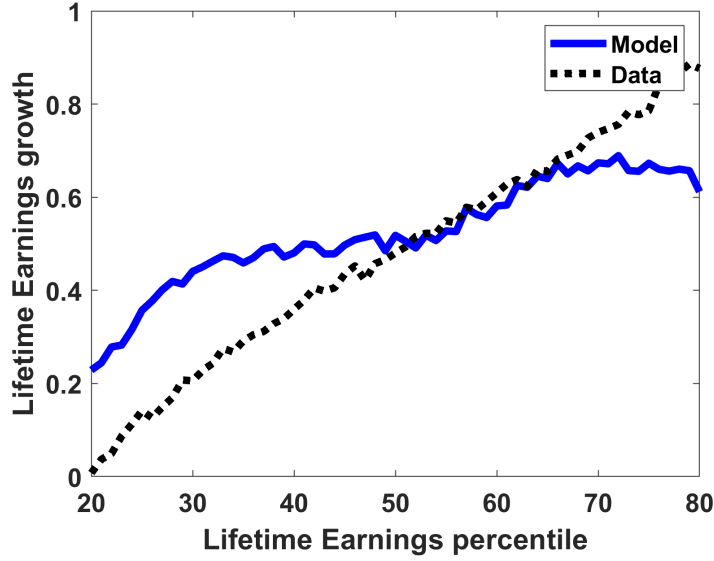
Parameter	Description	Value
α_0	Learning by doing	0.001
α_1	Learning from coworkers	0.020
α_u	Human capital depreciation	0.016
ρ	Production complementarity	0.810
A	Production efficiency	2.194
χ	Entrant distribution	2.623
λ_u	Meeting rate, unemployment	0.340
λ_e	Meeting rate, employed	0.238
δ	Separation rate	0.009
b	Flow value of unemployment	0.976
β	Discount factor	0.992
σ	Exit rate	0.002
γ	Bargaining power worker	0.5
h_N	Human capital ladder	5.0

source of human capital growth than learning by doing. The calibrated value of ρ is 0.8. The value of ρ implies that the production function is supermodular in the human capital of the firm’s employees. Quantitatively, the output produced by two firms that each employ one worker of type 1 and one worker of type 2 is 0.5% lower than the output produced by a firm employing two workers of type 1 and a firm employing two workers of type 2. Hence, the supermodularity of the production function is rather weak.

The reader may be surprised that the calibrated value of α_0 is so low relative to α_1 . These parameter values are those that are needed by the model in order to reproduce the empirical estimate of ϕ_2 in regression (7) and the empirical growth of average wages over the lifecycle. Yet, since α_0 is low relative to α_1 , there may be a concern the model generates a counterfactually large variation in wage growth across individuals, namely very low growth for those unlucky individuals who are never employed with more knowledgeable coworkers, and very high growth for those individuals who spend their life learning from others. In order to address this concern, we confront the model with disaggregated data on wage growth from Guvenen et al. (2021). Figure 2 plots the growth rate of average wages for individuals at different percentiles of the Lifetime Earnings distribution in the data and in the model, where the growth rate of average wages is defined as the difference of log-earnings at age 55 and 25, and Lifetime Earnings are defined as the average earnings over the entire worklife. The model does a decent job in reproducing the lifecycle growth of average wages for workers at different percentiles of the Lifetime Earnings distribution.⁸

⁸Figure 2 is cropped at the 20th percentile on the left and at the 80th percentile on the right. In the

Figure 2: Lifetime Earnings growth



More importantly, the model does not generate excessive variation in wage growth. If anything, the model does not generate quite enough variation in wage growth. This is hardly surprising considering that the model abstracts from heterogeneity in ability to learn on the job, heterogeneity that we ourselves have documented Section 3.1.

A concern may arise that α_1 is high relative to α_0 because we understate the extent of positive sorting in the data and that, hence, to some extent we misinterpret the coefficients in regression (7) as human capital growth rather than sorting. For instance, Freund (2022) shows that, given the same correlation between workers' human capital, the between-firm share of the cross-sectional wage variance may be mechanically lower for firms with two workers than for large firms. In particular, Freund (2022) shows that a between-firm share of wage variance of 0.4 for large firms (what we see in the data) is equivalent to a between-firm share of wage variance of 0.6 for firms with two workers. Even though Freund's correction does not directly apply to our model, we consider an alternative calibration in which we target a between-firm share of the wage variance of 0.6. In the alternative calibration, the production function is slightly more supermodular, in the sense that ρ falls from 0.8 to 0.5. The human capital accumulation function is such that learning from coworkers becomes slightly more important relative to learning on the job, in the sense that α_1 increases from 0.02 to 0.025, while α_0 barely changes. That is, targeting

data, individuals below the 20th percentile have negative wage growth, mainly due to early exit from the labor market. Individuals above the 80th percentile have sky-rocketing wage growth, due to limited labor market participation in the mid-20s and to extremely high earnings later in their 50s. Clearly, our model is not designed to reproduce the phenomena driving the outcomes in the tails of the Lifetime Earnings distribution.

a larger between-firm share of wage variance does not lower α_1 relative to α_0 . There is a simple explanation for this seemingly surprising finding. When we ask the model to generate a larger share of between-firm wage variance, the extent of positive sorting in the cross-section of workers must increase and, all else equal, the coefficient v_1 in (8) increases as well, reflecting an increase in the extent of sorting for workers coming out of an EUE transition. In order to bring the coefficient v_1 back down to its empirical counterpart, the model needs to reduce the extent of positive sorting among workers going through an EUE transition. The model accomplishes this task by making learning from coworkers more important. All details about this alternative calibration can be found in Appendix E.2.

A related concern is that we might be underestimating the extent of positive sorting among workers coming out of unemployment and, in doing so, we might overestimate α_1 . Indeed, one of the most surprising findings in a recent paper by Lentz, Piyapromdee and Robin (2023) is that the extent of sorting for short-tenure workers is not random and, on the contrary, it is already close to the extent of sorting for long-tenure workers. They find that the normalized Mutual Information⁹ for short-tenure workers ranges between 0.08 and 0.12 depending on the year, and that the normalized Mutual Information for long-tenure workers typically falls between 0.1 and 0.15. In our model, the normalized Mutual Information is 0.04 for short-tenure workers, 0.1 for long-tenure workers, 0.05 for workers coming out of unemployment, and 0.07 for workers in the cross-section. Consistent with the findings in Lentz Piyapromdee and Robin (2023), we find that sorting for workers coming out of unemployment is quite similar to sorting in the cross-section. Intuitively, sorting for workers coming out of unemployment is not random because, unlike in related models (e.g., Bagger and Lentz 2019), workers and firms both face an opportunity cost in forming a match in our model. The opportunity cost for workers is giving up on search efficiency. The opportunity cost for firms is having to fire an existing employee. The fact that sorting among workers coming out of unemployment is similar to sorting in the cross-section, together with the fact that we are targeting a measure of cross-sectional sorting, should reassure the reader that we are not underestimating sorting out of unemployment.

Even in light of the above robustness exercises, we acknowledge that identifying time-varying worker types, time-varying coworker types, and the extent to which workers' and coworkers' types are sorted is an extremely challenging empirical task. For this reason, our findings are unlikely to be the definitive answer about both sorting and the importance of learning from coworkers.

⁹We refer the reader to Lentz, Piyapromdee and Robin (2023) for the definition of normalized Mutual Information and for the details on the construction of worker and firm types in their data.

4 Counterfactuals and Welfare

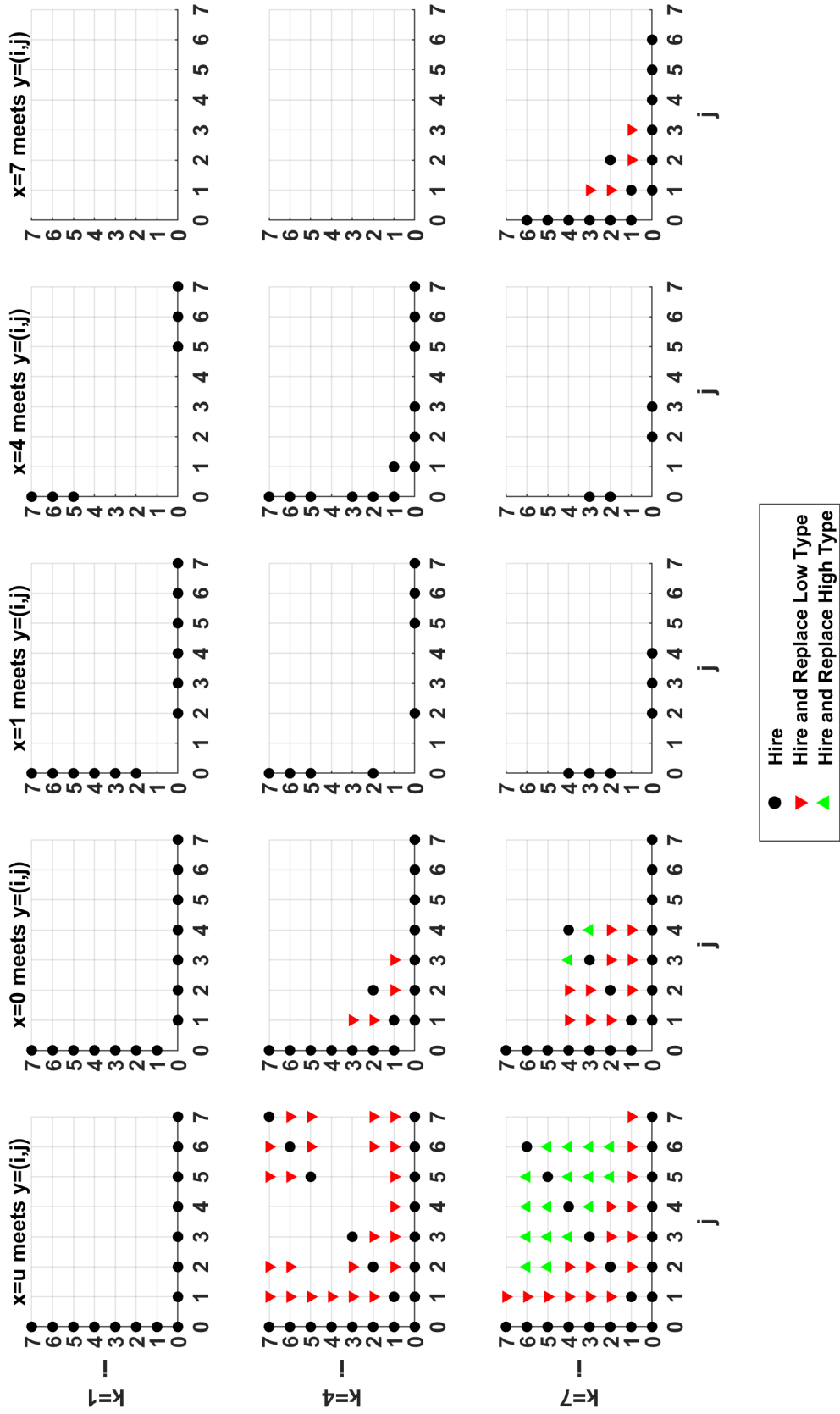
In this section, we use the calibrated model to carry out some counterfactuals and evaluate welfare. In Section 4.1, we establish some properties of the calibrated model in order to develop a benchmark for our counterfactual and welfare analysis. In Section 4.2, we consider a counterfactual in which workers learn by doing but not from their coworkers. In Section 4.3, we consider a counterfactual in which labor market segregation is caused by technological changes that increase the supermodularity of the production function. In Section 4.4, we compare, analytically and quantitatively, the properties of equilibrium with the properties of the solution of the social planner's problem.

4.1 Properties of equilibrium

In order to understand the properties of equilibrium, it is useful to start from the policy functions. Figure 3 shows whether the gains from trade between a worker at the k -th rung of the human capital ladder (the rows) in employment state x (the columns) and a firm in state y (inside each cell) are positive and, hence, whether the firm hires the worker. If a firm with two employees does hire the worker, Figure 3 also shows whether the worker is hired to replace the employee with the highest (green arrow up) or lowest (red arrow down) stock of human capital.

Consider the first column in Figure 3, which shows whether an unemployed worker ($x = u$) is hired by a firm in state y . When an unemployed worker meets a firm without employees (i.e. $y = (0, 0)$), the gains from trade are mainly determined by the productivity and the human capital growth of the worker when employed net of the productivity and the human capital growth of the worker when unemployed. These gains from trade are positive—as the worker gains human capital when employed and loses human capital when unemployed. When an unemployed worker meets a firm with a single employee (i.e. $y = (i, j)$ with i or $j = 0$), the gains from trade are mainly determined by the increase in the productivity and human capital growth of the worker when he gets paired with the firm's employee, and by the extra human capital growth of the firm's employee when he gets paired with the worker. These gains from trade are also positive. When an unemployed worker meets a firm with two employees (i.e. $y = (i, j)$ with $i, j > 0$), the firm has to fire one of them to hire the worker. The gains from trade depend on the increase in the productivity and human capital growth of the worker when he is paired with the firm's surviving employee, net of the decline in the productivity and human capital growth of the fired employee when he moves into unemployment. Moreover, the gains from trade depend on the net effect on the human capital growth of the firm's surviving employee

Figure 3: Equilibrium Policy Function



from changing partner. These gains from trade may be positive or negative.

In light of the above observations, we can make sense of the first column in Figure 3. An unemployed worker at the lowest rung of the human capital ladder (i.e. $k = 1$) is hired by firms without employees and by firms with a single employee. The worker, however, is not hired by firms with two employees, as the cost of sending one of the employees into unemployment always exceeds the benefit of lifting the worker out of unemployment. An unemployed worker at the middle rung of the ladder (i.e. $k = 4$) is hired by firms without employees and by firms with a single employee. The worker is also hired by some firms with fully-formed production units as a replacement for the employee with the lowest human capital. In most cases, the worker is hired to replace an employee that has less human capital than he does. In these cases, the gains from trade are positive because the benefit of lifting the worker out of unemployment exceeds the cost of sending the employee into unemployment. In a few cases, though, the worker is hired to replace an employee that has more human capital than he does. In these cases, which arise when the firm has two employees with high human capital, the gains of trade are positive because the worker has more to learn from the firm's surviving employee. An unemployed worker at the highest rung of the ladder (i.e. $k = 7$) is hired by firms without employees, by firms with a single employee, and by nearly all of the firms with two employees. Interestingly, the worker is sometimes hired as a replacement for the employee with the highest human capital. In these cases, which arise when the firm has two employees with relatively high human capital, the gains from trade are positive because the firm's surviving employee has more to learn from the newly hired worker than from his old partner.

Having understood the policy functions in the first column in Figure 3, we can make sense of the other columns. The gains from trade between a worker of type k and a firm in state y are smaller if the worker is employed ($x = 0, 1, 2 \dots$) rather than unemployed ($x = u$). Similarly, the gains from trade between a worker of type k and a firm in state y are smaller if the worker is employed with a coworker ($x = \ell$, with $\ell = 1, 2, \dots, N$) rather than employed by himself ($x = 0$). For these reasons, a worker of type k is hired by a smaller and smaller subset of firms as the worker's employment state x changes from unemployment (the first column in Figure 3) to employment without a coworker (the second column) and to employment with a coworker (the last three columns).

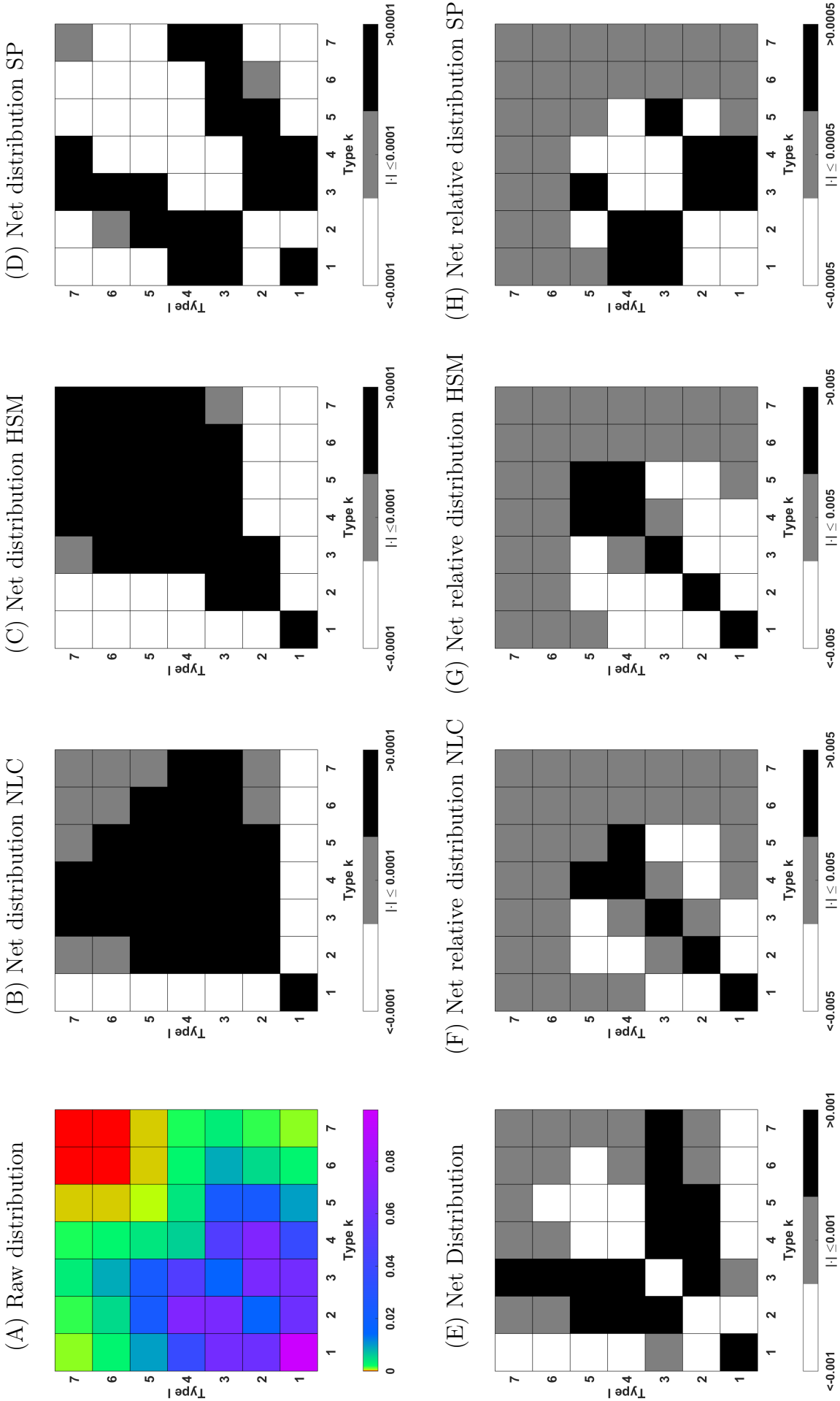
A worker of type k is hired in state $x = u$ but not in state $x = 0$ by firms that have marginal gains from trade, which are typically firms with a fully-formed production unit and at least one employee with high human capital. A worker of type k in state $x = \ell$ is never hired by a firm without employees, as this would involve breaking up the fully-formed production unit of the worker without creating a new one. A worker of

type k in state $x = \ell$ is almost never hired by a firm with two employees, as this would require sending one of the firm’s employees into unemployment without lifting anyone out of unemployment. Indeed, a worker of type k in state $x = \ell$ is almost exclusively hired by firms with a single employee. Sometimes the worker is hired by a firm with an employee that has more human capital than the worker’s current partner. This happens when there are both production and learning gains from reassigning the worker to a partner with more human capital (e.g., when the worker has less human capital than his current coworker) or when the production gains dominate the learning losses (e.g., when the distance between the human capital of the firm’s employee and the worker is not much smaller than the distance between the human capital of the worker and his current partner). Sometimes the worker is hired by a firm with an employee that has less human capital than the worker’s current coworker. This happens when the value of having the worker teach to a less knowledgeable coworker dominates any production losses (e.g., when both the worker and his current partner have more human capital than the firm’s employee).

Two key forces emerge from the analysis of the policy functions. First, production units in which both workers have the same stock of human capital are unstable—they are likely to break up when either one of the workers has the opportunity to move to a different production unit. This force—which is due to the convexity of the human capital accumulation function—pushes the equilibrium away from positive sorting. Second, production units in which one of the workers has low human capital are unstable—the worker with low human capital is likely to be replaced by an unemployed worker with more human capital. This force—which is generated by the fact that high human capital workers are more productive than low human capital ones—pushes the equilibrium towards positive sorting. A third force emerges mechanically from the properties of the human capital accumulation function. Since workers tend to catch up to more knowledgeable coworkers, the equilibrium allocation of workers tends to mechanically move towards positive sorting.

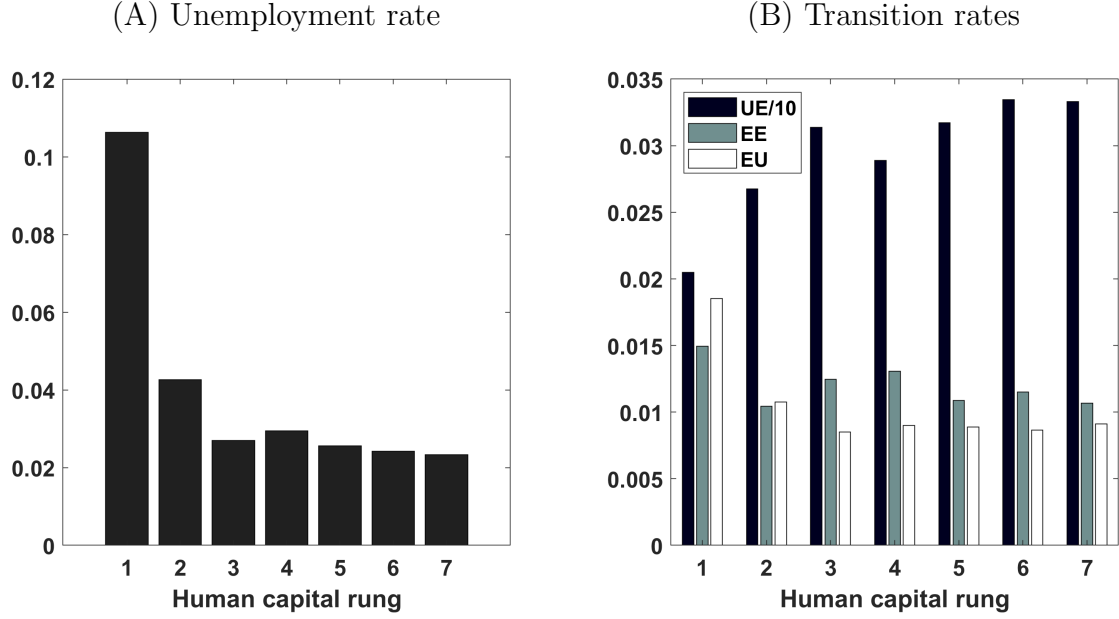
The forces above give rise to the equilibrium pattern of sorting. Panel A in Figure 4 plots the equilibrium distribution of workers and coworkers across production units. If the equilibrium featured PAM, all production units would be located along the off-diagonal. If the equilibrium featured NAM, all production units would be located along the main diagonal. The equilibrium features neither PAM or NAM, as the distribution of production units clearly has full support. Workers of type 1 are employed with all types of coworkers, but most often with coworkers of type 1. Workers of types 2 and 3 are employed with all types of coworkers, but most often with coworkers that have a similar but not identical level of human capital. Workers of type 4, 5, 6 and 7 are employed with

Figure 4: Net and relative distributions of workers and coworkers



Notes: Panel A is the baseline distribution of workers and coworkers. Panels B through E plot the difference between the fraction of teams with a worker of type k and a worker of type l net of the fraction of teams with a worker of type l if teams were formed at random in the NLC counterfactual and in the baseline. Panels F through H plot the net distribution after subtracting the benchmark model's net distribution.

Figure 5: Unemployment and transition rates



all types of coworkers, but most often with coworkers that have somewhat lower human capital.

The interpretation of the equilibrium pattern of sorting is complicated by the fact that there are different numbers of workers at different levels of human capital. To eliminate this complication, panel E in Figure 4 plots the equilibrium distribution of workers and coworkers net of the distribution that would obtain if workers and coworkers were matched at random. The panel makes it clear that the distribution of workers and coworkers is disproportionately concentrated around the off-diagonal, but not along the off-diagonal. Along the off-diagonal, the distribution has less density than under a random assignment of workers and coworkers, except for workers at the lowest rung of the human capital ladder. Around the off-diagonal, the distribution has more density than under a random assignment of workers and coworkers. In other words, workers are disproportionately matched with coworkers that have either slightly more or slightly less human capital than they do.

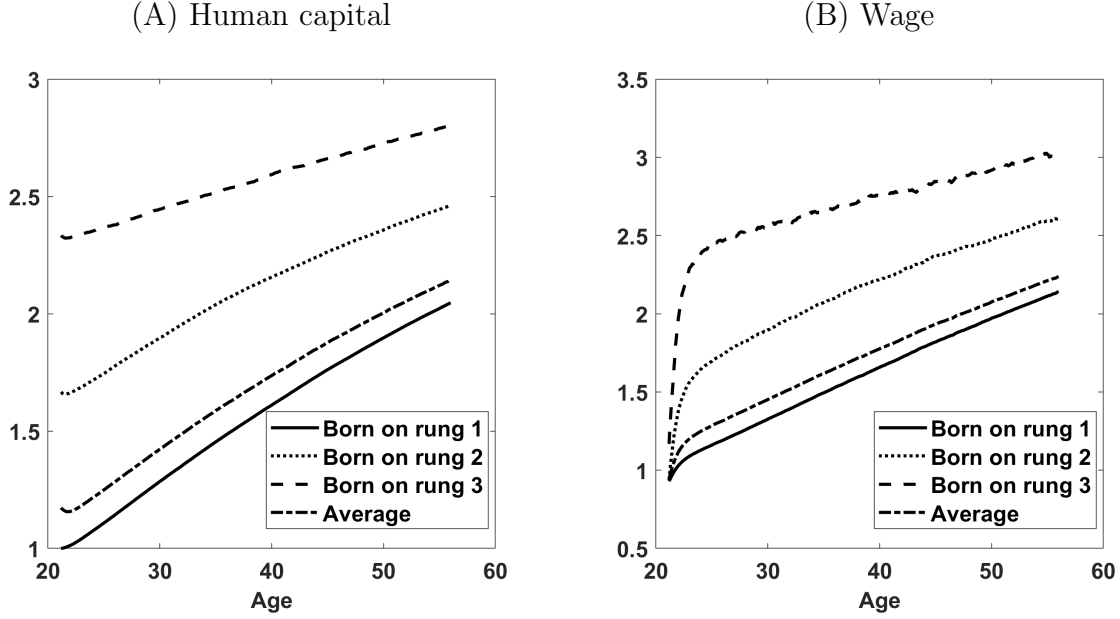
The policy functions also determine the UE, EU, EE and unemployment rates for different types of workers, which are plotted in Figure 5. The UE rate tends to be higher for workers with more human capital than for workers with less human capital. Conversely, the EU rate tends to be lower for workers with more human capital than for workers with less human capital. Workers with more human capital are more likely to be hired out of unemployment, as the benefit of lifting the worker out of unemployment is

more likely to exceed the cost of sending one of the firm’s employees into unemployment. For the same reason, workers with more human capital are less likely to be replaced by new hires and sent into unemployment. Since the UE rate is higher and the EU rate is lower for workers with more human capital, it follows that the unemployment rate is lower for workers with more human capital. Indeed, the unemployment rate is as high as 10% for workers at the bottom rung of the human capital ladder and as low as 2% for workers at the top of the human capital ladder. Hence, our model with heterogeneous workers and decreasing returns to scale provides a possible explanation for the observation that workers systematically differ with respect to their patterns of employment transitions (see, e.g., Ahn and Hamilton 2019, Hall and Kudlyak 2020, Gregory, Menzio and Wiczer 2021).

The policy functions and the human capital accumulation functions determine the growth of human capital and wages over the lifecycle (Figure 6). The growth of human capital over the lifecycle is due to learning by doing—which depends on the amount of time spent by workers in employment—and to learning from coworkers—which depends on the amount of time spent by workers in the company of different types of coworkers. The growth of wages over the lifecycle is due not only to the growth of human capital, but also to the fact that, over time, workers tend to move to firms where their marginal value is higher and they tend to capture an increasing share of their marginal value. For this reason, wages grow faster than human capital, especially at the initial stages of the lifecycle. Figure 6 also shows that initial differences in human capital and wages are very persistent over the lifecycle, despite the fact that low-human capital workers learn not only by doing but also from more knowledgeable peers. This finding is reassuring, since empirically wage differences across workers are very persistent.

Figure 7 plots the distribution of human capital among all workers who are active in the labor market, as well as the distribution of human capital among workers who have just entered the labor market. The distribution of human capital among active workers stochastically dominates the distribution of human capital among entering workers, reflecting both learning by doing and learning from coworkers. Quantitatively, the stock of human capital among active workers is 70% higher than the stock of human capital among entering workers. In this sense, learning on the job accounts for $0.7/1.7 = 41\%$ of the total stock of human capital in the economy, where 1 represents the stock of human capital among entering workers, 1.7 the stock of human capital among active workers, and 0.7 the stock of human capital accumulated on the job.

Figure 6: Human capital and wages over the lifecycle

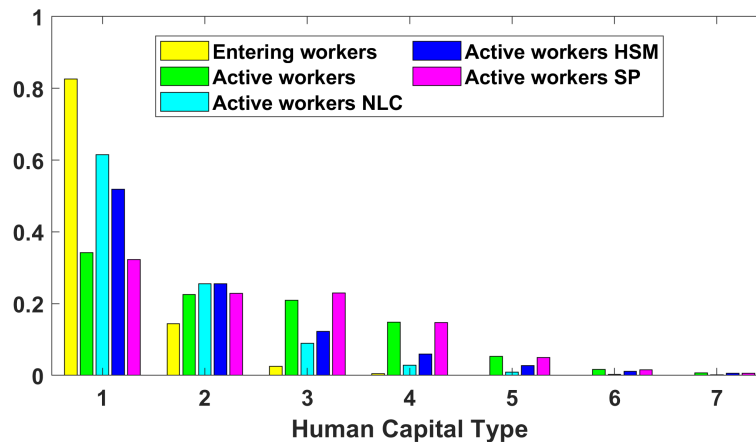


4.2 Learning from coworkers

We now want to use the calibrated model to assess the contribution of learning from coworkers to the aggregate stock of human capital, the aggregate flow of output, and the pattern of sorting between workers and coworkers. To this aim, we construct a counterfactual where the parameter α_1 —the parameter that controls the extent to which an individual learns from more knowledgeable coworkers—is set to 0, while all the other parameters are set as in the baseline calibration. We refer to this counterfactual as the No Learning from Coworkers (NLC) model.

Figure 7 plots the distribution of human capital among active workers in the equilibrium of the NLC model. The distribution of human capital among active workers is lower in the NLC than in the baseline model. The aggregate stock of human capital among active workers is 29% lower in the NLC than in the baseline model. In this sense, the learning-from-coworkers channel accounts for about one third of the aggregate stock of human capital. Furthermore, the aggregate stock of human capital among active workers is 21% higher than the aggregate stock of human capital among entering workers in the NLC model, while the aggregate stock of human capital among active workers is 70% higher than the aggregate stock of human capital among entering workers in the baseline model. In this sense, the learning-from-coworkers channel accounts for two thirds of the stock of human capital that is accumulated by workers on the job, while the learning-by-doing channel accounts for the remaining third.

Figure 7: Type Distribution



Panel B in Figure 4 plots the distribution of workers and coworkers in the NLC model net of the distribution that would obtain under random assignment. Panel E plots the distribution of workers and coworkers in the baseline model net of the distribution that would obtain under random assignment. The two figures are not directly comparable because the distribution of workers across rungs of the human capital ladder is different in the two models. To circumvent this problem, panel F plots the difference between the net distribution of workers and coworkers in the NLC model and in the baseline model. Panel F makes it clear that the distribution of workers and coworkers in the NLC model has more density along the off-diagonal and less density around the off-diagonal than in the baseline model. In this sense, the learning-from-coworkers channel makes the equilibrium pattern of sorting less positive.

The intuition behind these findings is simple. Since workers learn from more knowledgeable coworkers and are not slowed down by less knowledgeable ones, the value of production units in which workers have different stocks of human capital is higher than it would be if workers only learned by doing. Hence, the pattern of sorting of workers and coworkers is less positive than it would be if workers only learned by doing. Similarly, the human capital of individual workers grows faster than it would if workers only learned by doing. For this reason, the aggregate stock of human capital is higher than it would be if workers only learned by doing. Moreover, since the aggregate stock of human capital is higher, the aggregate flow of output is higher as well. Specifically, the aggregate flow of output is 27% higher than it would be if workers only learned by doing.

4.3 Higher supermodularity

Recent empirical work by Song et al. (2019) has documented a substantial rise in the extent of segregation in the US labor market—in the sense that high-wage individuals are more and more likely to work with other high-wage individuals, while low-wage individuals are more and more likely to work with other low-wage individuals. We want to use the calibrated model to show that the rise in labor market segregation can be explained by technological changes that increase the supermodularity of the production function—a hypothesis put forward by Kremer (1993). Moreover, we want to use the model to measure the impact of an increase in the supermodularity of the production function on human capital and output. To carry out these goals, we construct a counterfactual where ρ is lowered from its baseline value of 0.81 to -5 , which implies that the extra output produced by a firm in state $(1, 1)$ and a firm in state $(2, 2)$ relative to the output produced by two firms in state $(1, 2)$ increases from 0.5% to 15%. We refer to this counterfactual as the High Supermodularity (HSM) model.

In panel C of Figure 4, we plot the distribution of workers and coworkers in the HSM model net of the distribution that would obtain if workers and coworkers were assigned at random. In panel G, we plot the difference between the net distribution of workers and coworkers in the HSM and in the baseline model. Panel G shows that the distribution of workers and coworkers in the HSM model has more density along the off-diagonal and less density around the off-diagonal than in the baseline model. The increase in production supermodularity leads to an increase in segregation. High human capital workers become more likely to be employed with other high human capital workers, while low human capital workers become more likely to be employed with other low human capital workers. The finding is easy to understand. The increase in production supermodularity increases the value of production units where workers have similar stocks of human capital relative to production units where workers have different stocks of human capital and, hence, leads to more positive sorting.

In Figure 7, we plot the distribution of human capital among active workers in the HSM model. The distribution is lower in the HSM than in the baseline model. Quantitatively, the aggregate stock of human capital is 18% lower in the HSM than in the baseline model. Furthermore, the aggregate stock of human capital among active workers is 40% higher than among entering workers in the HSM model, and 70% higher in the baseline model. Hence, the increase in the supermodularity of the production function leads to an 18% loss in the aggregate stock of human capital, and to a 43% loss in the stock of human capital accumulated on the job. These findings are also easy to understand. An increase in production supermodularity leads to a rise in labor market segregation. Since low human

capital individuals become more likely to work with other low human capital individuals, they have fewer chances to learn from more knowledgeable coworkers and their human capital growth slows down. As a result, the stock of human capital declines.

The decline in aggregate output is approximately 23%. This number requires some interpretation. By construction of (9), the output produced by workers with the same human capital is independent of ρ . The output produced by workers with different human capital, however, does depend on ρ and, in particular, it falls when we lower ρ from 0.8 to -5 . For this reason, part of the decline in aggregate output is mechanically due to the decline in production function, and part of it is due to the change in the stock of human capital and in the pattern of sorting. To parse out the mechanical effect, we compute aggregate output at the ergodic distribution associated with the HSM model using the same production function as in the baseline model. We find that, when measured using the same production function as in the baseline model, aggregate output declines by 15%. The remaining 8 percentage points of output loss are due to the fact that by lowering ρ we lowered the production function (9) for any $h_k \neq h_\ell$.

4.4 Welfare

We now want to examine the welfare properties of equilibrium. In Appendix D, we formulate and solve the problem of a utilitarian social planner. We compare the stationary equilibrium with the steady state associated with the solution of the planner's problem, and we establish that the stationary equilibrium is always inefficient (Proposition 1). We then construct a system of employment subsidies such that the steady state associated with the solution to the planner's problem can be implemented as a stationary equilibrium (Proposition 2).

It is easy to understand why the equilibrium is inefficient. From the perspective of the social planner, the value of an additional unemployed worker includes the entirety of the gains from trade between the worker and a firm. Similarly, from the perspective of the social planner, the value of an additional production unit includes the entirety of the gains from trade between an employee and a poaching firm and between the firm and a potential new hire. In equilibrium, however, the value of unemployment to a worker includes only a fraction γ of the gains from trade between the worker and a firm. And the value of a production unit includes a fraction γ of the gains from trade between an employee and a poaching firm, and a fraction $1 - \gamma$ of the gains from trade between the firm and a potential new hire. Since γ and $1 - \gamma$ cannot be both equal to 1, the equilibrium values of unemployed workers and production units differ from the planner's marginal values and, in turn, the equilibrium hiring and firing decisions differ from the efficient hiring and

firing decisions. This type of inefficiency is common to models with two-sided search in which the measure of meetings between two types (say, unemployed workers of type k and production units in state (i, j)) is proportional to the product between the measures of the two types (see, e.g., Kiyotaki and Lagos 2007).

Having established that the equilibrium is inefficient, we quantify the nature and magnitude of the inefficiencies. In panel D of Figure 4, we plot the distribution of workers and coworkers in the solution to the planner’s problem net of the distribution that would obtain under random assignment. In panel H, we plot the difference between the net distribution of workers and coworkers in the solution to the planner’s problem and in the equilibrium. Panel H shows that the efficient distribution has more density around the off-diagonal and less density along the off-diagonal than the equilibrium distribution. In this sense, the equilibrium pattern of sorting is inefficiently positive. Intuitively, this is because the value of an unemployed worker of type k is lower in equilibrium than in the social plan, and the gap between the equilibrium and the social value of an unemployed worker is increasing in k . For this reason, firms in equilibrium replace low human capital employees with high human capital workers hired out of unemployment too often, and the equilibrium pattern of sorting becomes inefficiently positive.

In Figure 7, we plot the efficient distribution of human capital among active workers—i.e., the distribution of human capital among active workers in the steady state associated with the solution of the planner’s problem. In aggregate, the efficient stock of human capital is about 1% higher than in equilibrium and, as a result, the efficient flow of output is about 0.5% higher than in equilibrium. These findings are also easy to understand. Since the equilibrium pattern of sorting between workers and coworkers is inefficiently positive, low human capital workers are underexposed to more knowledgeable coworkers and, in turn, their human capital growth is inefficiently low, the aggregate stock of human capital is inefficiently low, and so is aggregate output. The equilibrium inefficiencies are, however, relatively small.

5 Conclusions

In this paper, we use theory and data to measure the extent to which the human capital growth of an individual depends on the quality of his coworkers. We first lay out a search-theoretic model of the labor market in which the productivity and the human capital accumulation of an individual both depend on the human capital of his coworkers. We then use an administrative matched employer-employee data set to uncover empirical evidence on the relation between the wage growth of an individual and the wage of his

past coworkers, as well as on the relation between the wages of the coworkers and the individual's past wage.

Empirically, we examine worker's wage changes based on their coworkers relative wages. We estimate that for workers who complete an EUE transition and earn less than their coworkers on the first job, a 10% increase in the coworkers' average wage forecasts a 1.23% higher wage on the second job. For workers who earn more than their coworkers on the first job, the effects of coworkers on wages in the second job are significantly weaker. These estimates suggest that learning on the job is convex in the human capital of the coworkers.

Given the sorting of workers and coworkers is not random, we document the relation between a worker's wage and the wages of future coworkers to measure the pattern of sorting. Among workers who complete an EUE transition, we find that regardless of whether an individual earns more or less than their coworkers on the first job, a 10% increase in their own wage forecasts roughly a 2% higher average coworkers' wage in the second job. These estimates reveal that sorting is positive.

Using the model, we translate our empirical evidence into structural parameters. We find that the production function is supermodular and that the human capital accumulation function is convex. Specifically, for an individual who is less knowledgeable than his coworkers, human capital growth depends positively on the human capital of the coworkers. For an individual who is more knowledgeable than his coworkers, human capital growth is independent from the human capital of the coworkers.

We find that learning from coworkers accounts for two thirds of the stock of human capital that is accumulated by workers while on the job, while the remaining third comes from learning by doing. We show that technological changes that increase the supermodularity of the production function cause labor market segregation and, by reducing the chances of low human capital workers to learn from more knowledgeable coworkers, eventually lead to lower aggregate human capital and output. We prove analytically that the equilibrium is always inefficient, and we show quantitatively that the equilibrium features too little mixing of high and low human capital workers.

References

- [1] Ahn, Hie Joo, and James Hamilton. 2019. “Heterogeneity and Unemployment Dynamics” *Journal of Business & Economic Statistics*, 38: 554-569
- [2] Anderson, Axel. 2015. “A Dynamic Generalization of Becker’s Assortative Matching Result.” *Journal of Economic Theory*, 159: 290-310.
- [3] Anderson, Axel, and Lones Smith. 2010. “Dynamic Matching and Evolving Reputations.” *Review of Economic Studies*, 77: 3-29.
- [4] Arrow, Kenneth. 1962. “The Economic Implications of Learning by Doing.” *Review of Economic Studies*, 29: 155-173.
- [5] Autor, David, and David Dorn. 2013. “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market.” *American Economic Review*, 103: 1552-1597.
- [6] Bagger, Jesper, Francois Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. 2014. “Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics.” *American Economic Review*, 104: 1551-1596.
- [7] Bagger, Jesper, and Rasmus Lentz. 2019. “An Empirical Model of Wage Dispersion with Sorting.” *Review of Economic Studies*, 86: 153-190.
- [8] Becker, Gary. 1962. “Investment in Human Capital: A Theoretical Analysis.” *Journal of Political Economy*, 70: 9-49.
- [9] Benhabib, Jess, Jesse Perla, and Christopher Tonetti. 2021. “Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier.” *Econometrica*, 89: 2261-2301.
- [10] Ben-Porath, Yoram. 1967. “The Production of Human Capital and the Life Cycle of Earnings.” *Journal of Political Economy*, 75: 352-365.
- [11] Bilal, Adrien, Niklas Engbom, Simon Mongey and Giovanni Violante. 2022. “Firm and Worker Dynamics in a Frictional Labor Market.” *Econometrica* 90, 1425-1462.
- [12] Booij, Adam, Edwin. Leuven, and Hessel Oosterbeek. 2017. “Ability Peer Effects in University: Evidence from a Randomized Experiment.” *Review of Economic Studies*, 84: 547-578.
- [13] Burdett, Kenneth, and Dale Mortensen. 1998. “Wage Differentials, Employer Size, and Unemployment.” *International Economic Review*, 39: 257-273.

- [14] Cahuc, Pierre, Fabien Postel-Vinay and Jean-Marc Robin. 2006. “Wage Bargaining with On-the-Job Search: Theory and Evidence.” *Econometrica*, 74: 323-364.
- [15] Carrell, Scott, Bruce Sacerdote, and James West. 2013. “From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation.” *Econometrica*, 81: 855-882.
- [16] Cornelissen, Thomas, Christian Dustmann, and Uta Schoenberg. 2018. “Peer Effects in the Workplace.” *American Economic Review*, 107: 425–56
- [17] Eeckhout, Jan and Philipp Kircher. 2011. “Identifying Sorting-In Theory.” *Review of Economic Studies*, 78: 872-906.
- [18] Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, John Robert Warren and Michael Westberry. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 CPS-IPUMS. Minneapolis, MN: IPUMS, 2022.
- [19] Freund, Lukas. 2022. “Superstar Teams: The Micro Origins and Macro Implications of Coworker Complementarities.” Cambridge Working Paper CWPE2276
- [20] Gregory Victoria. 2020. “Firms as Learning Environments: Implications for Earnings Dynamics and Job Search.” Manuscript, FRB Saint Louis.
- [21] Gregory, Victoria, Guido Menzio, and David Wiczer. 2021. “The Alpha Beta and Gamma of the Labor Market.” NBER Working Paper 28663.
- [22] Guvenen, Faith, Faith Karahan, Serdar Ozkan, and Jae Song. 2021. “What Do Data on Millions of U.S. Workers Reveal about Lifecycle Earnings Dynamics?” *Econometrica*, 89: 2303-2339.
- [23] Guvenen, Faith and Burhanettin Kuruscu. 2012. “Understanding the Evolution of the U.S. Wage Distribution: A Theoretical Analysis.” *Journal of the European Economic Association*, 10: 482-517.
- [24] Hagedorn, Marcus, Tzuo Hann Law, and Iourli Manovskii. 2017. “Identifying Equilibrium Models of Labor Market Sorting.” *Econometrica*, 85: 29-65.
- [25] Hall, Robert, and Marianna Kudlyak. 2020. “Job-Finding and Job-Losing: A Comprehensive Model of Heterogeneous Individual Labor-Market Dynamics Paper.” FRB San Francisco WP No. 19-05.
- [26] Hall, Robert, and Paul Milgrom. 2008. “The Limited Influence of Unemployment of the Wage Bargain.” *American Economic Review*, 98: 1653–74.
- [27] Herkenhoff, Kyle, Jeremy Lise, Guido Menzio and Gordon Phillips. 2018. “Production and Learning in Teams,” NBER Working Paper 25179.

- [28] Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER Working Paper 7867.
- [29] Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg. 2021. "Learning from Coworkers." *Econometrica*, 89: 647-676.
- [30] Jovanovic, Boyan. 2014. "Misallocation and Growth." *American Economic Review*, 104: 1149-1171.
- [31] Kiyotaki, Nobuhiro, and Ricardo Lagos. 2007. "A Model of Job and Worker Flows." *Journal of Political Economy*, 115: 770-815.
- [32] Kremer, Michael 1993. "The O-Ring Theory of Economic Development." *Quarterly Journal of Economics*, 108: 551-567.
- [33] Lise, Jeremy, Costas Meghir, and Jean-Marc Robin. 2016. "Matching, Sorting and Wages." *Review of Economic Dynamics*, 19: 63-87.
- [34] Lise, Jeremy, and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review*, 110: 2328-2376.
- [35] Lentz, Rasmus, Suphanit Piyapromdee, and Jean-Marc Robin. 2023. "On Worker and Firm Heterogeneity in Wages and Employment Mobility," *Econometrica*, Forthcoming.
- [36] Lopes de Melo, Rafael. 2018. "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence." *Journal of Political Economy*, 126: 313-346.
- [37] Lucas, Robert. 2009. "Ideas and Growth." *Economica*, 76: 1-19.
- [38] Lucas, Robert, and Benjamin Moll. 2014. "Knowledge Growth and the Allocation of Time." *Journal of Political Economy*, 122: 1-51.
- [39] Martellini, Paolo. 2021. "Local Labor Markets and Aggregate Productivity." Manuscript, UW Madison.
- [40] Menzio, Guido, Irina Telyukova, and Ludo Visschers. 2016. "Directed Search over the Life Cycle." *Review of Economic Dynamics*, 19: 38-62.
- [41] Mortensen, Dale and Christopher Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies*, 61: 397-415.
- [42] Nix, Emily. 2020. "Learning Spillovers in the Firm." Working Paper 2020:14, IFAU, Uppsala.
- [43] Perla, Jesse, and Christopher Tonetti. 2014. "Equilibrium Imitation and Growth." *Journal of Political Economy*, 122: 52-76.

- [44] Pissarides, Christopher 1985. “Short-run equilibrium dynamics of unemployment, vacancies, and real wages.” *American Economic Review*, 75: 676-690.
- [45] Postel Vinay, Fabien, and Jean-Marc Robin. 2002. “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity.” *Econometrica*, 70: 2295-2350.
- [46] Rosen, Sherwin. 1972. “Learning and Experience in the Labor Market.” *Journal of Human Resources*, 7: 326-342.
- [47] Shimer, Robert, and Lones Smith. 2000. “Assortative Matching and Search.” *Econometrica*, 68: 343-369.
- [48] Song, Jae, Faith Guvenen, David Price, Nicholas Bloom and Till von Wachter. 2019. “Firming Up Inequality.” *Quarterly Journal of Economics*, 134: 1-50.
- [49] U.S. Census Bureau. (2023). Longitudinal Household Employer Dynamics Data Base. Washington, DC: U.S. Census Bureau, Longitudinal-Employer Household Dynamics Program, accessed on 2018. (R2014)