

Production and Learning in Teams*

KYLE HERKENHOFF

University of Minnesota

JEREMY LISE

University of Minnesota and FRB Minneapolis

GUIDO MENZIO

NYU and NBER

GORDON PHILLIPS

Dartmouth College

October 2018

Abstract

The effect of coworkers on the learning and the productivity of an individual is measured combining theory and data. The theory is a frictional equilibrium model of the labor market in which production and the accumulation of human capital of an individual are allowed to depend on the human capital of coworkers. The data is a matched employer-employee dataset of US firms and workers. The measured production function is supermodular. The measured human capital function is non-linear: Workers catch-up to more knowledgeable coworkers, but are not dragged-down by less knowledgeable ones. The market equilibrium features a pattern of sorting of coworkers across teams that is inefficiently positive. This inefficiency results in low human capital individuals having too few chances to learn from more knowledgeable coworkers and, in turn, in a stock of human capital and a flow of output that are inefficiently low.

JEL Codes: E24, J24.

Keywords: Human Capital, Knowledge Diffusion, Search Frictions.

*We are grateful to Axel Anderson, Thibault Lamadon, Rasmus Lentz, Ellen McGrattan, Ricardo Lagos, Espen Moen, Tommaso Porzio, Jean-Marc Robin, Coen Teulings, Michela Tincani, and to participants at many conferences for comments and suggestions. 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.

1 Introduction

The theory of human capital has recognized the importance of learning on the job for understanding the lifecycle dynamics of wages ever since Becker (1962) and Ben-Porath (1967). The theory describes the process by which workers learn on the job by means of a human capital investment function. In some applications (e.g., Ben-Porath 1967), the investment function is assumed to include the worker’s time as the key input, thus leading to a trade-off between working and learning. In other applications (e.g., Bagger et al. 2014), the investment function is assumed to simply depend on whether the worker is employed or not, thus implying a mechanical process of accumulation. Much fewer applications, especially empirical ones, consider the quality of coworkers as an input in the human capital investment function. This is especially surprising considering that academics are very aware of the role played by colleagues in determining the growth of their knowledge.

The goal of this paper is to develop and quantify an equilibrium theory of the labor market in which the growth of an individual’s human capital is affected by his coworkers. The theory builds on the search-theoretic models of Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay and Robin (2006)—which have proved successful in explaining the patterns of workers’ transitions between unemployment, employment and across firms, as well as the dynamics of individual wages—and on the frictionless models of sorting with evolving types of Anderson (2015) and Anderson and Smith (2010)—which have explored the implications of the assumption that the human capital of an individual evolves depending on the quality of his coworkers. The theory is quantified using a matched employer-employee dataset of US workers and firms. The main finding of the paper is that coworkers matter for the human capital accumulation of an individual and they do so in a non-linear fashion. Specifically, an individual accumulates human capital at a rate that is increasing in the human capital of his coworkers, when his coworkers are more knowledgeable than he is. In contrast, an individual accumulates human capital at a rate that is nearly independent of the human capital of his coworkers, when his coworkers are less knowledgeable than he is. In other words, individuals catch-up to more knowledgeable coworkers but they are not dragged-down by less knowledgeable coworkers. In the aggregate, learning from coworkers accounts for 60% of all the learning on the job that takes place in the economy.

In the first part of the paper, we develop the theory. We consider a labor market populated by workers and firms. Workers are ex-ante heterogeneous with respect to their initial human capital. Firms are ex-ante homogeneous and produce by managing teams of

two workers. The firm’s production function may be supermodular or submodular in the human capital of the two team members. Workers and firms come into contact with each other as a result of a random search process. When a firm and a worker meet, the firm hires the worker if the gains from trade are positive—in the sense that the marginal value of the worker to the production unit of the firm exceeds the marginal value of the worker in his current production unit (if the worker is employed) or the value of unemployment (if the worker is unemployed). When a firm hires a worker, the wage allocates the gains from trade between the firm and the worker according to some given fractions. The wage remains constant until the worker’s outside option binds because he contacted some other firm. The human capital of a worker evolves over time according to a stochastic process which depends on whether the worker is employed or unemployed and, in the first case, on the human capital of the worker’s colleagues.

In the second part of the paper, we quantify the theory. Using a matched employer-employee dataset of US workers and firms, we look at workers who move from employment to unemployment and then back into employment. We find that, for individuals paid *less* than their coworkers in the first job, a 10% increase in the coworkers’ average wage forecasts a 1.5% higher wage in the second job (conditional on the individual’s wage in the first job). For individuals paid *more* than their coworkers in the first job, the coworkers’ average wage does not help forecast their wage in the second job. The logic behind these regressions is simple. The individual’s wage in the first job is a measure of his human capital at the start of the job, the coworkers’ wage in the first job is a measure of their human capital, and the individual’s wage in the second job is a measure of his human capital at the start of the second job. We focus on people going through an unemployment spell in between the two jobs, in order to purge the data from heterogeneity in their outside option when negotiating their wage in the second job. Even so, the correlation between the coworkers’ wage and the individual’s wage in the second job cannot be interpreted as the *effect* of the coworkers’ type on the human capital of the individual. In fact, sorting may imply some correlation between the coworkers’ human capital and the human capital of the individual at the start of the first job. To deal with this issue, we use the structure of the model. We calibrate the model so that it matches an empirical measure of sorting: the fraction of the wage variance due to differences in the average wage of different firms and the variance due to differences in the wage of different workers at the same firm. Additionally, we use information on the relationship between job-to-job transition and the wage gap between the worker and his coworkers.

The calibrated parameters reveal that workers catch-up to more knowledgeable coworkers, but are not dragged-down by less knowledgeable coworkers. That is, for workers with

less human capital than their coworkers, the accumulation of human capital is increasing in the coworkers' knowledge. For workers with more human capital than their coworkers, the accumulation of human capital is independent of the coworkers' knowledge. The properties of the calibrated human capital accumulation function tend to generate negative sorting of workers—that is, low human capital workers being employed alongside high human capital ones. However, this would result in a counterfactually large fraction of wage variance within firms. Therefore, the calibrated production function must be supermodular, which tends to generate positive sorting of workers and realigns the fraction of wage variance within firms in the model to the one observed in the data.

To assess the importance of learning from coworkers in the formation of human capital, in the production of output, and in the pattern of sorting, we remove the coworkers' type from the inputs of the human capital accumulation function. We find that learning from coworkers accounts for 24% of the stock of human capital held by workers and for more than 60% of the stock of human capital accumulated by workers on the job. We find that learning from coworkers accounts for about 20% of output. Finally, we find that learning from coworkers makes the sorting pattern less positive—in the sense that it causes high human capital workers to be employed less often alongside other high human capital workers. To assess the role played by the supermodularity of the production function, we consider a version of the model in which output is modular in the human capital of the workers. We find that the supermodularity of the production function has a negative effect on the stock of human capital and on output. This is so because the supermodularity of the production function makes the sorting pattern more positive and, for this reason, reduces the opportunities for low human capital individuals to learn from more knowledgeable coworkers.

In the last part of the paper, we examine the efficiency properties of equilibrium. We prove that the equilibrium is generally inefficient, in the sense that it does not coincide with the solution to the problem of a social planner. The source of inefficiency is easy to understand. The marginal value of a certain type of production unit to the planner includes the output of the unit as well as the entire value of the meetings generated by the agents attached to the unit (i.e., the firm and the employees). The value of a production unit in equilibrium includes only a fraction of the value of the meetings generated by the agents attached to the unit (i.e., a fraction equal to the firm's bargaining power for the meetings of the firm and a fraction equal to the worker's bargaining power for the meetings of the employees). We also prove that there exists a system of taxes and subsidies that makes the equilibrium efficient.

Given the calibrated value of the parameters, we find that the equilibrium pattern of

sorting is inefficiently positive. The pattern of sorting is inefficiently positive because, in equilibrium, a firm too often replaces its worst employee with a higher human capital worker. Since the pattern of sorting is inefficiently positive, low human capital individuals do not have enough opportunities to learn from more knowledgeable coworkers and, hence, the stock of human capital is inefficiently low and so is the level of output. The optimal system of taxes and transfers effectively removes the firm's incentive to replace the worst of its employees with a better worker.

The paper contributes to the literature on search-theoretic models of the labor market. Our paper builds on equilibrium models of search on the job such as Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay and Robin (2006). Bagger et al. (2014), Menzio et al. (2016) and Burdett et al. (2018) introduced human capital accumulation into equilibrium models of search on the job. Our model differs from these works in two dimensions. First, we assume that a worker's human capital evolves over time based not only on experience but also on the quality of his coworkers. The assumption implies that the joint value of a firm is not simply the sum of the joint value of the firm and each employee separately. Having lost the additive separability of the joint value function, we are free to relax other simplifying assumptions. Specifically, we are free to assume a production function in which a worker's human capital may interact with the coworker's human capital rather than a production function in which a worker's human capital only interacts with the firm's type. Thus, our model speaks to the sorting pattern of workers and coworkers in production teams rather than to the sorting pattern of workers and firms. We believe that it is more natural to think that the assignment of an individual worker is driven by the quality of the coworkers in different firms rather than by the quality of the firm itself. Lise and Postel-Vinay (2016) consider a version of Postel-Vinay and Robin (2002) where the skills of an individual evolve over time depending on the characteristics of the job he holds. Their focus is on understanding the effect of job characteristics on the skills of an individual. Our focus is on understanding the effect of coworkers' characteristics on the human capital of an individual.

The paper contributes to a recent literature on the diffusion of knowledge. Lucas (2008), Perla and Tonetti (2014), Lucas and Moll (2014), Jovanovic (2014) and Benhabib, Perla and Tonetti (2017) develop a variety of models in which—as a result of bilateral meetings—individuals can catch-up to the human capital of others. The literature is theoretical and focused on the effect of knowledge diffusion on the growth rate of the economy. Our paper contributes to this literature by trying to measure the process of knowledge diffusion among coworkers using data on the relationship between the wage of an individual and the wage of his past coworkers. Jarosch, Oberfield and Rossi-Hansberg

(2018) have a similar goal but implement it using a different theory (a frictionless model) and different data (a German dataset).

The paper contributes to the literature on sorting. Within this literature, our closest predecessors are Anderson and Smith (2010) and Anderson (2015) who study frictionless assignment problems with evolving types. They consider the problem of assigning different workers to 2-person production teams, given that output depends on the workers' types, and that the evolution of an individual worker's type depends on the type of his coworker. They show that the supermodularity of the production function is not sufficient to guarantee positive sorting, as the function controlling the evolution of a worker's type may push towards negative sorting. For instance, this is the case when low-type workers tend to catch-up to high-type coworkers. Our paper is a quantitative version of Anderson and Smith (2010) and Anderson (2015). We find that the production function is supermodular and that learning is such that low-type workers catch-up to high-type coworkers and high-type workers are not dragged-down by low-type coworkers. Since ours is a quantitative model, we want to make sense of the workers' transitions between unemployment, employment and across employers. For this reason, we assume that the labor market is frictional rather than frictionless. Ours is not the first paper on sorting in frictional labor markets (see, e.g., Shimer and Smith 2000, Gautier and Teulings 2004, Eeckhout and Kircher 2011, Lise, Meghir and Robin 2016, Gulyas 2016, Bagger and Lentz 2018, Lopes de Melo 2018, Hagedorn, Law and Manovskii 2018). However, ours is the first paper in which sorting is between workers within teams rather than between workers and firms. As mentioned above, we believe that this type of sorting is more natural. Moreover, inferring the complementarity between a worker and his coworkers is easier than inferring the complementarity between a worker and a firm, as one can observe the wage of coworkers more easily than the profit of a firm.

There is also a large literature concerned with measuring the effect of peers on academic achievement. As in our paper, the key measurement problem is the endogenous sorting of individuals into groups. This literature deals with the problem either by relying on natural variation that is presumably uncorrelated with the individual's choice of peers (e.g., Hoxby 2000) or by carrying out randomized experiments (e.g., Carrell, Sacerdote and West 2013 or Booij, Leuven and Oosterbeek 2017). Interestingly, their findings about the effect of peers on academic achievement are qualitatively similar to our findings on the effect of coworkers on human capital accumulation. As we do, they find that the quality of peers improves the academic performance of an individual and that the effect is non-linear. There is a much smaller literature that tries to measure the effect of coworkers on an individual worker (see, e.g., Mas and Moretti 2009 and Cornelissen, Dustmann

and Schoenberg 2017). The literature is smaller because of several additional difficulties, namely that running experiments is more costly and that wages are not direct measures of performance. We deal with these difficulties using theory. That is, we use the well-established theory of search to model wage determination and the process of sorting of workers into teams.

2 Theory

In this section, we build a search-theoretic model of the labor market in which the productivity and the human capital accumulation of an individual depends on the quality of his coworkers. In Section 2.1, we describe the environment and compare it with Postel-Vinay and Robin (2002), Cahuc, Postel-Vinay and Robin (2006) and with Anderson and Smith (2010). In Section 2.2, we define an equilibrium. In Section 2.3, we briefly illustrate some of the features of equilibrium to allow the reader to get acquainted with the mechanics of the model.

2.1 Environment

Time is discrete and continues forever. In any period, the economy is populated by a continuum of workers with measure 1 and by a continuum of firms with measure 1. A worker maximizes the present value of income, discounted at the factor $\beta \in (0, 1)$. At any point in time, a worker has some level of human capital $h_k \in H$, with $H = \{h_1, h_2, \dots, h_N\}$ where $0 < h_1 < h_2 < \dots < h_N$. We refer to $k \in K \equiv \{1, 2, \dots, N\}$ as the worker's human capital type. A worker is in some employment state $x \in X \equiv \{u, \bar{K}\}$ where $\bar{K} \equiv \{0, 1, 2, \dots, N\}$. If the worker is in state $x = u$, it means that he is unemployed. If the worker is in state $x = 0$, it means that the worker is employed by a firm and has no coworker. If the worker is in state $x = 1, 2, \dots, N$, it means that the worker is employed by a firm and has a coworker of type x . A worker's human capital h_k is valuable for two reasons. First, human capital affects the amount of output that the worker can produce in various employment states (i.e., unemployed, employed on his own, or employed in a team of workers). Second, human capital affects the growth rate of the human capital of the worker's colleagues (i.e., how much he can teach his colleagues).

A firm maximizes the present value of profits, discounted at the factor β . At any point in time, a firm is in some state $y = (\bar{k}, \bar{\ell})$ with $\bar{k}, \bar{\ell} \in \bar{K}$.¹ If in state $y = (0, 0)$, the firm has no employees. If in state $y = (k, 0)$ for $k > 0$, the firm has only one employee of type k . If

¹We refer to \bar{k} and $\bar{\ell} \in \bar{K}$ as the employees' extended types, with the interpretation that an employee of extended type 0 is a vacant spot on the firm's team.

in state $y = (k, \ell)$ for $k > 0$ and $\ell > 0$, it means that the firm has a team of two employees of type k and ℓ . The firm operates a production function $f : Y \rightarrow R_+$ with $Y \equiv \bar{K} \times \bar{K}$. The production function is symmetric, in the sense that $f(\bar{k}, \bar{\ell}) = f(\bar{\ell}, \bar{k})$. The production function is such that, if the firm does not employ any workers, it produces no output, i.e. $f(0, 0) = 0$. If the firm only employs a worker of type k , it produces $f(k, 0)$ units of output. If the firm employs two workers, one of type k and one of type ℓ , it produces $f(k, \ell)$ units of output. For the sake of simplicity, we assume that the firm cannot employ more than two workers.

Every period is divided into five stages. At the *learning stage*, the human capital of a worker evolves according to a probability distribution function that depends on the worker's current human capital $k \in K$ and on his employment status $x \in X$. Specifically, if the worker's current human capital type is k and his employment status is x , his human capital type becomes k_+ with probability $g(k_+|k, x) \in [0, 1]$, where $\sum_{k_+} g(k_+|k, x) = 1$.

At the *entry-and-exit stage*, a worker exits the labor market with some probability $\sigma \in (0, 1)$. In aggregate, a measure σ of workers exits the labor market and is replaced by an equal measure σ of workers who enter the market. A worker who enters the market has a human capital type $k \in K$ drawn from a distribution π_k , where $\pi_k \in [0, 1]$ and $\sum_k \pi_k = 1$. A worker enters the market unemployed.

At the *search-and-matching stage*, an employed worker becomes unemployed for exogenous reasons with probability $\delta \in (0, 1)$. Also, workers and firms come into contact with each other. In particular, an unemployed worker contacts a randomly selected firm with probability $\lambda_u \in (0, 1)$. An employed worker contacts a randomly selected firm with probability $\lambda_e \in (0, 1)$. Since the meeting process is random, a worker may meet a firm employing no workers, a firm employing one worker, or a firm employing a team of two workers. Similarly, a firm may meet a worker who is unemployed, a worker who is employed by himself, or a worker who is part of a team. Upon meeting, the firm and the worker match if and only if the gains from trade are positive. If the firm and the worker match, the worker becomes an employee of the firm. A fraction γ of the gains from trade is then captured by the worker and a fraction $1 - \gamma$ is captured by the firm, with $\gamma \in [0, 1]$.

The gains from trade are defined as the difference between the joint value of the match and the outside options of the firm and the worker. The joint value of the match is given by the sum of the values to the worker, the firm and its current employees when the worker joins their production unit. The outside option of the firm is given by the sum of the value to the firm and its current employees from not having the worker in their production unit. If the worker is unemployed, his outside option is the value of unemployment. If the worker is already employed, his outside option is given by the difference between the joint

value to the worker, his current employer and his current coworkers from being all in the same production unit and the joint value to the employer and the coworkers from being without the worker. Note that our definition of gains from trade includes the values to all the parties attached to the production units affected by the trade: not only the firm and the worker, but also the firm’s current employees and the worker’s current employer and coworkers.

We assume that the wage of a worker remains constant unless it is such that the worker’s value from staying with the firm is lower than the worker’s outside option—i.e., either the value of unemployment or, if the worker has contacted a poaching firm, the net value of the worker to the poaching firm and its employees—or such that the worker’s value from staying with the firm exceeds the net value of the worker to the firm and its other employees. In the first case, the wage is raised to the point where the worker’s value equals his outside option. In the second case, the wage is lowered to the point where the worker’s value equals his net value to the firm and its other employees. The first constraint on the wage guarantees that the worker has a private incentive to remain with the firm whenever he is more valuable inside than outside the firm. The second constraint on the wage guarantees that the worker has a private incentive to report to the firm being contacted by a poaching firm whenever he is more valuable to the poaching firm than to his current one. Taken together, the constraints on the wage guarantee that the worker’s behavior is incentive compatible.

At the *dismissal stage*, a firm and a worker who are currently matched may voluntarily separate. The firm and the worker separate if and only if the gains from trade are negative. That is, the firm and the worker separate if and only if the joint value to the worker, the firm and the other employee from being in the same production unit is smaller than the sum between the value to the worker from being unemployed and the value to the firm and the other employee from producing without the worker.

At the *production stage*, matched firms and workers generate output according to the production function $f(\bar{k}, \bar{\ell})$, workers receive part of the output as wages, and firms receive the remaining output as profits. Unemployed workers generate output according to the home-production function $b(k)$.

Our model is a version of Postel-Vinay and Robin (2002, henceforth PR) and Cahuc, Postel-Vinay and Robin (2006, henceforth CPR) in which the productivity and the human capital accumulation of an individual depend on the quality of his coworkers. As in PR and CPR, workers contact randomly selected firms both while unemployed and while employed. As in PR and CPR, workers move to a contacted firm if and only if the gains from trade are positive, where the gains from trade are defined by taking into account

the values of all of the parties affected by the trade. In order to capture the idea that workers learn from their coworkers, our model must depart from PR and CPR by having a well-defined notion of a firm. We obtain a notion of the firm by specifying a production function with decreasing returns to labor. We introduce decreasing returns to scale by assuming that the firm can hire at most two workers. We make this extreme assumption to keep the number of arguments in the state of the firm manageable (i.e. the human capital of the two workers). Another difference between our model and PR and CPR is that, in our model, all firms are ex-ante identical. Yet, the randomness of the search process together with the decreasing returns to scale in labor are enough to create ex-post heterogeneity in the value of a particular worker to different firms and, hence, to create a proper job ladder.

Our model is a version of Anderson (2015) and Anderson and Smith (2010, henceforth AS) with search frictions. As in AS, workers produce in teams of two and the output depends in a general fashion on their human capital. If the production function is super-modular in the workers' human capital, production pushes towards positive assortative matching (PAM) of coworkers. If the production function is submodular, production pushes towards negative assortative matching (NAM). As in AS, the evolution of the worker's human capital depends on the human capital of their coworkers. If the learning process is such that the coworker with the lower human capital catches up to the one with more human capital, learning pushes for NAM. If the learning process is such that the coworker with the higher human capital is dragged down by the one with less human capital, learning pushes for PAM. The main difference between our model and AS is that, in our model, workers come to form teams through a frictional search process rather than by frictionless assignment.

2.2 Definition of Equilibrium

In order to define an equilibrium, we need to introduce some notation to describe the value functions and the distribution of workers and firms across employment states. We denote as U_k the value of a worker of type k who is unemployed. We denote as $V_{0,0}$ the value of a firm who employs no workers at the production stage. We denote as $V_{k,0}$ the sum of the values to a firm and to a worker of type k from being in the same production unit (without anybody else). We denote as $V_{k,\ell}$ the sum of the values to a firm, a worker of type k and a worker of type ℓ from being in the same production unit. The value functions U and V are measured at the beginning of the production stage. We also find it useful to denote as \hat{V} the joint value of production units at the beginning of the dismissal stage.

Next, we introduce the notation for the stationary distributions of workers and firms. Let u_k denote the measure of unemployed workers of type k , let $e_{k,0}$ denote the measure of workers of type k who are employed without a coworker, and $e_{k,\ell}$ denote the measure of workers of type k who are employed with a coworker of type ℓ . The distribution of workers implies a distribution of firms. Let n_0 , n_1 and n_2 denote the measure of firms with 0, 1 or 2 employees. Then, $n_1 = \sum_k e_{k,0}$, $n_2 = \sum_{k,\ell} e_{k,\ell}/2$ and $n_0 = 1 - n_1 - n_2$. The measures u , e and n are computed at the beginning of the search-and-matching stage.

We are now in the position to formulate the Bellman Equations satisfied by the value functions U and V . The value $V_{0,0}$ of a firm without employees is such that

$$V_{0,0} = 0 + \beta \left\{ \left[\sum_{i,x} q(i,x)(1-\gamma) \max\{\hat{V}_{i,0} - \hat{V}_{0,0} - v(i,x), 0\} \right] + \hat{V}_{0,0} \right\}. \quad (1)$$

In the current period, the profit of the firm is $f(0,0) = 0$. In the next period, the firm contacts a worker of type i in employment state x with probability $q(i,x)$. The probability $q(i,u)$ is $\lambda_u u_i$, the probability $q(i,0)$ is $\lambda_e e_{i,0}$, and the probability $q(i,j)$ is $\lambda_e e_{i,j}$. Conditional on meeting a worker of type i in state x , the firm continuation value is given by $\hat{V}_{0,0}$ plus a fraction $1 - \gamma$ of the gains from trade $\hat{V}_{i,0} - \hat{V}_{0,0} - v(i,x)$, where $\hat{V}_{i,0}$ is the joint value of the match to the firm and to the worker, $\hat{V}_{0,0}$ is the outside option of the firm, and $v(i,x)$ is the outside option of the worker. The outside option of the worker depends on his employment status x . If $x = u$, the outside option of the worker is U_i . If $x = 0$, the outside option of the worker is $\hat{V}_{i,0} - \hat{V}_{0,0}$. If $x = j$, the outside option of the worker is $\hat{V}_{i,j} - \hat{V}_{j,0}$.

The joint value $V_{k,0}$ of a production unit consisting of a firm and a worker of type k is such that

$$\begin{aligned} V_{k,0} = f(k,0) + \beta \mathbb{E} & \left\{ \sigma [\hat{V}_{0,0} - \hat{V}_{k_+,0}] + \delta [\hat{V}_{0,0} + U_{k_+} - \hat{V}_{k_+,0}] \right. \\ & + \left[\sum_{i,x} q(i,x) (1-\gamma) \max\{\hat{V}_{k_+,i} - \hat{V}_{k_+,0} - v(i,x), 0\} \right] \\ & \left. + \left[\sum_y \lambda_e p(y) \gamma \max\{v(k_+,y) - (\hat{V}_{k_+,0} - \hat{V}_{0,0}), 0\} \right] + \hat{V}_{k_+,0} \right\} \end{aligned} \quad (2)$$

In the current period, the sum of the income to the worker and the profit to the firm is $f(k,0)$. In the next period, the worker's human capital type is k_+ , where the probability distribution over k_+ is given by the probability distribution function g . The worker exits the labor market with probability σ . In this case, the continuation value to the firm is $\hat{V}_{0,0}$ and the continuation value to the worker is zero. The worker moves into unemployment with probability δ . In this case, the continuation value to the firm is $\hat{V}_{0,0}$ and the continuation value to the worker is U_{k_+} .

The firm contacts a worker of type i in employment state x with probability $q(i,x)$. In

this case, the joint continuation payoff to the firm and to the existing worker is $\hat{V}_{k_+,0}$ plus a fraction $1 - \gamma$ of the gains from trade between the firm and the contacted worker, which are given by $\hat{V}_{k_+,i} - \hat{V}_{k_+,0} - v(i, x)$. The worker contacts a poaching firm in employment state $y \in Y$ with probability $\lambda_e p(y)$, where $p(0, 0) = n_0$, $p(i, 0) = e_{i,0}$, and $p(i, j) = e_{i,j}/2$. Conditional on meeting a poaching firm in state y , the continuation payoff to the worker and its current employer is $\hat{V}_{k_+,0}$ plus a fraction γ of the gains from trade between the worker and the poaching firm $v(k_+, y) - (\hat{V}_{k_+,0} - V_{0,0})$, where $v(k_+, y)$ denotes the joint value of the match between the worker and the poaching firm net of the outside option of the poaching firm, and $\hat{V}_{k_+,0} - \hat{V}_{0,0}$ is the outside option of the worker. If $y = (0, 0)$, $v(k_+, y)$ is $\hat{V}_{k_+,0} - \hat{V}_{0,0}$. If $y = (i, 0)$, $v(k_+, y)$ is $\hat{V}_{k_+,i} - \hat{V}_{i,0}$. If $y = (i, j)$, $v(k_+, y)$ is the maximum between $\hat{V}_{k_+,i} + U_j - \hat{V}_{i,j}$ and $\hat{V}_{k_+,j} + U_i - \hat{V}_{i,j}$, as the poaching firm needs to break-up with one of its current employees to make room for the new hire.²

The joint value $V_{k,\ell}$ of a production unit consisting of a firm, a worker of type k and a worker of type ℓ is such that

$$\begin{aligned}
V_{k,\ell} = & f(k, \ell) + \beta \mathbb{E} \{ \\
& + [\sigma(\hat{V}_{k_+,0} + \hat{V}_{\ell_+,0}) - 2\sigma\hat{V}_{k_+,\ell_+}] + [\delta(\hat{V}_{k_+,0} + \hat{V}_{\ell_+,0} + U_{k_+} + U_{\ell_+}) - 2\delta\hat{V}_{k_+,\ell_+}] \\
& + \left[\sum_{i,z} q(i, z) (1 - \gamma) \max \left\{ \max\{\hat{V}_{k_+,i} + U_{\ell_+}, \hat{V}_{\ell_+,i} + U_{k_+}\} - \hat{V}_{k_+,\ell_+} - v(i, z), 0 \right\} \right] \\
& + \left[\sum_y \lambda_e p(y) \gamma \max \left\{ v(k_+, y) - (\hat{V}_{k_+,\ell_+} - \hat{V}_{\ell_+,0}), 0 \right\} \right] \\
& + \left[\sum_y \lambda_e p(y) \gamma \max \left\{ v(\ell_+, y) - (\hat{V}_{k_+,\ell_+} - \hat{V}_{k_+,0}), 0 \right\} \right] + \hat{V}_{k_+,\ell_+} \}.
\end{aligned} \tag{3}$$

For the sake of brevity, we will only discuss the differences between expressions (3) and (2). First, both the worker of type k and the worker of type ℓ face a probability of exiting the labor market or moving into unemployment. Second, both the worker of type k and the worker of type ℓ meet poaching firms in state y with probability $\lambda_e p(y)$ and capture a fraction γ of the gains from trade in those meetings. Third, when the firm meets a worker of type i in employment state z , it captures a fraction $1 - \gamma$ of the gains from trade, in which the joint value to the worker, the firm and its employees from being in the same production unit is the maximum between $\hat{V}_{k_+,i} + U_{\ell_+}$ and $\hat{V}_{\ell_+,i} + U_{k_+}$, as the firm chooses which employee to dismiss in order to make room for the new hire.

At the dismissal stage, $\hat{V}_{0,0} = V_{0,0}$ as the firm has no employees to dismiss. The joint

²As the reader can infer from the formulation of (2), we assume that the events at the search-and-matching stage are mutually exclusive at the level of a production unit. That is, either the worker moves into unemployment, or the worker meets a poaching firm, or the firm meets another worker. The assumption would be without loss in generality in a continuous-time version of the model.

value $\hat{V}_{k,0}$ of a production unit consisting of a firm and a worker of type k is such that

$$\hat{V}_{k,0} = \max \{V_{k,0}, V_{0,0} + U_k\}. \quad (4)$$

Similarly, the joint value $\hat{V}_{k,\ell}$ of a production unit consisting of a firm, a worker of type k and a worker of type ℓ is such that

$$\hat{V}_{k,\ell} = \max \left\{ V_{k,\ell}, \hat{V}_{k,0} + U_\ell, \hat{V}_{\ell,0} + U_k \right\}. \quad (5)$$

The above expressions are easy to understand. The joint value of the production unit at the dismissal stage is the maximum between the joint value to the members of the unit from staying together and the joint from breaking up.

Finally, the value of unemployment U_k to a worker of type k is given by

$$U_k = b(k) + \beta \mathbb{E}_{k_+} \left\{ U_{k_+} + \sigma [0 - U_{k_+}] + \left[\sum_y \lambda_u p(y) \gamma \max \{v(k_+, y) - U_{k_+}, 0\} \right] \right\}. \quad (6)$$

In the current period, the income of the worker is $b(k)$. In the next period, the worker exits the labor market with probability σ . In this case, the worker's continuation value is zero. The worker contacts a firm in state y with probability $\lambda_u p(y)$. In this case, the worker's continuation value is U_{k_+} plus a fraction γ of the gains from trade $v(k_+, y) - U_{k_+}$.

The value functions above pin down the probability with which a worker transits across human capital and employment states. These transition probabilities are cumbersome to write, but follow immediately from the value functions. Therefore, we relegate them in Appendix A and here we only describe the transitions for a worker of type k who is in employment status $x = \ell$. At the search-and-matching stage, the worker is forced to move into unemployment for exogenous reasons with probability δ . With probability $\lambda_e p(y)$, the worker contacts a poaching firm in state y and, depending on whether the gains from trade are positive or negative, he may move to the poaching firm or remain with his current employer. With probability $\lambda_e p(y)$, the worker's teammate contacts a poaching firm and, depending on the sign of gains from trade, the worker may become employed by himself or keep his teammate. With probability $q(i, x)$, the firm contacts a worker of type i in employment status x and, depending on the sign of the gains from trade, the worker of type k may move into unemployment, be paired with a new coworker of type i , or remain with the old coworker of type ℓ . At the dismissal stage, the worker may move into unemployment or remain employed. At the learning stage, the worker's type evolves from k to k_+ according to the probability distribution function $g(k_+|k, x)$. At the entry-and-exit stage, the worker exits the labor market for exogenous reasons with probability σ . Otherwise, the worker reaches the next search-and-matching stage.

We are now in the position to define an equilibrium for our model.

Definition: *A Stationary Equilibrium is a list of value functions $\{U, V\}$ and a distribution of workers across employment states $\{u, e\}$ such that: (i) The value functions satisfy conditions (1)-(6) given $\{u, e\}$; (ii) The distribution $\{u, e\}$ is stationary given the transitions probabilities implied by the value functions.*

Note that, just as in PR and CPR, the equilibrium of our model is defined without making any reference to wages. There is a simple intuition behind this property of equilibrium. First, the hiring decisions are determined by the gains from trade, which are defined as the joint value to the worker, the poaching firm and its employees from hiring the worker minus the joint value to the worker, his current employer and his coworkers from retaining the worker. Second, the joint values are independent of wages because wages transfer utility at the rate of 1 for 1 from the firm to the worker. Thus, all gains from trade and all hiring decisions are independent of wages, and so is the stationary distribution of workers and firms. Yet, we do care about wages, as they play a key role in our calibration of the model. We refer the reader to Appendix B for the equilibrium conditions for wages.

2.3 Illustration of Equilibrium

In order to familiarize our reader with the mechanics of the model, we believe it is helpful to illustrate a sample path for the history of a firm and a sample path for the history of a worker. Figure 1(a) presents a sample path for the history of a firm. At the beginning of the path, the firm is idle (i.e., state $y = (0, 0)$). In period t_1 , the firm meets an unemployed worker of type 1. Since the net value of the worker to the firm is greater than the value of unemployment, the firm hires the worker and moves to the state $y = (1, 0)$. In period t_2 , the firm meets an unemployed worker of type 2. As the gains from trade are positive, the firm hires the worker and moves to the state $y = (1, 2)$. In period t_3 , the firm meets an unemployed worker of type 4. The firm finds it optimal to hire the worker as a replacement for its employee of type 1 and moves to state $y = (4, 2)$. In period t_4 , the firm meets another unemployed worker of type 4. The firm finds it optimal to hire the worker as a replacement for its employee of type 2 and moves to state $y = (4, 4)$. In period t_5 , one of the employees of the firm contacts a poacher in state $(7, 0)$. The employee finds it optimal to leave the firm, which moves to state $y = (0, 4)$. In period t_6 , the remaining employee of the firm moves into unemployment for exogenous reasons. The firm is back to state $y = (0, 0)$.

The sample path in Figure 1(a) illustrates several important features of equilibrium. First, it takes time for a firm to build a team of workers, as it takes time to find each team

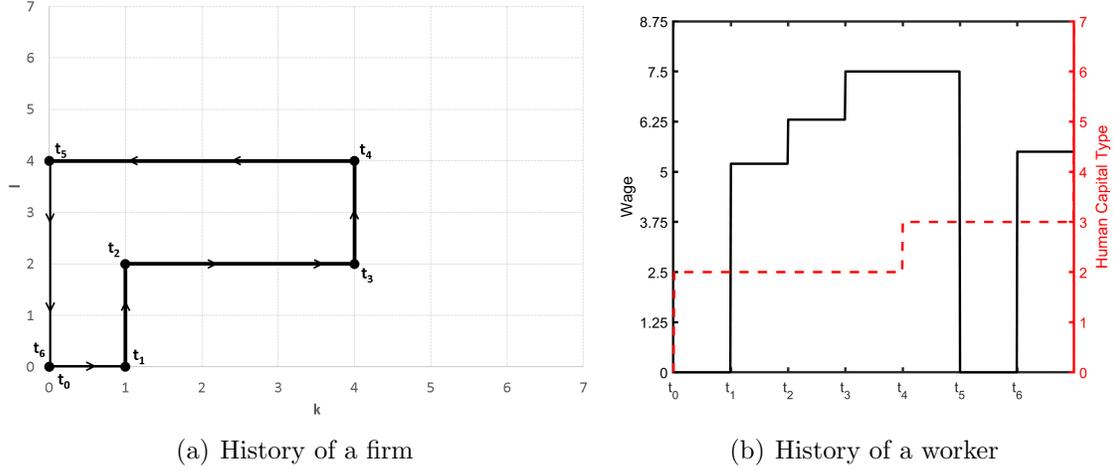


Figure 1: Sample Histories

member. Second, the firm may replace current employees with new hires. Replacement hiring is a natural consequence of the assumption that the firm keeps contacting workers even when it has formed a team and of the assumption that the firm faces decreasing returns to labor. Replacement hiring leads to a process of upgrading of the firm’s workforce that is analogue to the process of climbing the job ladder for workers. Replacement hiring also causes endogenous transitions of workers into unemployment. Third, Figure 1 shows that the firm may lose employees for a variety of reasons. In fact, employees may leave to move to a poaching firm, they may leave to unemployment, and they may exit the labor market.

Figure 1(b) presents a sample path of the history of a worker. At the beginning of the path, the worker has the second lowest level of human capital (i.e. $k = 2$) and is unemployed (i.e., $x = u$). In period t_1 , the worker meets a firm with one employee of type 6 (i.e. $y = (6, 0)$). Since the net value of the worker to the firm is greater than the value of unemployment, the worker is hired by the firm. The worker captures a fraction γ of the gains from trade, which implies a wage of $w = 5.2$. In period t_2 , the worker meets a poaching firm with one employee of type 4 (i.e. $y = (4, 0)$). Since the net value of the worker to the poaching firm is smaller than to his current employer, the worker does not move. However, since the net value of the worker to the poaching firm exceeds the value currently earned by the worker, the worker’s wage is raised to $w = 6.3$. In period t_3 , the worker meets a poaching firm with one employee of type 7 (i.e. $y = (7, 0)$). Since the net value of the worker to the poaching firm is greater than the net value to his current employer, the worker moves and his wage becomes $w = 7.5$. In period t_4 , the worker moves up the human capital ladder. However, his wage does not change because

his outside option (i.e. the value of unemployment) does not bind. In period t_5 , the firm meets an unemployed of type 4. The firm finds it optimal to hire the unemployed as a replacement to the worker who, then, becomes unemployed. Finally, in period t_6 , the worker meets a firm with one employee of type 6. The worker is hired at the wage $w = 5.5$.

The sample path in Figure 1(b) illustrates several important features of equilibrium wages. In period t_2 , we see that the wage of a worker during his tenure with a firm increases when the worker meets a poacher and the value of the worker to the poacher is lower than the value to the incumbent, but higher than the worker's value of remaining with the incumbent at the previous wage. In period t_3 , we see that the wage of a worker during a particular employment spell may increase when the worker moves from one firm to another firm. In general, whether the worker's wage goes up or down upon moving from one firm to another firm depends on the gains from trade and on the worker's old wage. In period t_4 , we see that the wage of a worker does not immediately reflect changes in his human capital. In fact, the wage of a worker only reflects his human capital at the time he was hired or at the time he last receives a sufficiently attractive outside offer from a poacher. In period t_6 , we see that the wage of a worker does eventually reflect changes in his human capital. In fact, in period t_2 , the worker has human capital h_2 and he moved from unemployment into a firm at the wage $w = 5.2$. In period t_6 , the worker has human capital h_3 and he moved from unemployment into the same type of firm at the wage $w = 5.5$.

3 Data and Measurement

In this section, we calibrate the model to the US labor market. To this aim, our main source of information is a matched employer-employee dataset, which contains information about US workers and firms. The purpose of the calibration is to measure the learning and the production functions, objects about which we have little prior knowledge. In Section 3.1, we introduce assumptions on functional forms and review all the parameters of the model that need to be calibrated. In Section 3.2, we describe the empirical moments used to calibrate the parameters. In Section 3.3, we report the values of the calibrated parameters, interpret their meaning, and illustrate the properties of the Stationary Equilibrium. In Section 3.4, we use the calibrated model to carry out two counterfactual exercises. In the first exercise, we shut down the effect of coworkers on an individual's human capital accumulation. The purpose of this exercise is to quantify the role of learning from coworkers in the formation of human capital, in the production output and in shaping

the pattern of sorting of workers across firms. In the second exercise, we examine how the economy would look like if the production function was modular. The purpose of this exercise is to assess the role of production complementarities between high human capital workers in the pattern of sorting of workers and, in turn, in the formation of human capital and in the production of output.

3.1 Parameters

We set the length of a period in the model to be one month. The human capital of a worker is an h_k from the set $H = \{h_1, h_2, \dots, h_N\}$ where $h_1 = 1$, $h_{k+1} = h_k + \Delta$ for $k = 1, 2, \dots, N-1$, $\Delta > 0$ and $N = 7$. The human capital accumulation process is described by the function $g(k_+|k, u)$. We assume that $g(k_+|k, u)$ takes the form

$$\begin{aligned} g(k-1|k, u) &= \alpha_u \text{ if } k > 1, 0 \text{ else;} \\ g(k|k, u) &= 1 - g(k-1|k, u). \end{aligned} \quad (7)$$

That is, an unemployed worker moves down the human capital ladder by one step with probability $\alpha_u \geq 0$, and otherwise maintains his stock of human capital. The specification of g in (7) captures the standard view that human capital depreciates during unemployment.

We assume that $g(k_+|k, 0)$ takes the form

$$\begin{aligned} g(k+1|k, 0) &= \alpha_0 \text{ if } k < N, 0 \text{ else;} \\ g(k|k, 0) &= 1 - g(k+1|k, 0). \end{aligned} \quad (8)$$

That is, a worker employed on his own moves up the human capital ladder by one step with probability $\alpha_0 \geq 0$, and otherwise maintains his stock of human capital. The specification of g in (8) captures the standard view that workers accumulate human capital by doing.

We assume that $g(k_+|k, \ell)$ takes the form

$$\begin{aligned} g(k+1|k, \ell) &= \begin{cases} \alpha_0 + \alpha_1^+ \max\{h_\ell - h_k, 0\} - \alpha_1^- \max\{h_k - h_\ell, 0\} & \text{if } k < N, \\ 0 & \text{else;} \end{cases} \\ g(k|k, \ell) &= 1 - g(k+1|k, \ell). \end{aligned} \quad (9)$$

That is, a worker employed in a team moves up the human capital ladder with a probability that depends on the human capital of his coworker relative to his own human capital. Specifically, the probability that the worker moves up the human capital ladder is α_0 if the coworker has the same human capital as the worker. The probability is $\alpha_0 + \alpha_1^+(h_\ell - h_k)$, with $\alpha_1^+ \geq 0$, if the coworker has more human capital than the worker. The probability is $\alpha_0 - \alpha_1^-(h_k - h_\ell)$, with $\alpha_1^- \geq 0$, if the coworker has less human capital than the worker. The parameter α_0 captures again the effect of learning by doing. The parameter α_1^+

captures the idea that a worker might learn more quickly by being in the company of a more knowledgeable coworker. The parameter α_1^- captures the idea that a worker might be hindered by being in the company of a less knowledgeable coworker. As long as the parameters α_1^+ and α_1^- are different, the effect of coworkers on learning is globally non-linear, which implies that the aggregate amount of learning depends on the composition of teams in the economy. We restrict α_1^+ and α_1^- to be non-negative based on the prior that a coworker with more human capital can be no worse company than a coworker with less human capital.³

The production process is described by the functions $b(k)$ and $f(\bar{k}, \bar{\ell})$. We assume that the production function for unemployed workers, i.e. the home-production function, has the form bh_k . That is, we assume that home-production is proportional to human capital. We assume that the production function for workers employed on their own to be $f(k, 0) = h_k$. That is, we assume that the output of a worker employed on his own is equal to his human capital. We assume that the production function for teams of workers is $f(k, \ell) = h_k + h_\ell + (h_k^{1/\rho} + h_\ell^{1/\rho})^\rho$. That is, we assume that the output of a team of two workers is equal to the output of each individual workers plus a term that, depending on the value of ρ , may be sub or super-modular in the human capital of the two workers. Note that the output of a firm with two workers always exceeds the output of two firms each employing one of the two workers.

The search-and-bargaining process is described by the parameters λ_u , λ_e , δ and γ , where λ_u is the probability that an unemployed worker meets a firm, λ_e is the probability that an employed worker meets a firm, δ is the probability that an employed worker moves into unemployment for exogenous reasons, and γ is the worker's bargaining power.

Finally, the entry-and-exit process is described by the parameter σ and by the probability distribution π_k . The parameter σ is the probability that a worker exits the labor market. The probability distribution π_k describes the human capital of a worker who enters the labor market. We specialize π_k to be a discretized exponential distribution with shape coefficient χ . We assume that workers enter the labor market at age 21.

3.2 Data

Compared to a canonical model in the vein of PR and CPR, ours has some additional parameters that need to be measured. Specifically, we need to measure the parameters α_1^+ and α_1^- , which capture the effect of the human capital of coworkers on the accumulation of human capital of an individual, and the parameter ρ , which captures the production

³The parameters α_1^+ and α_1^- are also restricted to guarantee $g(k_+|k, \ell) \in [0, 1]$.

complementarity between the human capital of coworkers and the human capital of an individual. To measure these parameters, we adopt a structural approach. That is, instead of looking for a “natural experiment” that might allow us to directly measure the parameters, we make use of the full equilibrium structure of the model to measure the parameters from the raw data. For instance, the problem with measuring α_1^+ and α_1^- is that the sorting of workers across firms is not random. Hence, observing an individual with highly-paid coworkers confounds information about the treatment received by the individual and the individual’s unobserved stock of human capital. In the face of this confounding effect, one could measure α_1^+ and α_1^- by looking for a natural experiment in which sorting is random. This is the approach typically used in the literature on peer effects in the classroom (see, e.g., Hoxby 2000). We do not follow this approach. Instead, we make use of the fact that the model provides us with a theory of sorting, and use the theory to infer α_1^+ and α_1^- from the raw data.

Let us explain in detail our approach to measuring α_1^+ and α_1^- . We use the Longitudinal Employer-Household Dynamics (LEHD) database between 2001 and 2008. The LEHD is a matched employer-employee dataset that covers 95% of jobs in the U.S. private sector. The LEHD includes identifiers for workers, identifiers for firms, demographic characteristics of workers, characteristics of firms, the identity of workers employed by firms and their earnings. We have access to the LEHD between the years 2001 and 2008 for 11 states: California, Illinois, Indiana, Maryland, Nevada, New Jersey, Oregon, Rhode Island, Texas, Virginia, and Washington. Even though we only use a 10% random sample of the data, for each worker in the sample we have summary statistics for the earnings of all his coworkers.

Using the 10% random sample of the LEHD, we construct a dataset in which the basic unit of observation is an EUE transition, i.e. a transition of an individual i from employment at a firm j into unemployment and then back into employment at a different firm. We refer to this as the EUE dataset. We restrict attention to EUE transitions that meet the following criteria: (i) individual i is a male between the ages of 24 and 65; (ii) firm j is a single-unit firm with a number of employees between 2 and 250; (iii) in year t , individual i was employed at the firm j for the whole year; in year $t + 1$, individual i was unemployed for at least one quarter (i.e., he had at least one quarter with less than \$1k of earnings); in year $t + 2$, individual i was employed for the whole year at a firm different from j . For each EUE transition that meets these criteria, we record the log annual earnings $w_{i,t}$ of individual i at firm j in year t , the log annual earnings $w_{i,t+2}$ of individual i in year $t + 2$, and the log of the average annual earnings $w_{-i,j,t}$ of the coworkers of individual i at firm j in year t . We also record the demographic characteristics of worker i and the characteristics of firm j . With a slight abuse of language we shall refer to w ’s

as wages. Table 1.D in Appendix D contains summary statistics for the EUE dataset.

We decompose the EUE dataset into observations where the wage of individual i at firm j in year t is lower than the average wage of his coworkers, and into observations where it is higher. Separately for the two subsets of the EUE dataset, we run the following OLS regression

$$w_{i,t+2} = \phi_0 + \phi_1 w_{i,t} + \phi_2 w_{-i,j,t} + \Phi X_{i,j,t} + \epsilon_{i,t}. \quad (10)$$

where $X_{i,j,t}$ includes dummies for calendar years, dummies for States, dummies for the 1-digit Standard Industrial Classification (SIC) code of firm j , firm j size, dummies for the race and education of worker i , quadratic terms for worker i 's age and tenure. Moreover, $X_{i,j,t}$ includes a dummy for the type of firm j , which is constructed by clustering of all firms into 10 groups based on the mean and the standard deviation of the firms' wages.

The main coefficient of interest is ϕ_2 , which measures the relationship between the wage of the coworkers of individual i in the current job and the wage that individual i earns in the next job (after going through an unemployment spell). While the estimates of ϕ_2 do not recover structural parameters of the model, they are related to α_1^+ and α_1^- . In fact, the wage of the coworkers of individual i is a (noisy) measure of the human capital of the coworkers to whom i has been exposed, and the wage of individual i in the next job is a (noisy) measure of his human capital after being exposed to those coworkers. We run the regression separately for individuals whose wage is lower and higher than the wage of their coworkers, as we suspect that the human capital of the coworkers might affect the accumulation of human capital of an individual differently depending on whether the coworkers are better or worse than the individual. We look at the individual's wage in the next job because in our model—as in all models with search frictions—the wage in the current job does not necessarily reflect changes in the individual's human capital. We look at individuals who pass through unemployment because we want to eliminate heterogeneity in outside options. Having eliminated heterogeneity in outside options, the individual's distribution of wages in the next job depends exclusively on the human capital of the worker at the time he is hired. We add the firm's type in the controls $X_{i,j,t}$ to purge the estimate of ϕ_2 from firm-specific effects.

Table 1 reports the result of the OLS regression (10). Columns (1) and (3) report the results of the regressions for observations in which individual i 's wage in year t is, respectively, lower and higher than the average wage of his coworkers. Columns (2) and (4) report the results of the same regressions when we add the controls $X_{i,j,t}$ and two lags of the wage of individual i . From columns (1) and (3), we can see that estimate of ϕ_2 is 0.145 when the individual's wage is lower than the average wage of his coworkers, and it is much lower—specifically, equal to 0.02—when the individual's wage is higher than

Table 1: EUE Sample: Regression (10)

| Dependent variable | (1) | (2) | (3) | (4) |
|----------------------|----------------------|---------------------------------------|------------------------|---------------------------------------|
| | Wage t+2 | Wage t+2 | Wage t+2 | Wage t+2 |
| Sample | $w_{i,t} < w_{-i,t}$ | $w_{i,t} < w_{-i,t}$ | $w_{i,t} > w_{-i,t}$ | $w_{i,t} > w_{-i,t}$ |
| Coworker Wage, t | 0.153*** (0.0176) | 0.145*** (0.0240) | 0.0202*** (0.00487) | 0.0406*** (0.0116) |
| Individual Wage, t | 0.460*** (0.0182) | 0.456*** (0.0185) | 0.731*** (0.00606) | 0.678*** (0.00686) |
| Individual Wage, t-1 | | 0.00565** (0.00252) | | 0.00210 (0.00158) |
| Individual Wage, t-2 | | -0.00221 (0.00160) | | -0.00249*** (0.000904) |
| Controls | None | Firm Type Dummies and Demographics | None | Firm Type Dummies and Demographics |
| R-squared | 0.268 | 0.317 | 0.470 | 0.488 |
| Round N | 15000 | 15000 | 46000 | 46000 |

Notes: SE clustered at SEIN level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Demographic controls include: firm size, state dummies, 1-digit sic dummies, race dummies, gender dummies, education dummies, quadratics in age and tenure, as well as year fixed effects. Firm Type Dummies: 10 types constructed as k-means clusters of coworker wage and coefficient of variation of coworker wage.

the average wage of his coworkers. These estimates imply that, for an individual paid less than his coworkers, a 10% increase in the coworkers' wage forecasts a 1.53% higher wage for the individual in the next job. For an individual paid more than his coworkers, a 10% increase in the coworkers' wage forecasts a 0.2% higher wage for the individual in the next job. From columns (2) and (4), we see that the estimates of ϕ_2 do not change much if we add the controls $X_{i,j,t}$ and lagged wages.

As mentioned above, the estimates of ϕ_2 do not recover the structural parameters α_1^+ and α_1^- . This is easy to understand. Since workers are not allocated to different firms at random, there will be a systematic relationship between the human capital of an individual and the human capital of his coworkers. If sorting is positive, the human capital of an individual will be positively correlated with the human capital of his coworkers. If sorting is negative, the human capital of an individual will be negatively correlated with the human capital of his coworkers. In either case, the wage of the individual and the wage of his coworkers are both (imperfectly) correlated with the individual's unobserved stock of human capital and, hence, they will both help forecast the individual's wage in the next job. In order to tease out α_1^+ and α_1^- from the estimates of ϕ_2 , it is then critical to calibrate the production function so that the model matches well the pattern of sorting in the data.

We measure the sorting of workers by breaking down the cross-sectional dispersion of wages into the variance of the average wage paid by different firms (*between-firm wage variance*) and the average variance of wages paid by the same firm to different workers (*within-firm wage variance*). This decomposition of wage dispersion is a good summary statistic for the sorting of workers across firms. In fact, if firms typically employ workers

with similar human capital, the wages paid by a firm to its employees will tend to be similar and, hence, the within-firm variance of wages will be low relative to the between-firm variance. If, in contrast, firms typically employ workers with different human capital, the wages paid by a firm to its employees will tend to be different and, hence, the within-firm variance of wages will be high relative to the between-firm variance. Using the LEHD, Spletzer (2017) finds that both between and within-firm variance account for approximately half of the overall cross-sectional dispersion in wages. We calibrate the parameter ρ —which controls the degree of supermodularity of the production function—so that the model matches the fraction of within-firm wage variance.

The parameters α_1^+ , α_1^- and ρ determine the value of an individual with a given stock of human capital in different types of firms and, in turn, the rate at which that individual leaves different types of firms (i.e., firms with different types of employees). For example, if workers do not learn from each other (i.e. α_1^+ and α_1^- are close to zero) and the production function is strongly supermodular (i.e. ρ is high), the value of an individual with high human capital is higher in a firm that employs other high human capital workers than in a firm that employs low human capital workers. Thus, the individual will be less likely to leave a firm employing high rather than low human capital workers. In contrast, if workers learn a lot from each other (i.e. α_1^+ is positive and large and α_1^- is close to zero) and the production function is modular (i.e. $\rho = 1$), the value of an individual with high human capital is higher in a firm that employs low human capital workers than in a firm that employs high human capital workers. Thus, the individual will be less likely to leave a firm employing low rather than high human capital workers.

Motivated by the above observations, we construct a second dataset using the LEHD. In this dataset, the basic unit of observation is an individual i and a year t . We refer to this as the EE dataset. We restrict attention to observations that meet the following criteria: (i) individual i is a male between the ages of 24 and 65; (ii) individual i was employed at the same firm j for the entire duration of year t ; (iii) firm j is a single-unit firm with a number of employees between 2 and 250. For each observation that meets these criteria, we record the log annual earnings $w_{i,t}$ of individual i at firm j in year t , the log of the average annual earnings $w_{-i,j,t}$ of individual i 's coworkers at firm j in year t , and an indicator function $EE_{i,t+1}$ which takes the value 1 if and only if, in year $t + 1$, the individual moves from firm j to a different firm without an intervening spell of unemployment. Table 2.D in Appendix D contains some summary statistics for the EE dataset.

Again, we decompose the EE dataset into observations where the wage of individual i is lower than the average wage of his coworkers, and into observations where it is higher.

Table 2: EE Sample: Regression (11)

| Dependent variable | (1) | (2) | (3) | (4) |
|----------------------|--------------------------|---------------------------------------|---------------------------|---------------------------------------|
| | EE, t+1 | EE, t+1 | EE, t+1 | EE, t+1 |
| Sample | $w_{i,t} < w_{-i,t}$ | $w_{i,t} < w_{-i,t}$ | $w_{i,t} > w_{-i,t}$ | $w_{i,t} > w_{-i,t}$ |
| Wage Distance | -0.0121*** (0.000674) | -0.0117*** (0.00112) | -0.000135 (0.000286) | 0.00648*** (0.000626) |
| Individual Wage | 0.00242*** (0.000657) | 0.00476*** (0.00107) | -0.00761*** (0.000391) | -0.00149** (0.000647) |
| Individual Wage, t-1 | | 0.000713*** (5.54e-05) | | 0.00174*** (4.81e-05) |
| Individual Wage, t-2 | | -0.000384*** (4.42e-05) | | 0.000163*** (3.68e-05) |
| Controls | None | Firm Type Dummies and Demographics | None | Firm Type Dummies and Demographics |
| R-squared | 0.001 | 0.009 | 0.001 | 0.015 |
| Round N | 2.442e+06 | 2.442e+06 | 7.206e+06 | 7.206e+06 |

Notes: SE clustered at SEIN level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Demographic controls include: firm size, state dummies, 1-digit sic dummies, race dummies, gender dummies, education dummies, quadratics in age and tenure, as well as year fixed effects. Firm Type Dummies: 10 types constructed as k-means clusters of coworker wage and coefficient of variation of coworker wage.

Separately for the two subsets, we run the following OLS regression

$$EE_{i,t+1} = \psi_0 + \psi_1 |w_{i,t} - w_{-i,j,t}| + \Psi X_{i,j,t} + \epsilon_{i,t}. \quad (11)$$

Table 2 reports the result of the regression. Columns (1) and (3) report the results of the regressions for observations in which the individual's wage is, respectively, lower and higher than the wage of his coworkers. The main coefficient of interest is ψ_1 . The estimate of ψ_1 is $-.012$ when the individual's wage is lower than the wage of his coworkers. The estimate of ψ_1 is not statistically different from zero when the individual's wage is higher than the wage of his coworkers. These estimates imply that, for individuals paid less than their coworkers, a 10% increase in the wage of the coworkers forecasts a 0.12 percentage point decline in the probability that the individual goes through an EE transition next year. In contrast, for individuals paid more than their coworkers, an increase in the wage of the coworkers does not forecast any change in the probability that the individual goes through an EE transition next year. In columns (2) and (4), we see that adding the controls $X_{i,j,t}$ and two lags of the wage of worker i does not change much the estimates of ψ_1 .

We now turn to the calibration of the other parameters describing the accumulation of human capital of an individual. The parameter α_0 denotes the probability that an individual moves up the human capital ladder when employed with someone who has the same knowledge as he does. We calibrate α_0 by targeting a measure of life-cycle wage growth. In particular, we target the ratio between the average wage of male workers aged 54 to 56 to the average wage of male workers aged 22 to 24. In the pooled 2000-2016

Current Population Survey (CPS), the ratio is about 1.9. The parameter α_u denotes the probability that an individual moves down the human capital ladder when unemployed. We calibrate α_u by targeting the correlation between the duration of the unemployment spell of an individual and the wage that the individual earns once he re-enters employment. In the pooled 2000-2016 CPS, the correlation is about -7% .

The calibration of the parameters describing the search-and-bargaining process is standard. The parameter δ denotes the probability with which an employed worker moves into unemployment for exogenous reasons. We calibrate δ so that the model matches the average EU rate in the 2000-2016 CPS (1.1% per month). The parameter λ_u denotes the probability with which an unemployed worker contacts a firm. We calibrate λ_u so that the model matches the average UE rate in the CPS (22% per month). The parameter λ_e denotes the probability with which an employed worker contacts a firm. We calibrate λ_e to match the average EE rate in the CPS (1.2% per month). The parameter γ denotes the fraction of the gains from trade accruing to the worker. We calibrate γ by targeting a measure of wage loss after an unemployment spell. To this aim, we isolate all workers who are employed in month 4 of the 2000-2016 CPS, unemployed in months 13, 14 or 15 and employed in month 16. For these workers, we compute the ratio of their wage in month 16 relative to their wage in month 4. We find that the average wage ratio is approximately 90%.

We calibrate the parameters describing the structure of the human capital ladder and the process of entry and exit of workers into the labor market as follows. We normalize the lowest level of human capital, h_1 , to be equal to 1 and choose the highest level of human capital, h_N , by targeting a measure of the cross-sectional variance of wages. In particular, we target the ratio between the 90th percentile and the 10th percentile of wages in the 2000-2016 CPS (4.57). We calibrate the shape parameter, χ , of the exponential distribution for the human capital of workers entering the labor market by targeting the ratio between the 90th percentile and the 10th percentile of wages for workers aged 22 to 24 in the 2000-2016 CPS (2.88). We calibrate the exit rate of workers, σ , so that a worker spends, on average, 35 years in the labor market.

We want to make sure that we also match some properties of the dynamics of wages for individual workers. To this aim, we include two targets about individual wage dynamics. Using the CPS, we isolate workers who are employed in month t and in month $t + 12$. First, we target the average wage growth between months t and $t + 12$ for these workers, which we find to be equal to 2%. We also target the variance of the change in the wage between months t and $t + 12$ for these workers, which we find to be equal to 0.43.

Finally, we need to choose a value for the discount factor β . We choose the value of

β to be 0.988, which implies a discount factor of 15% per year. It is well-known that models in the style of PR and CPR generate some negative wages when β is set to a more standard value of 5% per year and when workers have linear utility in income (and, hence, an infinite elasticity of intertemporal substitution). Intuitively, when β is high and workers have linear utility, wages out of unemployment are sometimes negative—as workers expect large raises when they receive an outside offer and do not need to be compensated for these large changes in income. We do not want to introduce curvature in the worker’s utility function, as we would not be able to characterize the equilibrium by simply computing the joint value of different production units. Instead, to avoid the problem of negative wages, we choose to set the discount factor β to 15% per year.

3.3 Measurement and Properties of Equilibrium

Table 3 reports the calibrated value of the parameters, the empirical moments used to calibrate them and the fit of the calibrated model to these empirical moments. The calibrated value of the parameter ρ is 2.35. Since $\rho > 1$, the calibrated production function is supermodular and, as explained before, it pushes the equilibrium pattern of sorting towards PAM. To better gauge the meaning of $\rho = 2.35$, note that it implies that the sum of the output generated by one production unit with workers $(k, \ell) = (7, 7)$ and another production unit with workers $(k', \ell') = (1, 1)$ is 5% higher than the output generated by two production units with workers $(k', \ell) = (\ell', k) = (1, 7)$. The value of ρ is the one required by the model to guarantee that half of the variance of wages is due to dispersion in the average wage of different firms and half is due to dispersion in the wage paid by the same firm to different employees.

The calibrated values of the parameters α_0 , α_1^+ and α_1^- are, respectively, 0.003, 0.017 and 0.0002. The calibrated value for α_0 (together with the calibrated value of a step in the human capital ladder) implies that the expected growth rate of human capital for a worker of type $k = 4$ employed with a coworker of the same type is approximately 0.66% per year. The value of α_0 is the one required by the model to match the lifecycle wage growth. The calibrated value for α_1^+ implies that the expected growth rate of human capital for a worker of type $k = 4$ employed with a coworker of type $\ell = 5$ (which has 15% more human capital) is approximately 1.61% per year. Similarly, the calibrated value for α_1^- implies that the expected growth rate of human capital for a worker of type $k = 4$ employed with a coworker of type $\ell = 3$ (which has 15% less human capital) is approximately 0.65% per year. As expected, the calibrated values of α_1^+ and α_1^- are different from the regression coefficients in Table 1 even though the model matches these coefficients well. Since α_1^+ is positive and α_1^- is basically zero, workers tend to catch up

Table 3: Calibration: Model Moments vs. Data Moments (Source: CPS and LEHD)

| Parameter | Description | Value | Moment Description | Data | Model |
|--------------|-----------------------------------|-------|---|--------|--------|
| ρ | Complementarity in production | 2.355 | Between firm wage variance | 0.503 | 0.435 |
| α_0 | Base prob. moving up hc ladder | 0.004 | Mean wage at 54/ Mean wage at 24 | 1.854 | 1.877 |
| α_1^+ | Effect of coworker hc for $l > k$ | 0.018 | Regression of $EE_{i,t+1}$ on $ w_{-i,t} - w_{i,t} (w_{i,t} < w_{-i,t})$ | -0.012 | -0.014 |
| α_1^- | Effect of coworker hc for $k > l$ | 0.000 | Regression of $EE_{i,t+1}$ on $ w_{-i,t} - w_{i,t} (w_{i,t} > w_{-i,t})$ | 0.000 | -0.003 |
| | | | Regression of $w_{i,t+2}$ on $w_{-i,t} (w_{i,t} < w_{-i,t})$ | 0.153 | 0.130 |
| | | | Regression of $w_{i,t+2}$ on $w_{-i,t} (w_{i,t} > w_{-i,t})$ | 0.020 | 0.021 |
| | | | Regression of $w_{i,t+2}$ on $w_{i,t} (w_{i,t} < w_{-i,t})$ | 0.460 | 0.120 |
| | | | Regression of $w_{i,t+2}$ on $w_{i,t} (w_{i,t} > w_{-i,t})$ | 0.731 | 0.874 |
| | | | Mean annual wage growth | 0.020 | 0.035 |
| | | | Std of annual wage growth | 0.440 | 0.513 |
| α_u | Prob. of moving down hc ladder | 0.059 | Wage ratio before and after unemp. spell | 0.903 | 0.710 |
| h_N | Highest human capital type | 2.714 | p90 wage/ p10 wage | 4.597 | 6.164 |
| χ | Shape of new entrant dist. | 3.016 | Mean wage at 24/ Mean wage p90 wage at 24/p10 wage at 24 | 0.582 | 0.623 |
| λ_u | Meeting rate for unemployed | 0.324 | UE rate | 3.026 | 4.432 |
| λ_e | Meeting rate for employed | 0.293 | EE rate | 0.224 | 0.228 |
| δ | Exogenous separation rate | 0.005 | EU rate | 0.017 | 0.012 |
| γ | Worker's bargaining power | 0.646 | Correlation of log wage and duration (in weeks) for job finders | 0.011 | 0.009 |
| | | | | -0.073 | -0.049 |

with more knowledgeable coworkers, but are not dragged down by less knowledgeable ones. As explained in Anderson (2015), this feature of the learning function pushes the equilibrium pattern of sorting towards NAM.

The calibrated values of the search parameters λ_u , λ_e and δ are, respectively, 0.32, 0.29 and 0.005. Given these values, the model does a pretty good job in matching the empirical UE, EU and EE rates. The calibrated bargaining power of the worker, γ , is 0.64. Given this value, the model generates a 70% ratio (compared to 90% in the data) between the wage before and after an unemployment spell.⁴

Given the calibrated parameter values, we now examine the equilibrium policy functions. Figure 2 shows whether a worker of type k (vertical axis) in employment status x (horizontal axis) is hired by a firm in state y and, in case the firm has two employees, which one the worker is replacing. Consider a worker with the lowest human (i.e. $k = 1$) who is unemployed (i.e. $x = u$). The policy functions for this worker are shown in the top-left quadrant of Figure 2. The worker is hired by firms in state $(0, 0)$ or in state $(i, 0)$. The worker is not hired by firms in state (i, j) . A firm in state $(0, 0)$ hires the worker because, when employed, the worker accumulates human capital (by doing) and, when matched with a coworker, the worker will have a higher marginal product than in unemployment.⁵ A firm in state $(i, 0)$ hires the worker because, when employed with a coworker of type i , the worker accumulates human capital (by doing and by learning from the coworker) and has a higher marginal product than in unemployment. A firm with employees (i, j) does not hire the worker. Hiring the worker requires sending one of the firm's employees into unemployment. Since the worker's productivity on the job (relative to his productivity in unemployment) is lower than the relative productivity of either employee, hiring the worker would reduce overall output. Since the worker's human capital is lower than the human capital of either employee, hiring the worker would increase the overall human capital accumulation. However, this benefit is smaller than the cost of lowering production.

Now, consider a worker with the lowest human capital (i.e. $k = 1$) who is employed on his own (i.e. $x = 0$). The policy functions for this worker are shown in the second quadrant in the top row of Figure 2. The worker is not hired by firms in state $(0, 0)$, since the gains from trade are zero. The worker is hired by firms in state $(i, 0)$, as the marginal product of the worker when employed alongside a coworker of type i is greater than his marginal product when employed on his own. The worker is not hired by firms in state

⁴The model does not match the targeted moments exactly because there are more moments than parameters.

⁵Since workers search almost as well on as off the job, the consequences of hiring on the option value of searching are negligible.

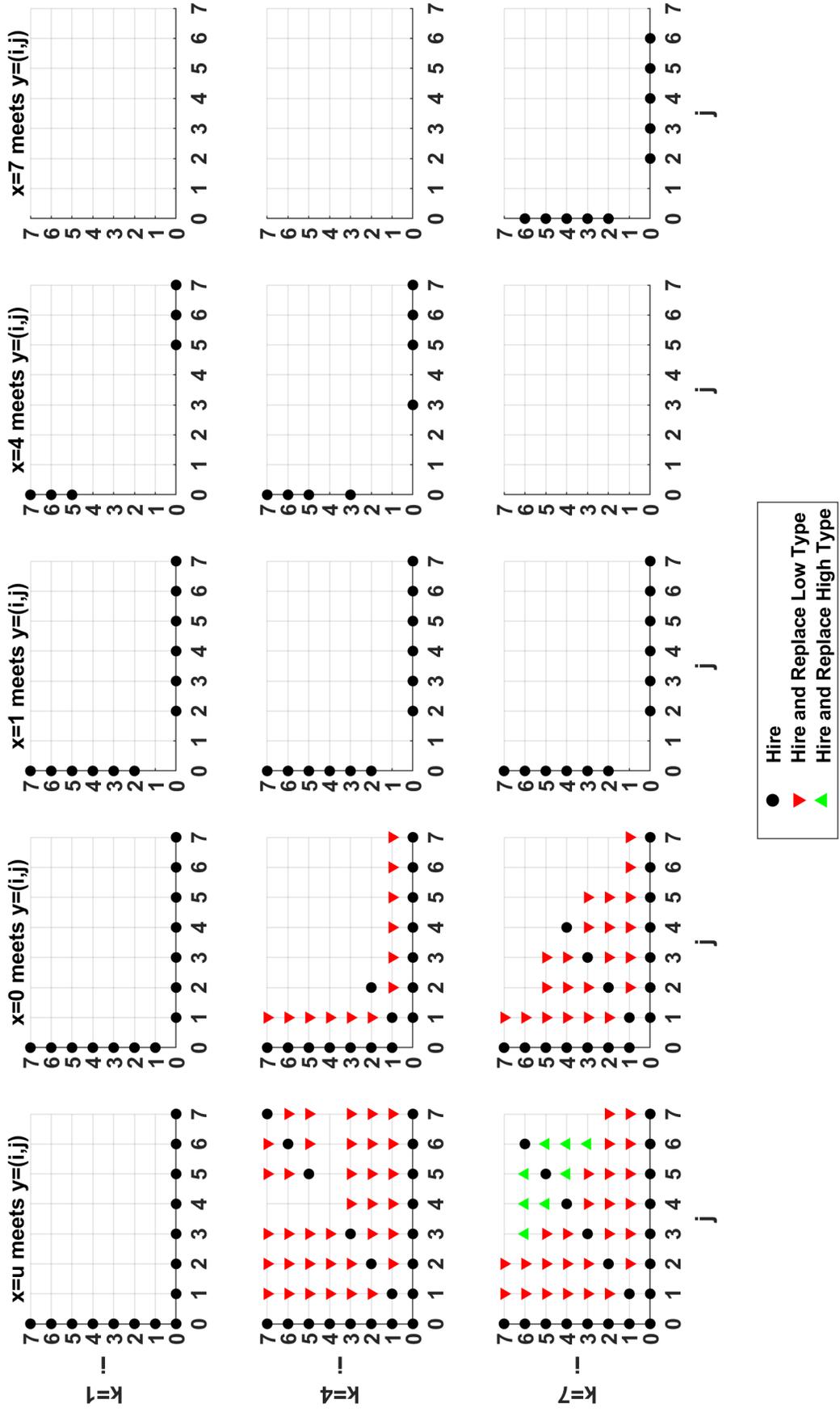


Figure 2: Equilibrium Policy Functions

(i, j) . Intuitively, hiring the worker would have the costly consequence of sending one of the firm's employees into unemployment where they would produce less and lose human capital. This cost is greater than the benefit of allowing the worker to learn not only by doing, but also from being in the company of a more knowledgeable coworker.

Next, consider a worker with the lowest human capital (i.e. $k = 1$) who is employed with a coworker of type $\ell = 1, 4$ or 7 (i.e. $x = \ell$). The policy functions for this worker are shown respectively in the third, fourth and fifth quadrants in the top row of Figure 2. The worker is never hired by firms in state $(0, 0)$. This is because productivity and human capital accumulation are lower when the worker is employed on his own than when he is employed with a coworker. The worker is hired by firms in state $(i, 0)$ only if $i > \ell$. This is so because the marginal productivity and the human capital accumulation are higher when the worker is employed with a better coworker.

The second row in Figure 2 shows the policy functions for a worker with an intermediate stock of human capital (i.e. $k = 4$). First, suppose that the worker is unemployed (i.e. $x = u$). The worker is hired by firms in state $(0, 0)$ or $(i, 0)$. These are the same policies as for a worker of type 1, and the economics behind them is the same. The worker is hired by a firm in state (i, j) if $\min\{i, j\} \leq 3$ or if $\min\{i, j\} \geq 5$. These policies are different than for a worker of type 1, and they deserve some explanation. A firm in state (i, j) with $\min\{i, j\} \leq 3$ hires an unemployed worker of type 4 to replace the worst of its employees. Since the worker's productivity on the job (relative to his productivity in unemployment) is higher than the relative productivity of the replaced employee, hiring the worker increases overall output. Since the worker's human capital is higher than the human capital of the replaced employee, the other employee accumulates human capital more quickly. These benefits are larger than the cost of the lost human capital accumulation for the replaced employee. An unemployed worker of type 4 is also hired by a firm in state (i, j) with $\min\{i, j\} \geq 5$ to replace the worst of the firm's employees. The firm finds it optimal to do so because the worker has more to learn from the firm's best employee than the firm's worst employee.

Next, suppose that the worker is employed on his own (i.e. $k = 4, x = 0$). The worker is not hired by firms in state $(0, 0)$, as the gains from trade are zero. The worker is hired by firms in state $(i, 0)$, as the value of the worker is higher when he is employed with a coworker of type i than by himself. The worker is also hired by a firm in state (i, j) if $\min\{i, j\} = 1$ or $i = j = 2$. Intuitively, the firm finds it optimal to do so because the cost of sending one of its employees into unemployment is lower than the additional output that the worker can produce with a coworker than on his own.

Finally, suppose that the worker is employed with a coworker of type $\ell = 1, 4, 7$ (i.e.

$k = 4, x = \ell$). These policies follow a similar pattern as those for a worker of type 1. That is, the worker moves when he meets a poaching firm with only one employee who happens to be better than his current coworker. However, when a worker of type 4 is employed with someone who is just like him, he also moves to a poaching firm with a single employee of type 3. This is because the benefit of the worker teaching to an individual of type 3 exceeds the cost of the worker producing with an individual of type 3 rather than 4.

The third row in Figure 2 shows the policy functions for a worker with the highest stock of human capital (i.e. $k = 7$). These policy functions follow the same general pattern as those for workers of type 1 and 4. There are two exceptions to this pattern. First, an unemployed worker of type 7 is sometimes hired to replace the firm's best rather than the firm's worst employee. The firm finds this optimal because the worker can teach more to the worst employee than to the best employee. Second, a worker of type 7 does not always move from a team to a poaching firm with a single employee who is better than the worker's current teammate. This is because the worker can teach more to his current teammate.

Taking a bird's eye view on Figure 2, we see forces pushing equilibrium towards NAM and forces pushing towards PAM. Transitions of high human capital individuals from firms employing other high human capital workers to firms employing low human capital workers push equilibrium towards NAM. These transitions are caused by the non-linearity of the human capital accumulation function, which implies that a better worker can teach a worse worker without any consequences on his own human capital accumulation. Firms replacing low human capital employees with high human capital individuals hired from unemployment and from firms with incomplete teams pushes the equilibrium towards PAM. The extent to which firms upgrade their workforce through replacement hiring depends on the degree of supermodularity of the production function (which determines the productivity gains caused by upgrading) and on the asymmetry of the human capital accumulation function (which determines the loss in human capital growth caused by upgrading). Interestingly, some upgrading would take place even with a modular production function. This is because the productivity gains caused by upgrading depend on a comparison between the productivity when employed relative to the productivity when unemployed for the new hire and the old employee. And the relative productivity of a worker is increasing in his human capital stock even when the production function is modular.

Having examined in detail the policy functions, we can now easily make sense of the other features of equilibrium. Figure 3 contains the histogram of the unemployment rate (left panel) and of the monthly transition rates (right panel) for workers with different

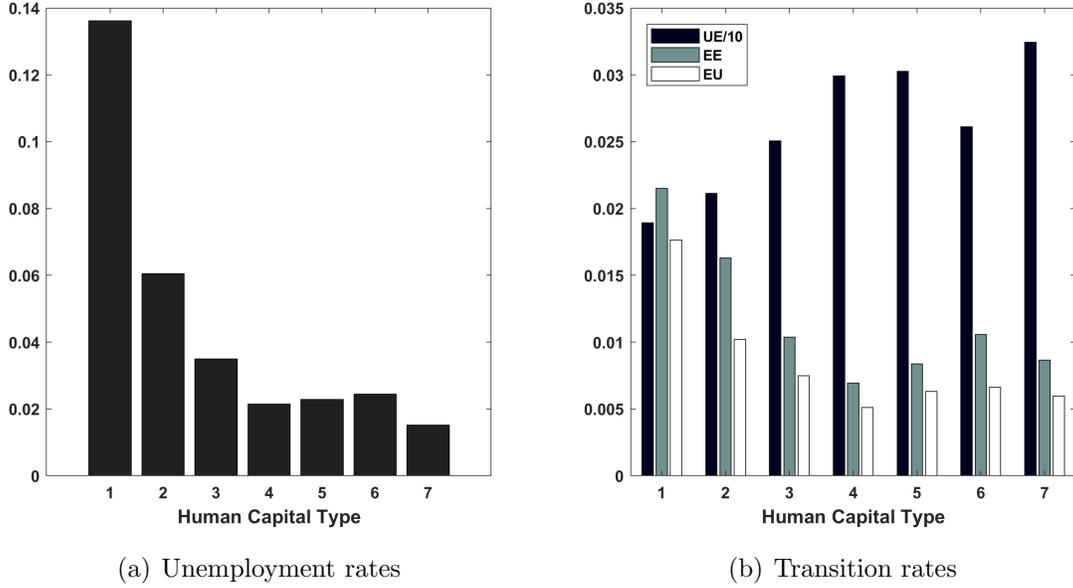


Figure 3: Unemployment and Transition Rates by Type

human capital. For workers with the lowest human capital, the unemployment rate is 13.5%, the unemployment to employment rate (UE) is about 18%, the employment to unemployment rate (EU) is 1.7%, and the rate of transition from one employer to another (EE) is 2%. For workers with more human capital, the unemployment rate is lower, the UE rate is higher and the EU rate is lower. Indeed, for workers with the highest human capital, the unemployment rate is about 1.5%, the UE rate is 27%, the EU rate is 0.5% and the EE rate is 0.85%. The UE rate is higher because workers with more human capital are more frequently hired out of unemployment as replacements for existing employees. The EU rate is lower because workers with more human capital are less frequently replaced by new hires and sent into unemployment. Hence, workers with more human capital end up having a lower unemployment rate.

Figure 4 is a histogram of the distribution of human capital among workers entering the labor market (Entering workers) and among the entire population (All workers: baseline). The distribution of human capital in the population is higher—in the sense of first-order stochastic dominance—than the distribution of human capital among new entrants. The difference between the two distributions reflects the amount of human capital accumulation that takes place over the life-cycle, which is due to the standard learning by doing mechanism (as captured by the parameter α_0) and to learning from more knowledgeable coworkers (as captured by the parameter α_1^+). In aggregate, the stock of human capital among all workers is more than 50% higher than among new entrants.

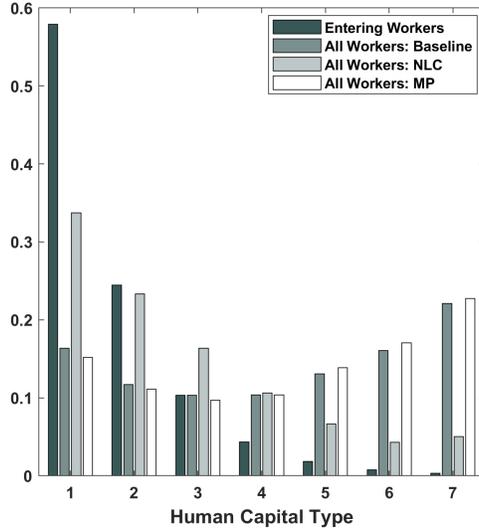


Figure 4: Human Capital Distribution

Figure 5 shows the average wage for workers of different age (left panel) and the average human capital for workers of different age (right panel). Consistently with the data, the model is such that the average wage doubles over the lifecycle. The average wage increases with age because older workers have more human capital (as they climbed the human capital ladder), they are employed in firms in which their marginal value is higher (as they moved towards better jobs), and they capture a larger share of the marginal value (as they received more outside offers). Figure 5 also shows the lifecycle profile of the average wage and human capital for workers with different initial conditions.⁶ Note that initial conditions have persistent effects. Even after 30 years in the labor market, there is no full convergence in the human capital and in the wage of workers who entered the market with different amounts of knowledge.

The model also does a decent job at matching the dispersion of wages in the data. The ratio between the 90th percentile and the 10th percentile in the overall wage distribution is around 6, while in the data is 4.5. The dispersion of wages is generated by fundamental and by circumstantial differences among workers. As a matter of fundamentals, workers are paid different wages because they have different human capital. As a matter of circumstances, workers with the same human capital are paid different wages because they are employed in jobs in which their marginal value is different and because they capture different fractions of their marginal values. The ratio between the 90th percentile and the 10th percentile in the wage distribution for workers aged 24 is around 4, while in the data

⁶We only plot the lifecycle for workers entering the labor market with human capital type 1, 2 and 3. These types account for about 90% of all workers entering the market.

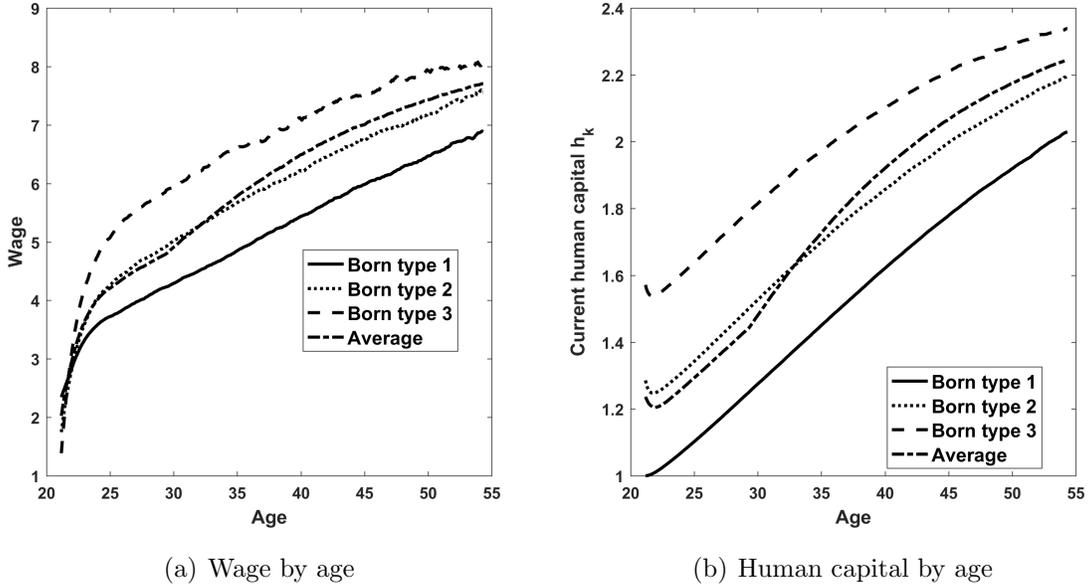


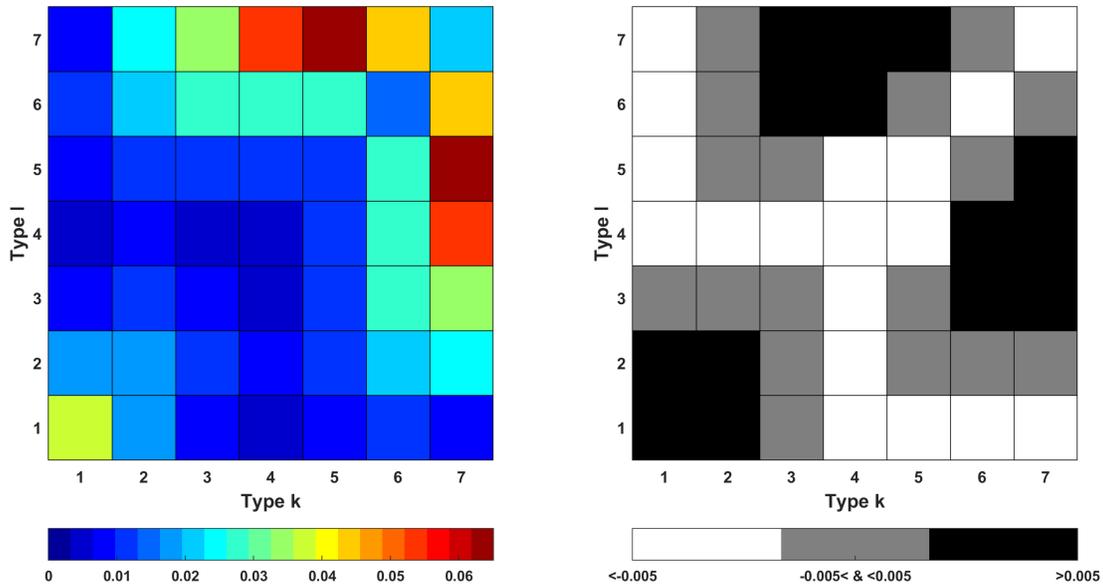
Figure 5: Lifecycle of Wages and Human Capital

is 3. There is less wage dispersion among young workers than in the overall population simply because the 90-10 percentile ratio of human capital among young workers is lower than in the overall population.

Finally, we turn to the analysis of the equilibrium pattern of sorting. Figure 6(a) illustrates the distribution of coworkers.⁷ If the equilibrium was PAM, the entire distribution would lie along the anti-diagonal (the one going from the South West to the North East corner). If the equilibrium was NAM, the entire distribution would lie along the main diagonal (the one going from the North West to the South East corner). The distribution in Figure 6(a) is neither PAM nor NAM. To understand the equilibrium pattern of sorting, it is useful to look at Figure 6(b), which shows the distribution of coworkers in equilibrium *net* of the distribution of coworkers that would obtain under random assignment.⁸ Consider workers with the lowest human capital ($k = 1$). The equilibrium distribution of workers of type 1 across coworkers is, relative to a random-assignment, decreasing in the coworker's type. That is, workers of type 1 are more than proportionally matched with low human capital coworkers. The equilibrium distribution of workers of type 4 across coworkers is, relative to random assignment, negative for all coworker's types except 6 and 7. That is, workers of type 4 are disproportionately matched with the highest types of coworkers. The equilibrium distribution of workers of type 7 across coworkers is, net of random-assignment, hump-shaped with respect to the coworker's type. That is, workers

⁷In Appendix D, the reader can find a 3D illustration of the equilibrium distribution of coworkers.

⁸To construct the distribution under random assignment, we only consider workers employed in teams.



(a) Distribution

(b) Distribution net of random assignment

Figure 6: Equilibrium Pattern of Sorting: Baseline

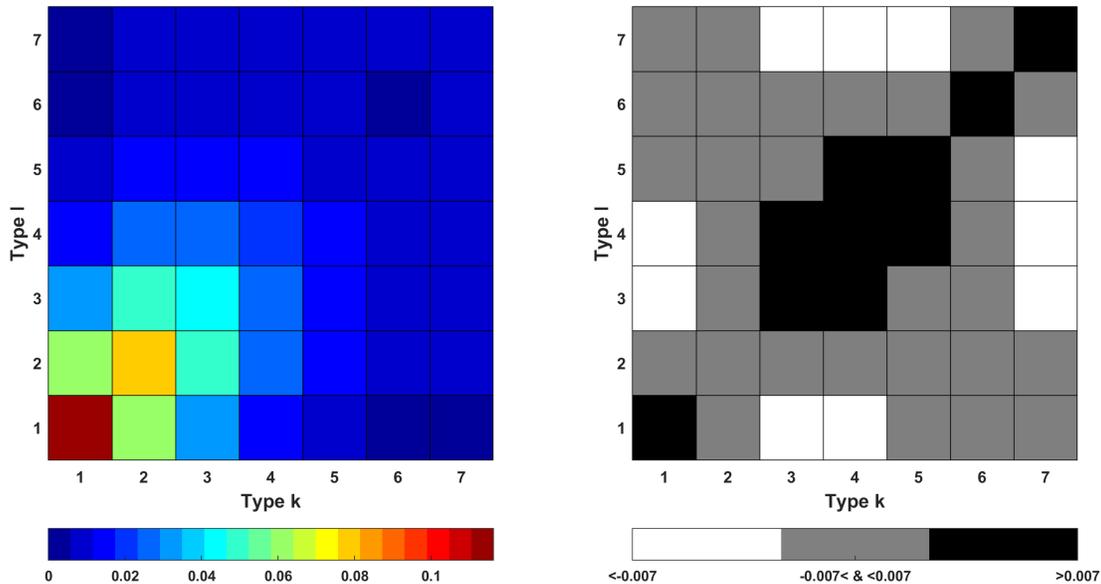
of type 7 are disproportionately matched with workers with average human capital.

3.4 Counterfactuals

We now carry out two counterfactual exercises, designed to understand and quantify the role of the two distinctive features of our model—i.e. learning from coworkers and production complementarities between coworkers—in determining the formation of human capital, the production of output, and the shape of the pattern of sorting of workers across firms.

3.4.1 Learning from Coworkers

In the first counterfactual, we set the learning parameters α_1^+ and α_1^- to zero, so that all employed workers accumulate human capital at the same speed irrespective of the identity of their coworker. The purpose of this counterfactual is to quantify the importance of learning from coworkers for the accumulation of human capital in the population, the production of output in the economy, and the equilibrium pattern of sorting between coworkers. We refer to this counterfactual as the No Learning from Coworkers (NLC) model.



(a) Distribution

(b) Net distribution in NLC minus Baseline

Figure 7: Equilibrium Pattern of Sorting: NLC

Figure 4 contains histograms for the distribution of human capital among workers entering the labor market, among all workers in the economy in the NLC model, and among all workers in the economy in the baseline model. Not surprisingly, the distribution of human capital in the economy is lower in the NLC than in the baseline model. The average human capital in the NLC model is about 24% lower than in the baseline. This number understates the importance of learning from coworkers in the accumulation of human capital. Indeed, the difference between the average human capital in the population and the average human capital at entry is 60% lower in the NLC than in the baseline model. In other words, learning from coworkers accounts for 60% of all the stock of human capital that is accumulated during the work-life!⁹ Clearly, as human capital is lower in the NLC model, so is output. Specifically, aggregate output is about 20% lower in the NLC than in the baseline model.

Figure 7(a) illustrates the distribution of coworkers in the NLC model.¹⁰ Since the NLC and the baseline models differ with respect to the distribution of worker types (as shown in Figure 4), comparing Figure 7(a) with Figure 6(a) does not reflect only

⁹Recall that about half of the stock of human capital is accumulated on the job. Of this half, 60% is accumulated because of more knowledgeable coworkers teaching to less knowledgeable ones.

¹⁰We refer the reader to Appendix D for a 3D illustration of the distribution of coworkers for the NLC and MP models.

differences in the pattern of sorting between the two models. To isolate differences in the pattern of sorting, Figure 7(b) compares the distribution of coworkers *net* of random assignment in the NLC model with the distribution net of random assignment in the baseline model. From this comparison we see that the distribution has more density (relative to random assignment) along the 45°-line and less density away from the 45°-line in the NLC than in the baseline model. Relatedly, we find that the correlation between the human capital of coworkers is 30% in the NLC model, while it is only 10% in the baseline model. Informally, we could say that in the NLC model sorting is more positive than in the baseline.¹¹

It is easy to understand why the pattern of sorting is more positive in the NLC model than in the baseline. The calibrated values of α_1^+ and α_1^- imply that the human capital accumulation function is non-linear. Workers accumulate human capital more quickly when paired with someone who is better than them. However, workers do not accumulate human capital less quickly when paired with someone who is worse than them. This implies that pairing workers with different human capital leads to more human capital accumulation than pairing workers with the same human capital. In the NLC model, we make the growth rate of human capital of an individual independent of the identity of his coworker and, hence, this effect disappears. As a result, the pattern of sorting becomes more positive. In Appendix D, the reader can find the equilibrium policy function for the NLC model.

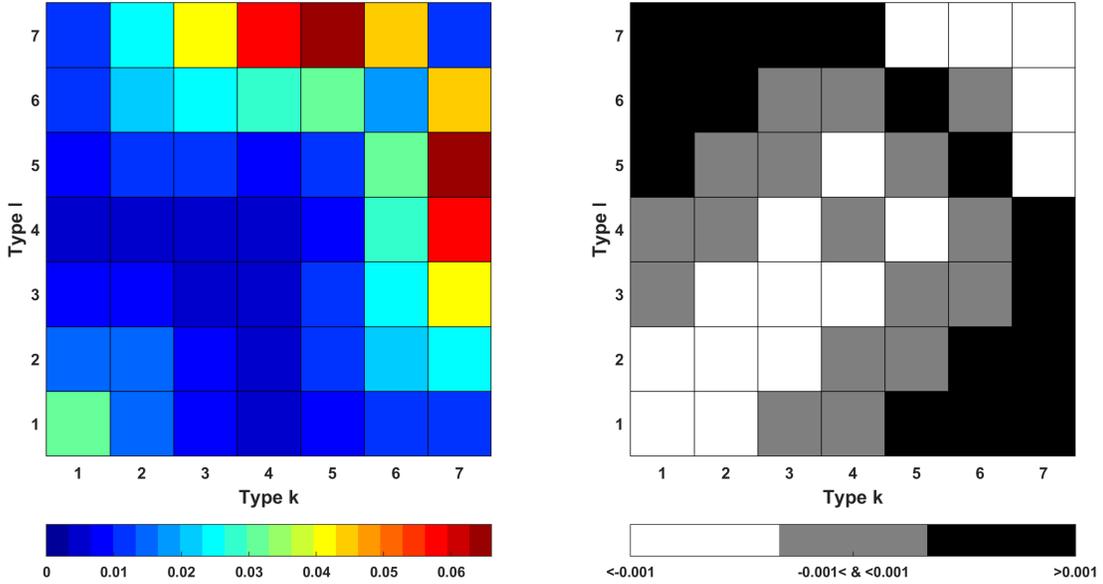
3.4.2 Modular Production

In the second counterfactual, we set the parameter ρ in the production function $f(k, \ell)$ to 1, so that the output produced by a team of workers becomes additive in the workers' human capital. As changes in ρ affect not only the supermodularity but also the level of the production function $f(k, \ell)$, we rescale $f(k, \ell)$ so that the output produced by two workers of the same type remains the same as in the baseline calibration.¹² The purpose of the counterfactual is to quantify the importance of production supermodularity for the accumulation of human capital in the population, the production of output in the economy, and the equilibrium pattern of sorting. We refer to this counterfactual as the Modular Production (MP) model.

¹¹Anderson and Smith (2018) expose some of the difficulties associated with formalizing the notion of “more positive” sorting.

¹²Formally, in this counterfactual the production function is

$$f(k, \ell) = h_k + h_\ell + \frac{2^{2.35}}{2^1}(h_k + h_\ell).$$



(a) Distribution

(b) Net distribution in MP minus Baseline

Figure 8: Equilibrium Pattern of Sorting: MP

Figure 8(a) illustrates the distribution of coworkers in the MP model. To isolate differences in the pattern of sorting from differences in the distribution of types in the MP and in the baseline models, Figure 8(b) compares the distribution of coworkers *net* of random assignment in the MP model with the distribution net of random assignment in the baseline model. We immediately see from this panel that the distribution has less density (relative to random assignment) along from the 45°-line and more density away from the 45°-line in the MP than in the baseline model. Relatedly, we find that the correlation between the human capital of coworkers is only 1% in the MP model, while it is 10% in the baseline model. Informally, we could say that in the MP model sorting is less positive than in the baseline. This finding is easy to explain. The calibrated value of ρ implies that production function $f(k, \ell)$ is supermodular, which gives firms an incentive to create teams of workers who have a similar stock of human capital. In the MP model, we make the production function modular and, thus, we remove this incentive. As a result, the pattern of sorting becomes less positive. In Appendix D, the reader can find the equilibrium policy function for the MP model.

As sorting across coworkers is less positive in the MP than in the baseline model, human capital diffuses more rapidly from more to less knowledgeable individuals. Figure 4 contains histograms for the distribution of human capital among workers entering the

labor market, among all workers in the economy in the MP model, and among all workers in the economy in the baseline model. The average human capital is about 2% higher in the MP than in the baseline model. The difference between the average human capital and the average human capital at entry is 4% higher in the MP than in the baseline model. In other words, the supermodularity of the production function reduces the accumulation of human capital during the work-life by 4%. Similarly, aggregate output is 2% higher in the MP than in the baseline model.

4 Planner Problem and Optimal Taxation

In this section, we solve the social planner problem and characterize its steady state. In Section 4.1, we show that the stationary equilibrium is inefficient, in the sense that it is different from the steady state of the planner’s problem. In Section 4.2, we construct a system of taxes and subsidies for unemployed workers and different units of production with the property that the steady state of the planner’s problem is a stationary equilibrium. In Section 4.3, we show that—given the calibrated parameter values—the equilibrium is such that low human capital employees are inefficiently often replaced by better workers. As a result of this inefficiency, the equilibrium pattern of sorting of coworkers is too positive, the stock of human capital in the economy is too low, and so is output. The optimal system of taxes and subsidies serves the purpose of discouraging firms from replacing low human capital employees with better workers.

4.1 Inefficiency of Equilibrium

The value for a utilitarian social planner, evaluated at the beginning of the production stage, is such that

$$\begin{aligned}
 & S(u_p, e_p) \\
 = & \max \sum_k [u_p(k)b(k) + e_p(k, 0)f(k, 0) + \sum_\ell e_p(k, \ell)f(k, \ell)/2] + \beta S(u_{p+}, e_{p+}). \tag{12}
 \end{aligned}$$

The state of the problem is $\{u_p, e_p\}$, where $u_p(k)$ is the measure of unemployed workers of type k , $e_p(k, 0)$ is the measure of workers of type k who are employed without a coworker, and $e_p(k, \ell)$ is the measure of workers of type k who are employed with a coworker of type ℓ . Consistency requires $e_p(k, \ell) = e_p(\ell, k)$. The objective of the planner is to maximize the present value of output discounted at the factor β . At the learning stage, the planner takes as given the evolution of the workers’ human capital. At the entry-and-exit stage, the planner takes as given the death and birth of workers. At the search-and-matching stage, the planner takes as given the meetings between workers and firms. For every

meeting between a worker of type k in state x and a firm in state y , the planner chooses whether the worker should be hired and, if the firm has already two employees, which one the worker should replace. At the dismissal stage, the planner chooses which firm-worker matches to break. The planner's choices determine the state $\{u_{p+}, e_{p+}\}$ in the next period. In Appendix C, the reader can find a complete formulation of the planner's problem, as well as the details of the solution.

We want to characterize the properties of the steady-state of the solution to the planner's problem. To this aim, we denote as $\{u_p^*, e_p^*\}$ the steady-state distribution of workers measured at the production stage, and as $\{u^*, e^*\}$ the steady-state distribution of workers measured at the search-and-matching stage. We denote as $V_{0,0}^*$ the value of an additional idle firm to the planner. We denote as $V_{k,0}^*$ the value of an additional production unit with a single worker k . We denote as $V_{k,\ell}^*$ the value of an additional production unit with a team of workers (k, ℓ) . We denote as U_k^* the value of an additional unemployed worker of type k . These objects are evaluated at the steady state and, for this reason, we omit their dependence from the state of the problem. As we did for the equilibrium, we find it useful to denote \hat{V} the value of a production unit at the dismissal stage.

The marginal value $V_{0,0}^*$ of an idle firm is such that

$$V_{0,0}^* = 0 + \beta \left\{ \left[\sum_{i,x} q^*(i,x) \max\{\hat{V}_{i,0}^* - \hat{V}_{0,0}^* - v^*(i,x), 0\} \right] + \hat{V}_{0,0}^* \right\}. \quad (13)$$

The value to the planner of an idle firm is the sum of the output generated by the firm in the current period (zero) and the present value of the output generated by the firm from the next period onwards. In the next period, the firm meets a worker of type i in state x with probability $q^*(i,x)$, where $q^*(i,u) = \lambda_u u_i^*$, $q^*(i,0) = \lambda_e e_{i,0}^*$ and $q^*(i,j) = \lambda_e e_{i,j}^*$. In this case, the continuation value of the firm is given by $\hat{V}_{0,0}^*$ plus the difference between the marginal value of the worker at the firm, $\hat{V}_{i,0}^* - \hat{V}_{0,0}^*$, and the marginal value of the worker in his current state, $v^*(i,x)$.

The marginal value of a production unit with a single worker of type k is such that

$$\begin{aligned} V_{k,0}^* = f(k,0) + \beta \mathbb{E} \left\{ \right. & \left. \sigma[\hat{V}_{0,0}^* - \hat{V}_{k+,0}^*] + \delta[\hat{V}_{0,0}^* + U_{k+}^* - \hat{V}_{k+,0}^*] \right. \\ & + \left[\sum_{i,x} q^*(i,x) \max\{\hat{V}_{k+,i}^* - \hat{V}_{k+,0}^* - v^*(i,x), 0\} \right] \\ & \left. + \left[\sum_y \lambda_e p^*(y) \max\{v^*(k+,y) - (\hat{V}_{k+,0}^* - \hat{V}_{0,0}^*), 0\} \right] + \hat{V}_{k+,0}^* \right\} \end{aligned} \quad (14)$$

The first term on the right-hand side is the output generated by the production unit in the current period. The second term is the present value of output generated by the production unit starting next period. The worker exits the market with probability

σ , in which case the continuation value of the unit is $\hat{V}_{0,0}^*$. The worker is forced into unemployment with probability δ , in which case the continuation value of the unit is $\hat{V}_{0,0}^* + U_{k_+}^*$. The worker contacts a poaching firm in state y with probability $\lambda_e p^*(y)$, where $p^*(i) = e_{i,0}^*$, $p^*(i, j) = e_{i,j}^*/2$ and $p^*(0) = n_0^*$ with $n_0^* = 1 - \sum_i e_{i,0}^* - \sum_{i,j} e_{i,j}^*/2$. In this case, the continuation value of the production unit is given by $\hat{V}_{k_+,0}^*$ plus the difference between the marginal value of the worker at the poaching firm, $v^*(k_+, y)$, and at his current firm, $\hat{V}_{k_+,0}^* - \hat{V}_{0,0}^*$. The firm contacts a worker of type i in state x with probability $q^*(i, x)$. In this case, the continuation value of the production unit is given by $\hat{V}_{k_+,0}^*$ plus the difference between the marginal value of the worker at the production unit, $\hat{V}_{k_+,i}^* - \hat{V}_{k_+,0}^*$, and the marginal value of the worker in his current state, $v^*(i, x)$.

The marginal value of a production unit with workers of type k and ℓ is such that

$$\begin{aligned}
V_{k,\ell}^* &= f(k, \ell) + \\
&\beta \mathbb{E} \left\{ [\sigma(\hat{V}_{\ell_+,0}^* + \hat{V}_{k_+,0}^*) - 2\sigma\hat{V}_{k_+,\ell_+}^*] + [\delta(\hat{V}_{k_+,0}^* + \hat{V}_{\ell_+,0}^* + U_{k_+}^* + U_{\ell_+}^*) - 2\delta\hat{V}_{k_+,\ell_+}^*] \right. \\
&+ \left. \left[\sum_{i,z} q^*(i, z) \max \left\{ \max \{ \hat{V}_{k_+,i}^* + U_{\ell_+}^*, \hat{V}_{\ell_+,i}^* + U_{k_+}^* \} - \hat{V}_{k_+,\ell_+}^* - v^*(i, z), 0 \right\} \right] \right. \\
&+ \left. \left[\sum_y \lambda_e p^*(y) \max \left\{ v^*(k_+, y) - (\hat{V}_{k_+,\ell_+}^* - \hat{V}_{\ell_+,0}^*), 0 \right\} \right] \right. \\
&+ \left. \left[\sum_y \lambda_e p^*(y) \max \left\{ v^*(\ell_+, y) - (\hat{V}_{k_+,\ell_+}^* - \hat{V}_{k_+,0}^*), 0 \right\} \right] + \hat{V}_{k_+,\ell_+}^* \right\} \quad (15)
\end{aligned}$$

The expression above is the analogue to (14) for a production unit with two rather than one worker.

The value $\hat{V}_{0,0}^*$ is obviously equal to $V_{0,0}^*$. The values $\hat{V}_{k,0}^*$ and $\hat{V}_{k,\ell}^*$ are respectively given by

$$\hat{V}_{k,0}^* = \max \{ V_{k,0}^*, V_{0,0}^* + U_k^* \}, \quad (16)$$

and

$$\hat{V}_{k,\ell}^* = \max \left\{ V_{k,\ell}^*, \hat{V}_{k,0}^* + U_\ell^*, \hat{V}_{\ell,0}^* + U_k^* \right\}. \quad (17)$$

Finally, the marginal value U_k^* of an unemployed workers of type k is such that

$$U_k = b(k) + \beta \mathbb{E}_{k_+} \left\{ U_{k_+}^* + \sigma [0 - U_{k_+}^*] + \left[\sum_y \lambda_u p^*(y) \max \left\{ v^*(k_+, y) - U_{k_+}^*, 0 \right\} \right] \right\}. \quad (18)$$

The first term on the right-hand side of (18) is the output generated by the worker in the current period. The second term is the present value of output generated by the worker starting next period. The worker meets a firm in state y with probability $\lambda_u p^*(y)$. When the worker meets a firm in state y , the present value of output generated by the worker is $U^*(k)$ plus $\max\{v^*(k_+, y) - U^*(k_+), 0\}$.

The equations for the marginal value of unemployed workers, idles firms and production units in the planner's problem are very similar to the equations for the value of

unemployment, idle firms and production units in the stationary equilibrium. The sole difference is that the equations in the planner’s problem do not depend on the bargaining powers γ and $1 - \gamma$. In fact, the marginal value of an unemployed worker to the planner includes the full value of the meetings between that worker and firms (rather than a fraction γ). The marginal value of an idle firm to the planner includes the full value of the meetings between that firm and workers (rather than a fraction $1 - \gamma$). The marginal value of a production unit to the planner includes the full value of the meetings between the firm associated with that production unit and outside workers (rather than a fraction $1 - \gamma$), and the full value of the meetings between the workers associated with that production unit and poaching firms (rather than a fraction γ).

It is easy to understand the difference between the marginal values in the planner’s problem and the private values in the equilibrium. At the margin, the value to the planner of an additional unemployed worker, idle firm, or production unit must include the full value of the meetings generated by these entities. In equilibrium, the private value of an unemployed worker, idle firm, or production unit only includes the fraction of the value of a meeting that is captured by these entities. For the private values in equilibrium to coincide with the marginal values in the planner’s problem, it would have to be the case that—in a meeting between two parties—both parties are rewarded with the entire value of the meeting. Clearly, this is not possible as it requires distributing more resources than are available. 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 with a team (i, j)) is proportional to the product between the measures of the two types (see, e.g., Kiyotaki and Lagos 2007).

The above observation suggest that the Stationary Equilibrium is not efficient. Formally, let $\{U^*, V^*\}$ and $\{u^*, e^*\}$ denote the steady-state value functions and distributions of the planner’s problem and let $\{U, V\}$ and $\{u, e\}$ denote the steady-state value functions and distributions in the Stationary Equilibrium. To see that the equilibrium is not efficient, it is sufficient to note that—even if $\{u, e\} = \{u^*, e^*\}$ —the equilibrium values $\{U, V\}$ are different from the social values $\{U^*, V^*\}$ as they solve a different system of Bellman Equations. Moreover, since $\{U, V\}$ are different from $\{U^*, V^*\}$, the equilibrium policies will be typically different¹³ from the social planner’s choices and, hence, $\{u, e\}$ will also be different from $\{u^*, e^*\}$.

We have thus established the following result.

¹³Of course, it may be possible that—even though the equilibrium values are different from the social values—their ranking is the same so that the planner’s steady-state distribution is a steady-state of the market economy. In our numerical examples, we find that this is not the case.

Theorem 1: (Inefficiency of Equilibrium) *Any Stationary Equilibrium $\{U, V\}$ and $\{u, e\}$ is inefficient. That is, $\{U, V\}$ and $\{u, e\}$ is different from the planner's steady-state $\{U^*, V^*\}$ and $\{u^*, e^*\}$.*

4.2 Optimal Taxation

We now construct a system of subsidies/taxes for which there is an efficient equilibrium, in the sense that the steady-state of the planner's problem, $\{U^*, V^*\}$ and $\{u^*, e^*\}$, satisfies the conditions for a stationary equilibrium. To this aim, let $s_u(k)$ denote the subsidy (if positive) or tax (if negative) paid by the government to an unemployed worker of type k . Let s_0 denote the subsidy paid to an idle firm. Let $s_1(k)$ denote the subsidy paid to a production unit with one worker of type k . Finally, let $s_2(k, \ell)$ denote the subsidy paid to a production unit with workers of type (k, ℓ) .

Consider the subsidy $s_u^*(k)$ for an unemployed worker of type k such that

$$U_k = b(k) + s_u^*(k) + \beta \mathbb{E}_{k_+} \left\{ U_{k_+}^* + \sigma [0 - U_{k_+}^*] + \left[\sum_y \lambda_u p^*(y) \gamma \max \left\{ v^*(k_+, y) - U_{k_+}^*, 0 \right\} \right] \right\}. \quad (19)$$

In words, $s_u^*(k)$ is a subsidy such that the equilibrium value of unemployment for a worker of type k —given that the equilibrium distribution of workers and firms is the same as in the steady state of the planner's problem and that the equilibrium continuation values are the same as the planner's marginal values—coincides with the marginal value of an unemployed worker of type k to the planner.

Next, consider the subsidy s_0^* for an idle firm such that

$$V_{0,0}^* = s_0^* + \beta \left\{ \left[\sum_{i,x} q^*(i, x) (1 - \gamma) \max \{ \hat{V}_{i,0}^* - \hat{V}_{0,0}^* - v^*(i, x), 0 \} \right] + \hat{V}_{0,0}^* \right\}. \quad (20)$$

In words, s_0^* is such that the equilibrium value of an idle firm—given that the distribution of workers and firms is the same as in the planner's steady state and that the continuation values are the same as the planner's marginal values—coincides with the marginal value of an idle firm to the planner.

Similarly, the subsidy $s_1^*(k)$ for a production unit with one worker of type k is

$$\begin{aligned} V_{k,0}^* &= f(k, 0) + s_1^*(k) \\ &+ \beta \mathbb{E} \left\{ \sigma [\hat{V}_{0,0}^* - \hat{V}_{k_+,0}^*] + \delta [\hat{V}_{0,0}^* + U_{k_+}^* - \hat{V}_{k_+,0}^*] \right. \\ &+ \left. \left[\sum_{i,x} q^*(i, x) (1 - \gamma) \max \{ \hat{V}_{k_+,i}^* - \hat{V}_{k_+,0}^* - v^*(i, x), 0 \} \right] \right\} \\ &+ \left[\sum_y \lambda_e p^*(y) \gamma \max \{ v^*(k_+, y) - (\hat{V}_{k_+,0}^* - \hat{V}_{0,0}^*), 0 \} \right] + \hat{V}_{k_+,0}^* \end{aligned} \quad (21)$$

The subsidy $s_2^*(k, \ell)$ for a production unit with two workers of type (k, ℓ) is

$$\begin{aligned}
V_{k,\ell}^* &= f(k, \ell) + s_2^*(k, \ell) \\
&+ \beta \mathbb{E} \left\{ [\sigma(\hat{V}_{\ell_+,0}^* + \hat{V}_{k_+,0}^*) - 2\sigma\hat{V}_{k_+,\ell_+}^*] + [\delta(\hat{V}_{k_+,0}^* + \hat{V}_{\ell_+,0}^* + U_{k_+}^* + U_{\ell_+}^*) - 2\delta\hat{V}_{k_+,\ell_+}^*] \right\} \\
&+ \left[\sum_{i,z} q^*(i, z)(1 - \gamma) \max \left\{ \max\{\hat{V}_{k_+,i}^* + U_{\ell_+}^*, \hat{V}_{\ell_+,i}^* + U_{k_+}^*\} - \hat{V}_{k_+,\ell_+}^* - v^*(i, z), 0 \right\} \right] \\
&+ \left[\sum_y \lambda_e p^*(y) \gamma \max \left\{ v^*(k_+, y) - (\hat{V}_{k_+,\ell_+}^* - \hat{V}_{\ell_+,0}^*), 0 \right\} \right] \\
&+ \left[\sum_y \lambda_e p^*(y) \gamma \max \left\{ v^*(\ell_+, y) - (\hat{V}_{k_+,\ell_+}^* - \hat{V}_{k_+,0}^*), 0 \right\} \right] + \hat{V}_{k_+,\ell_+}^* \}
\end{aligned} \tag{22}$$

The subsidies $s_1^*(k)$ and $s_2^*(k, \ell)$ are constructed so that the equilibrium value of a production unit coincides with its marginal value to the planner, given that the distribution of workers and firms is the same as in the planner's steady state and that the continuation values are the same as the planner's marginal values.

Given the system of subsidies $s^* \equiv \{s_u^*, s_0^*, s_1^*, s_2^*\}$, the planner's values $\{U^*, V^*\}$ and the planner's steady-state distribution $\{u^*, e^*\}$ constitute a stationary equilibrium. By construction of s^* , the planner's values and the planner's distribution satisfy all of the Bellman Equations for a stationary equilibrium. Moreover, since the planner's distribution $\{u^*, e^*\}$ is stationary and the law of motion for the distribution is the same in the planner's problem as in equilibrium, $\{u^*, e^*\}$ also satisfies all of the inflow-outflow equations for a stationary equilibrium. Notice that the optimal system of subsidies s^* can be recovered using equations (19)-(22) and computed using only the planner's solution, $\{U^*, V^*\}$ and $\{u^*, e^*\}$, and the bargaining powers, γ and $1 - \gamma$.

Nothing guarantees that the system of subsidies s^* balances the budget of the government. To address this issue, let T^* be given by

$$T^* = \{s_0^* n_0^* + \sum_k [s_u^*(k) u_p^*(k) + s_1^*(k) e_p^*(k, 0) + \sum_\ell s_2^*(k, \ell) e_p^*(k, \ell) / 2]\} / 2. \tag{23}$$

Given T^* , let the system of subsidies/taxes $\hat{s} \equiv \{\hat{s}_u, \hat{s}_0, \hat{s}_1, \hat{s}_2\}$ be defined as

$$\hat{s}_u = s_u^* - T^*, \hat{s}_0 = s_0^* - T^*, \hat{s}_1 = s_1^* - 2T^*, \hat{s}_2 = s_2^* - 3T^*. \tag{24}$$

In words, T^* is equal to the total subsidies paid out by the government to unemployed workers, idle firms, and production units divided by the total measure of workers and firms. Then, \hat{s}_u is equal to the optimal subsidy s_u^* for an unemployed worker net of the average subsidy T^* . Similarly, \hat{s}_0 is equal to the optimal subsidy s_0^* for an idle firm net of T^* . Finally, \hat{s}_1 and \hat{s}_2 are equal to the optimal subsidies s_1^* and s_2^* for production units net of the average subsidy T^* multiplied by the number of agents attached to the unit. Effectively, \hat{s} is equal to s^* net of a per-capita constant T^* . Since \hat{s} is the same as the

optimal system s^* net of a per-capita constant, \hat{s} induces the same hiring choices as s^* and, hence, leads to the efficient allocation of workers and firms. Since T^* is equal to the average subsidy given the system s^* , \hat{s} balances the budget of the government.

We have thus established the following result.

Theorem 2: (Optimal Subsidies and Taxes).

(i) *Given the system of subsidies/taxes s^* in (19)-(22), the steady-state of the planner's problem, $\{U^*, V^*\}$ and $\{u^*, e^*\}$, is a Stationary Equilibrium.*

(ii) *Given the system of subsidies/taxes \hat{s} in (24), $\{U, V\}$ and $\{u^*, e^*\}$ is a Stationary Equilibrium, where $U = U^* - T^*/(1 - \beta(1 - \sigma))$, $V_0 = V_0^* - T^*/(1 - \beta)$, $V_1 = V_1^* - T^*/(1 - \beta) - T^*(1 - \beta(1 - \sigma))$ and $V_2 = V_2^* - T^*/(1 - \beta) - 2T^*/(1 - \beta(1 - \sigma))$. Moreover, the system of subsidies/taxes \hat{s} balances the government budget.*

4.3 Measuring Inefficiencies and Optimal Taxes

We now compare the equilibrium to the solution of the planner's problem given the calibrated value of the parameters. Figure 9 contains histograms for the distribution of human capital among workers entering the labor market, the efficient distribution of human capital in the population, and the equilibrium distribution of human capital in the population. The distribution of human capital is lower in equilibrium than in the planner's solution. The average human capital is 2.5% lower in equilibrium than in the planner's solution. The difference between the average human capital in the population and at entry is about 7% lower in equilibrium than in the planner's solution. Similarly, output is 2% lower in equilibrium.

To understand why human capital is lower in equilibrium, it is useful to compare the equilibrium policy functions in Figure 2 with the planner's policy functions in Figure 10. There are three qualitative differences. First, in equilibrium, high human capital workers who are unemployed are too often hired as replacements of the worst rather than the best of the firm's employees. Second, medium and high human capital workers who are employed on their own are too often hired as replacements of the worst of the firm's employees rather than being left on their own. Third, medium and high human capital workers who are paired with good teammates do not move often enough to firms with worse employees. All three differences between equilibrium and planner's policies have the effect of reducing the time that low human capital workers spend together with more knowledgeable coworkers and, given the asymmetry of the learning function, lowering the overall accumulation of human capital. Intuitively, the value of production units made up of two workers with different human capital is lower in equilibrium because the value

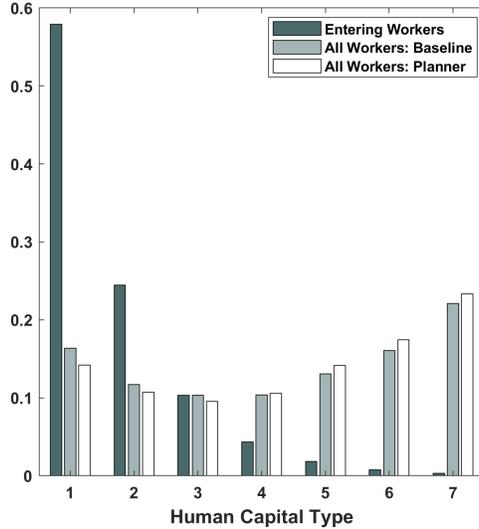


Figure 9: Efficient Human Capital Distribution

of the knowledge transmitted by the high to the low human capital worker is eventually captured by poaching firms.

Figure 11(a) illustrates the distribution of coworkers in the planner’s solution.¹⁴ Figure 11(b) compares the distribution of coworkers net of random assignment in the equilibrium and in the planner’s solution. The figure is clear. In equilibrium, there is an inefficiently high density along the 45°-line and an inefficiently low density away from the 45°-line. This means that in equilibrium there is an inefficiently large measure of teams composed of similar workers and an inefficiently small measure of teams composed of workers with different human capital. The correlation of the human capital of coworkers is 10% in equilibrium and -2% in the planner’s problem.

Figure 12 shows the system of subsidies that makes the equilibrium efficient and balances the budget of the government. The left panel shows the subsidies for unemployed workers of type k and for production unit with a single worker of type k . These subsidies are large (relative to the flow output) and they are increasing in the worker’s human capital. The right panel shows the subsidies for production units with workers of type k and ℓ . These subsidies are negative (they are taxes), small (relative to output), and their absolute value tend to increase with the average human capital of the coworkers. The key property of the system is that the subsidies for unemployed workers and for workers employed on their own are increasing in the worker’s human capital. Such subsidies dissuade firms from upgrading their workforce, which is the main source of inefficiency

¹⁴We refer the reader to Appendix D for a 3D illustration of the efficient distribution of coworkers.

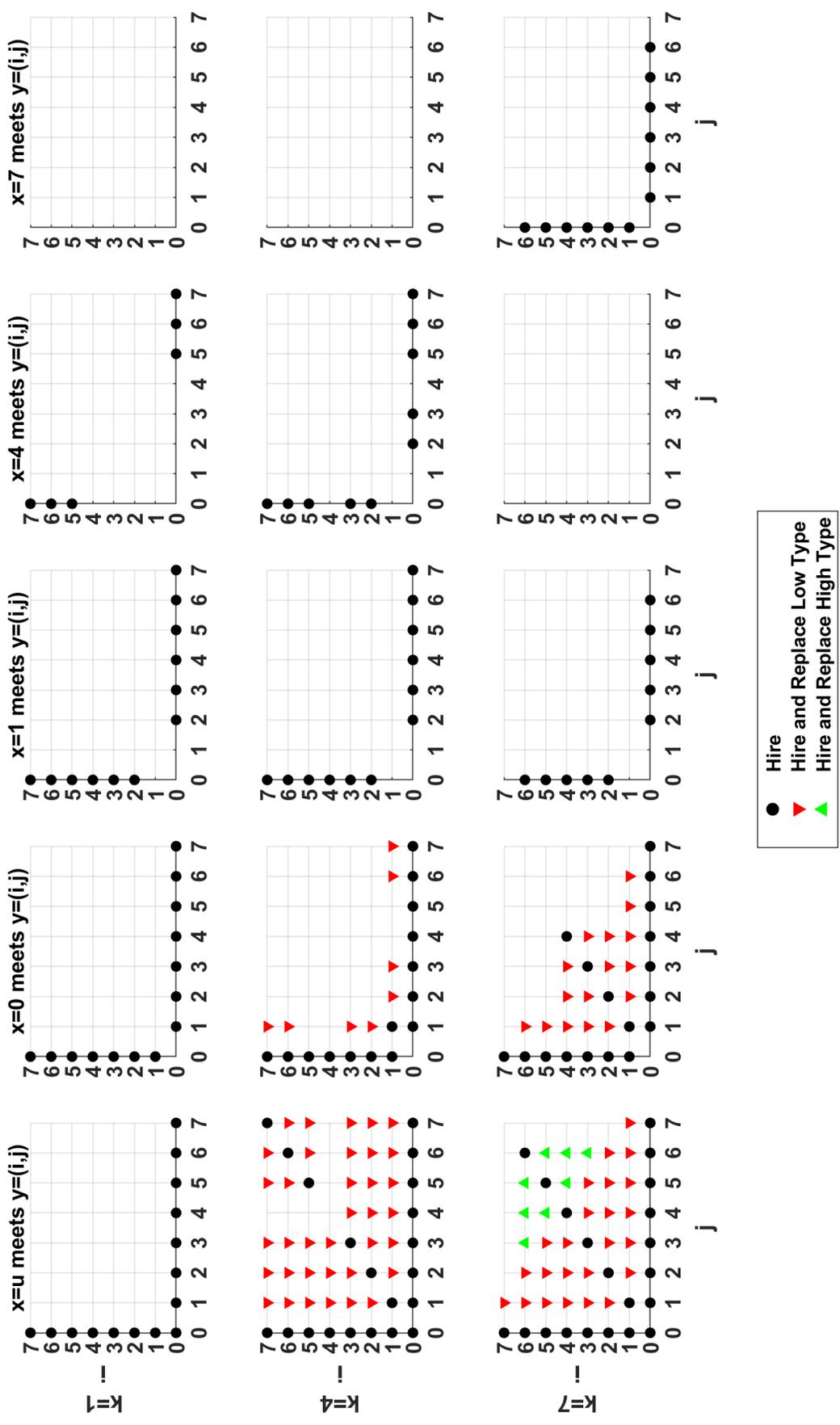
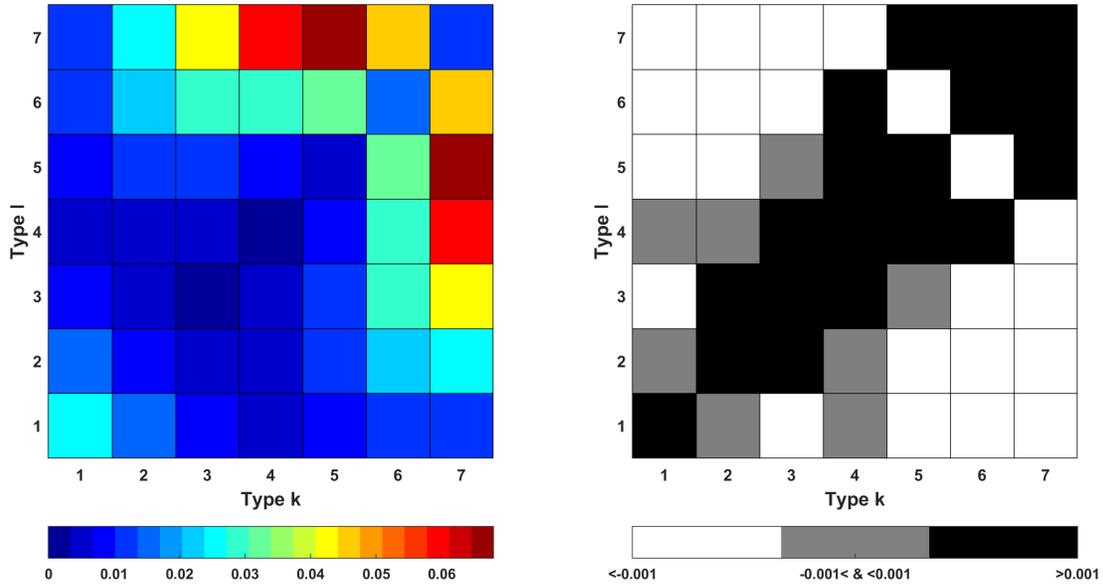


Figure 10: Efficient Policy Functions



(a) Distribution

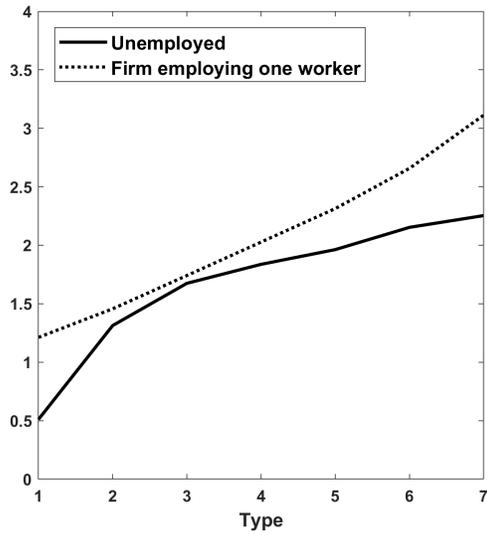
(b) Net distribution in Baseline minus Plan

Figure 11: Efficient Pattern of Sorting

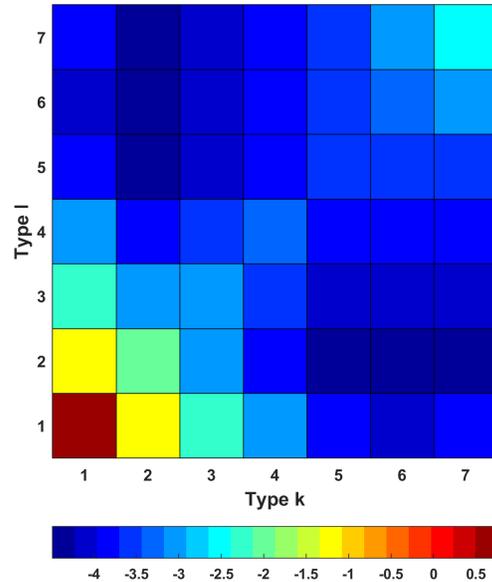
of equilibrium. Indeed, if a firm replaces a low human capital employee with a more knowledgeable new hire, it has to compensate the new hire for the loss of a large subsidy while the replaced employee who moves into unemployment receives a smaller subsidy.

The system of subsidies in Figure 12 is designed to make sure that, in any meeting between a firm and a worker, the private gains from trade coincide *in level* with the planner's gains from trade. Given this design, the system of subsidies guarantees that the socially efficient distribution of workers and firms is a stationary equilibrium of the market economy. However, the system is also highly invasive. A less invasive system can be obtained by finding hiring subsidies/taxes to make sure that the private gains from trade coincide *in sign* with the planner's gains from trade.

Figure 13 reports the optimal hiring subsidies and taxes conditional on the worker's state—i.e. his human capital and his employment status—and on the firm's state—i.e., idle, operating a production unit with a single employee or a production unit with two employees. The three rows in Figure 13 report the hiring subsidies and taxes for a worker of type 1, 4 and 7. The five columns report the hiring taxes and subsidies for a worker who is unemployed, employed on his own, or employed with a coworker of type 1, 4 or 7. Each cell defined by the intersection of a row and a column, reports the hiring taxes and subsidies if the firm in state (\bar{i}, \bar{j}) . A green dot means that the hiring is subsidized.



(a) Subsidies



(b) Subsidies for teams

Figure 12: Optimal Subsidies and Taxes

A green triangle pointing up (*down*) means that the hiring is subsidized to replace the firm's best (*worst*) employee. A red dot means that the hiring is taxed so that it does not take place. A red triangle pointing up (*down*) means that the hiring is taxed in a way to induce the firm to replace the firm's best (*worst*) employee.

Glancing at Figure 13, one immediately notices that the system does not prescribe any subsidies or taxes for firms considering hiring workers of type 1. The private gains from trade between a firm and a worker of type 1 always have the same sign as the planner's gains from trade and, thus, the equilibrium hiring decisions are efficient even without hiring subsidies or taxes. The system prescribes some subsidies and taxes for firms considering hiring workers of type 4. For instance, a firm with an employee of type 1 and another employee of type 4 or 5 needs to be dissuaded from hiring a worker of type 4 currently employed on his own. A firm with a single employee of type 2 or 3 needs to be given a subsidy to hire a worker of type 4 currently employed with another worker of type 4. The taxes and subsidies for firms considering hiring a worker of type 7 are more prevalent. For instance, firms with employees of type 1 and 7 need to be dissuaded by a tax from hiring an unemployed worker of type 7. Firms with an employee of type (2, 6) need to be faced with a tax in order to be dissuaded from replacing the best employee and, instead, replace the best employee. Firms with a single employee of type 1 or 2 need

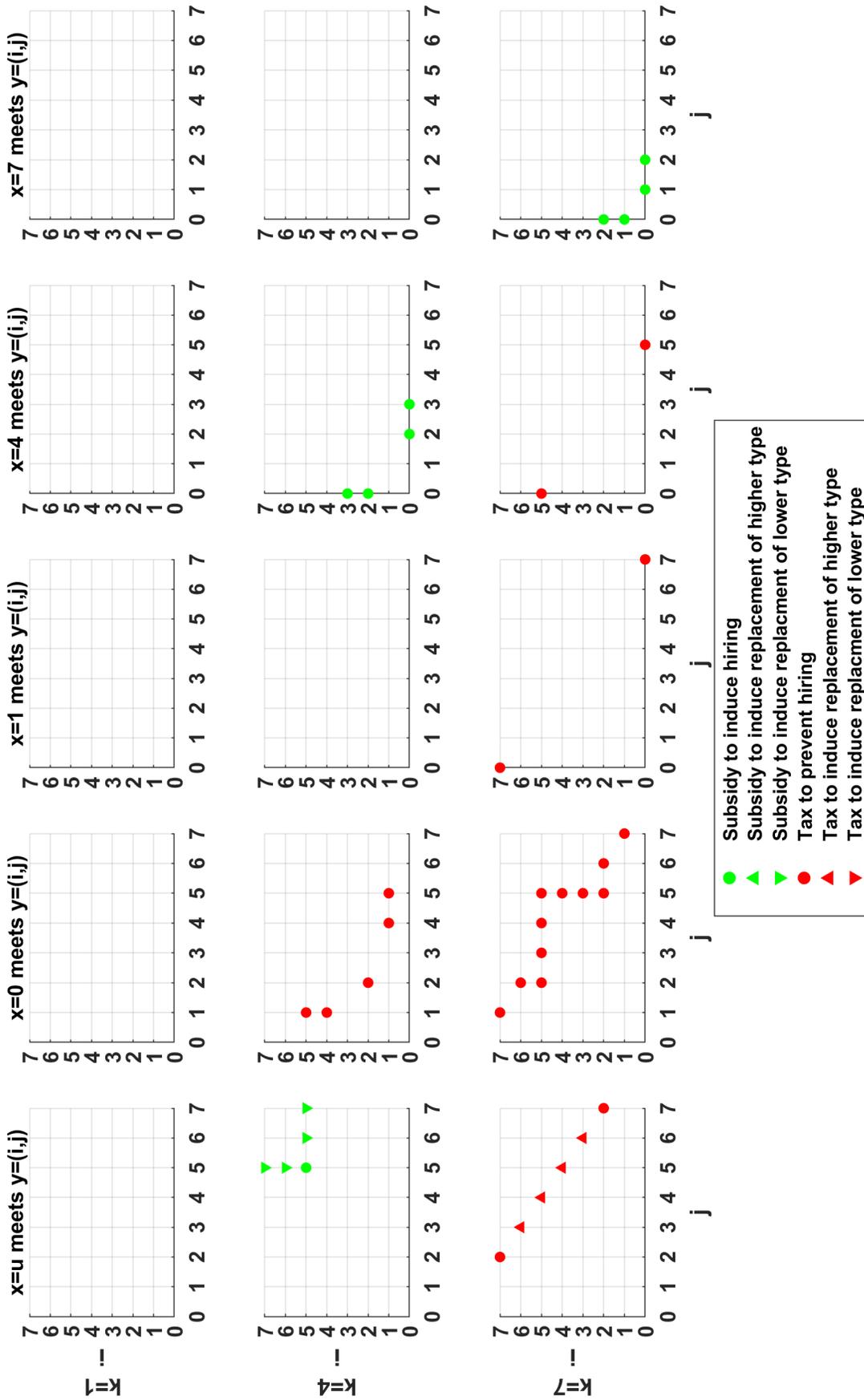


Figure 13: Hiring Subsidies and Taxes

a subsidy to poach a worker of type 7 currently employed with another high productivity individual. Overall, hiring taxes and subsidies tend to reduce the positive sorting in the equilibrium.

5 Conclusions

We developed a version of the search-theoretic equilibrium model of the labor market of Postel-Vinay and Robin (2002) in which production involves teams of workers and in which a worker's human capital accumulation depends on the human capital of his coworkers. The model implies a theory of how workers with different human capital sort themselves into different teams, and a theory of how a worker's wage is affected by his human capital, his coworkers' human capital and the extent of competition in a frictional labor market. The theory allows us to infer the properties of the production function and of the human capital accumulation function using a matched employer-employee dataset of US workers and firms. We find that the production function is supermodular. We find that the human capital function is asymmetric. Namely, if a worker is less knowledgeable than his coworkers, his human capital grows faster the better are his coworkers. If a worker is more knowledgeable than his coworkers, his human capital grows at the same rate irrespective of the quality of his coworkers. Having developed and estimated an equilibrium theory of the labor market, we can run proper counterfactuals. We find that learning from coworkers accounts for 60% of all of the human capital that is accumulated by workers on the job. We find that the supermodularity of the production function has a negative effect on the stock of human capital and on output by making the pattern of sorting more positive. Finally, we study the efficient assignment of workers to production teams, taking into consideration both the supermodularity of production and the asymmetry of learning. We find that the efficient sorting of coworkers across production teams is less positive than in the data, allowing for more learning opportunities for low human capital workers and, consequently, for a higher aggregate stock of human capital and higher output.

Much work remains to be done. From the perspective of theory, it would be important to extend the model to allow firms to operate production teams with more than two workers. The difficulty in carrying out this extension is the growth in the dimensionality of the state space of the firm. The reward from this extension is a more direct mapping between the notion of coworkers in the model and in the data. We experimented with this extension using a directed search model in the spirit of Menzies and Shi (2011) and Schaal (2017), but we still needed to keep the number of human capital level low. Second, it would be interesting to extend the model to allow for ex-ante differences in the productivity of

firms. Since our model generates, because of search frictions, ex-post differences in the value of a worker to different firms we did not need to pursue this extension to match data on EE transitions. Third, it would also be interesting to find out whether a simple system of taxes and subsidies—in particular a system that does not require knowledge of the human capital of different workers—can bring the economy close to the efficient allocation. From the perspective of measurement, it would be interesting to dig deeper into the properties of the human capital accumulation function. For instance, we would want to know whether the effect of coworkers in an individual’s human capital accumulation function is affected by the distance between worker and coworkers in terms of race, gender, age and education.

References

- [1] Anderson, A. 2015. “A Dynamic Generalization of Becker’s Assortative Matching Result.” *Journal of Economic Theory*, 159: 290-310.
- [2] Anderson, A., and L. Smith. 2010. “Dynamic Matching and Evolving Reputations.” *Review of Economic Studies*, 77: 3-29.
- [3] Anderson, A., and L. Smith. 2018. “Robust Comparative Statics in Matching Models.” Manuscript, Georgetown University.
- [4] Bagger, J., F. Fontaine, F. Postel-Vinay, and J. Robin. 2014. “Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics.” *American Economic Review*, 104: 1551-1596.
- [5] Bagger, J., and R. Lentz. 2018. “An Empirical Model of Wage Dispersion with Sorting.” *Review of Economic Studies*, Forthcoming.
- [6] Becker, G. 1962. “Investment in Human Capital: A Theoretical Analysis.” *Journal of Political Economy*, 70: 9-49.
- [7] Benhabib, J., J. Perla, and C. Tonetti. 2017. “Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier.” Manuscript, New York University.
- [8] Ben-Porath, Y. 1967. “The Production of Human Capital and the Life Cycle of Earnings.” *Journal of Political Economy*, 75: 352-365.

- [9] Booij, A., E. Leuven, and H. Oosterbeek. 2017. “Ability Peer Effects in University: Evidence from a Randomized Experiment.” *Review of Economic Studies*, 84: 547-578.
- [10] Burdett, K., Carillo-Tudela, C., and M. Coles. 2018 “The Cost of Job Loss.” Manuscript, University of Essex.
- [11] Cahuc, P. F. Postel-Vinay and J. Robin. 2006. “Wage Bargaining with On-the-Job Search: Theory and Evidence.” *Econometrica*, 74: 323-364.
- [12] Carrell, S., B. Sacerdote, and J. West. 2013. “From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation.” *Econometrica*, 81: 855-882.
- [13] Cornelissen, T., C. Dustmann, and U. Schoenberg. 2018. “Peer Effects in the Workplace.” *Review of Economic Studies*, Forthcoming.
- [14] Eeckhout, J., and P. Kircher. 2011. “Identifying Sorting-In Theory.” *Review of Economic Studies*, 78: 872-906.
- [15] Gautier, P., and C. Teulings. 2004. “The Right Man for the Job.” *Review of Economic Studies*, 71: 553-580.
- [16] Gulyas, A. 2016. “Identifying labor market sorting with firm dynamics.” Manuscript, University of Mannheim.
- [17] Hagedorn, M., T. Law, and I. Manovskii. 2018. “Identifying Sorting.” *Econometrica*, Forthcoming.
- [18] Hoxby, C. 2000. “Peer Effects in the Classroom: Learning from Gender and Race Variation.” NBER Working Paper 7867.
- [19] Jarosch, G., E. Oberfield, and E. Rossi-Hansberg. 2018. “Learning from Coworkers.” Manuscript, Princeton University.
- [20] Jovanovic, B. 2014. “Misallocation and Growth.” *American Economic Review*, 104: 1149-1171.
- [21] Kiyotaki, N., and R. Lagos. 2007. “A Model of Job and Worker Flows.” *Journal of Political Economy*, 115: 770-815.
- [22] Lise, J., C. Meghir, and J. Robin. 2016. “Matching, Sorting and Wages.” *Review of Economic Dynamics*, 19: 63-87.

- [23] Lise, J., and F. Postel-Vinay. 2016. "Multidimensional Skills, Sorting, and Human Capital Accumulation." Manuscript, University of Minnesota.
- [24] Lopes de Melo, R. 2018. "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence." *Journal of Political Economy*, Forthcoming.
- [25] Lucas, R. 2009. "Ideas and Growth." *Economica*, 76: 1-19.
- [26] Lucas, R., and B. Moll. 2014. "Knowledge Growth and the Allocation of Time." *Journal of Political Economy*, 122: 1-51.
- [27] Mas, A., and E. Moretti. 2009. "Peers at Work." *American Economic Review*, 99: 112-145.
- [28] Menzio, G., and S. Shi. 2011. "Efficient Search on the Job and the Business Cycle." *Journal of Political Economy*, 119: 468-510.
- [29] Menzio, G., I. Telyukova, and L. Visschers. 2016. "Directed Search over the Life Cycle." *Review of Economic Dynamics*, 19: 38-62.
- [30] Perla, J., and C. Tonetti. 2014. "Equilibrium Imitation and Growth." *Journal of Political Economy*, 122: 52-76.
- [31] Postel Vinay, F., and J. Robin. 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica*, 70: 2295-2350.
- [32] Schaal, E. 2017. "Uncertainty and Unemployment." *Econometrica*, 85: 1675-1721.
- [33] Shimer, R., and L. Smith. 2000. "Assortative Matching and Search." *Econometrica*, 68: 343-369.

Appendix

A Equilibrium Laws of Motion

Let u and e denote the distribution of workers across employment states at the beginning of the current search stage. In order to formally express how this distribution evolves over time, we need to introduce some notation to describe the equilibrium policy functions. Specifically, let $h_y(k, x)$ denote the probability that a firm in state $y \in Y$ hires a worker of type $k \in K$ in state $x \in X$. The probability $h_y(k, x)$ is 1 if the marginal value of the worker to the firm, $v(k, y)$, is greater than the value of the worker in his current employment state, $v(k, x)$. Otherwise, $h_y(k, x)$ is zero. Also, let $r_{i,j}(k, i)$ denote the probability that a firm in state $(i, j) \in K \times K$ replaces the employee i with a new hire of type k . The probability $r_{i,j}(k, i)$ is equal to 1 if $\hat{V}_{k,j} + U_i > \hat{V}_{k,i} + U_j$, it is equal to 0 if $\hat{V}_{k,j} + U_i < \hat{V}_{k,i} + U_j$, and it is equal to 1/2 if $\hat{V}_{k,j} + U_i = \hat{V}_{k,i} + U_j$.

Let u_p and e_p denote the distribution of workers across employment states at the beginning of the production stage. At this stage, the measure $u_p(k)$ of unemployed workers is given by

$$\begin{aligned}
 & u_p(k) \\
 = & u(k) \left\{ (1 - \lambda_u) + \left[\sum_y \lambda_u p(y) (1 - h_y(k, u)) \right] \right\} \\
 + & e(k, 0) \delta + \sum_\ell e(k, \ell) \left\{ \delta + \left[\sum_{i,x} q(i, x) h_{k,\ell}(i, x) r_{k,\ell}(x, k) \right] \right\}
 \end{aligned} \tag{25}$$

The first term on the right-hand side is the measure of unemployed workers at the beginning of the search stage who either did not contact a firm or contacted a firm but were not hired. The second term is the measure of workers who were employed on their own at the beginning of the search stage and lost their job for exogenous reasons. The last term is the measure of workers who were employed with a coworker of type ℓ at the beginning of the search stage and either lost the job for exogenous reasons or were replaced by a new hire.

At the beginning of the production stage, the measure $e_p(k, 0)$ of workers employed on their own is given by

$$\begin{aligned}
& e_p(k, 0) \\
&= u(k) \{ \lambda_u p(0, 0) h_{0,0}(k, u) \} + e(k, 0) \left\{ \left[\sum_y \lambda_e p(y) (1 - h_y(k, 0)) \right] \right\} \\
&+ e(k, 0) \left\{ \left[\sum_{i,x} q(i, x) (1 - h_{k,0}(i, x)) \right] + \left[1 - \delta - \lambda_e - \sum_{i,x} q(i, x) \right] \right\} \\
&+ \sum_\ell e(k, \ell) \left\{ \lambda_u p(0, 0) h_{0,0}(k, \ell) + \left[\sum_y p(y) h_y(\ell, k) \right] + \delta \right\}.
\end{aligned} \tag{26}$$

The first term on the right-hand side is the measure of workers who were unemployed at the beginning of the search stage and who were hired by a firm in state $(0, 0)$. The second, third and fourth terms are the measure of workers who were employed on their own at the beginning of the search stage and remained in the same employment position. The last term is the measure of workers who were employed with a coworker of type ℓ at the beginning of the search stage and who were either hired by a firm in state $(0, 0)$ or who lost their coworker to a poaching firm or to unemployment.

At the beginning of the production stage, the measure $e_p(k, \ell)$ of workers employed with a coworker of type ℓ is given by

$$\begin{aligned}
& e_p(k, \ell) \\
&= u(k) \{ \lambda_u p(\ell, 0) h_{\ell,0}(k, u) + \lambda_u [\sum_s p(\ell, s) h_{\ell,s}(k, u) r_{\ell,s}(k, s) + p(s, \ell) h_{s,\ell}(k, u) r_{s,\ell}(k, s)] \} \\
&+ e(k, 0) \{ \lambda_e [\sum_s p(\ell, s) h_{\ell,s}(k, 0) r_{\ell,s}(k, s) + p(s, \ell) h_{s,\ell}(k, 0) r_{s,\ell}(k, s)] \} \\
&+ e(k, 0) \{ \lambda_e p(\ell, 0) h_{\ell,0}(k, 0) + [\sum_x q(\ell, x) h_{k,0}(\ell, x)] \} \\
&+ e(k, \ell) \left\{ \lambda_e \left[\sum_y p(y) (1 - h_y(k, \ell)) \right] + \lambda_e \left[\sum_y p(y) (1 - h_y(\ell, k)) \right] \right\} \\
&+ e(k, \ell) \left\{ \left[\sum_{i,x} q(i, x) h_{k,\ell}(i, x) \right] + \left[1 - 2\delta - 2\lambda_e - \sum_{i,x} q(i, x) \right] \right\} \\
&+ \sum_s e(k, s) \{ \lambda_e p(\ell, 0) h_{\ell,0}(k, s) + [\sum_t \lambda_e p(\ell, t) h_{\ell,t}(k, s) r_{\ell,t}(k, t)] \} \\
&+ \sum_s e(k, s) \{ \sum_x q(\ell, x) h_{k,s}(\ell, x) r_{k,s}(s, \ell) \}.
\end{aligned} \tag{27}$$

The right-hand side includes all the ways in which a worker finds himself employed with a coworker of type ℓ at the production stage. First, the worker could have been unemployed and been hired by a firm in state $(\ell, 0)$ or hired by a firm in state (ℓ, s) as a replacement for s . Second, the worker could have been employed by himself and been hired by a firm in state $(\ell, 0)$, by a firm in state (ℓ, s) as a replacement for s , or his employer could have hired a worker of type ℓ . Third, the worker could have been employed with a coworker of type ℓ and the team had survived. Fourth, the worker could have been employed with a coworker of type s and ended up, because he moved or the firm replaced his coworker, with a coworker of type ℓ .

Let u_d and e_d denote the distribution of workers across employment states at the end of the period (i.e., after the dismissal stage). The measures $u_d(k)$, $e_d(k, 0)$ and $e_d(k, \ell)$ are given by

$$\begin{aligned} u_d(k) &= u_p(k) + e_p(k, 0)d_{k,0}(k) + \sum_{\ell} e_p(k, \ell)d_{k,\ell}(k), \\ e_d(k, 0) &= e_p(k, 0)(1 - d_{k,0}(k)) + \sum_{\ell} e_p(k, \ell)d_{k,\ell}(\ell), \\ e_d(k, \ell) &= e_p(k, \ell)(1 - d_{k,\ell}(k))(1 - d_{k,\ell}(\ell)), \end{aligned} \quad (28)$$

where $d_y(k)$ denotes the probability that a firm in state y fires an employee of type k . The above expressions are easy to understand.

Let u_g and e_g denote the distribution of workers across employment states at the end of the learning stage of next period. The measures $u_g(k)$, $e_g(k, 0)$ and $e_g(k, \ell)$ are given by

$$\begin{aligned} u_g(k) &= \sum_s g(k|s, u)u_d(s), \\ e_g(k, 0) &= \sum_s g(k|s, 0)e_d(s, 0), \\ e_g(k, \ell) &= \sum_{s,t} g(k|s, t)g(\ell|t, s)e_d(s, t). \end{aligned} \quad (29)$$

In words, the measure of unemployed workers of type k at this stage is equal to the sum of the measures of unemployed workers of type s at the previous stage whose human capital type goes from s to k . The measure of workers of type k employed on their own is the sum of the measures of workers of type s employed on their own whose human capital type goes from s to k . The measure of workers of type k employed with a coworker of type ℓ is the sum of the measures of workers of type s employed with a coworker of type t who transition from s to k and from t to ℓ .

Finally, let u_+ and e_+ denote the distribution of workers across employment states at the end of the entry and exit stage. The measures $u_+(k)$, $e_+(k, 0)$ and $e_+(k, \ell)$ are given by

$$\begin{aligned} u_+(k) &= (1 - \sigma)u_g(k) + \sigma\pi(k), \\ e_+(k, 0) &= (1 - \sigma)e_g(k, 0) + \sum_{\ell} \sigma e_g(k, \ell), \\ e_+(k, \ell) &= (1 - 2\sigma)e_g(k, \ell). \end{aligned} \quad (30)$$

The distribution u_+ , e_+ is also the distribution of workers across employment states at the beginning of the search stage of next period. Thus, the distribution is stationary if and only if the following inflow-outflow conditions hold

$$\begin{aligned} u_+(k) - u(k) &= 0, \\ e_+(k, 0) - e(k, 0) &= 0, \\ e_+(k, \ell) - e(k, \ell) &= 0. \end{aligned} \quad (31)$$

B Wage Equations

Let $\hat{W}_{k,0}(w)$ denote the value for a worker of type k who is employed at the wage w and does not have a coworker. Similarly, let $\hat{W}_{k,\ell}(w)$ denote the value for a worker of type k who is employed at the wage w and has a coworker of type ℓ . The value functions $\hat{W}_{k,0}(w)$ and $\hat{W}_{k,\ell}(w)$ are measured at the beginning of the dismissal stage. The value functions $W_{k,0}(w)$ and $W_{k,\ell}(w)$ are measured at the beginning of the production stage.

The value functions $\hat{W}_{k,0}(w)$ and $\hat{W}_{k,\ell}(w)$ are given by

$$\hat{W}_{k,0}(w) = \max\{U_k, \min\{W_{k,0}(w), \hat{V}_{k,0} - \hat{V}_{0,0}\}\}, \quad (32)$$

and

$$\hat{W}_{k,\ell}(w) = \max\{U_k, \min\{W_{k,\ell}(w), \hat{V}_{k,\ell} - \hat{V}_{\ell,0}\}\}. \quad (33)$$

In words, $\hat{W}_{k,\cdot}(w)$ is the maximum between the value of unemployment to a worker of type k and the lowest between the value to the worker of staying with his current employer at the wage w and the marginal value of the worker to the production unit operated by the firm.

The value function $W_{k,0}(w)$ is given by

$$\begin{aligned} W_{k,0}(w) = & w + \beta \mathbb{E} \left\{ \hat{W}_{k_+,0}(w) \right. \\ & + \sigma \left[0 - \hat{W}_{k_+,0}(w) \right] + \delta \left[U_{k_+} - \hat{W}_{k_+,0}(w) \right] \\ & + \sum_{i,x} q(i,x) \left[h_{k_+,0}(i,x) \left[\hat{W}_{k_+,i}(w) - \hat{W}_{k_+,0}(w) \right] \right] \\ & \left. + \sum_y \lambda_e p(y) \left[\max\{\hat{W}_{k_+,0}(w), \min\{v(k_+,y), \gamma v(k_+,y) + (1-\gamma)v(k_+,x)\}\} - \hat{W}_{k_+,0}(w) \right] \right\} \end{aligned} \quad (34)$$

In the current period, the worker is paid the wage w . At the exit stage of next period, the worker leaves the labor market with probability σ . In this case, the worker's continuation value is 0. At the search stage, the worker moves into unemployment with probability δ . In this case, the worker's continuation value is U_{k_+} . The firm contacts a worker of type i in state x with probability $q(i,x)$ and hires him with probability $h_{k_+,0}(i,x)$. In this case, the worker's continuation value is $\hat{W}_{k_+,i}(w)$. The worker contacts a poaching firm in state y with probability $\lambda_e p(y)$. In this case, three different situations may arise depending on the value of the worker to the poaching firm, $v(k_+,y)$, the value of the worker to the incumbent firm, $v(k_+,x) = \hat{V}_{k_+,0} - \hat{V}_{0,0}$, and the value to the worker of being employed at the incumbent firm at the wage w , $\hat{W}_{k_+,0}(w)$. If $v(k_+,y) \leq \hat{W}_{k_+,0}(w)$, the worker remains with the incumbent at the wage w . Hence, his continuation value is $\hat{W}_{k_+,0}(w)$.

If $v(k_+, y) > \hat{W}_{k_+,0}(w)$ and $v(k_+, y) \leq v(k_+, x)$, the worker remains with the incumbent but his wage is increased to match his outside option. Hence, his continuation value is $v(k_+, y)$. If $v(k_+, y) > v(k_+, x)$, the worker moves to the poacher and captures a fraction γ of the gains from trade. Hence, his continuation value is $v(k_+, x) + \gamma[v(k_+, y) - v(k_+, x)]$.

Similarly, the value function $W_{k,\ell}(w)$ is given by

$$\begin{aligned}
W_{k,\ell}(w) &= w + \beta \mathbb{E} \left\{ \hat{W}_{k_+, \ell_+}(w) \right. \\
&+ \sigma \left[0 - \hat{W}_{k_+, \ell_+}(w) \right] + \delta \left[U_{k_+} - \hat{W}_{k_+, \ell_+}(w) \right] + \\
&+ \left(\sigma + \delta + \sum_y \lambda_e p(y) h_y(\ell_+, k_+) \right) \left[\hat{W}_{k_+, 0}(w) - \hat{W}_{k_+, \ell_+}(w) \right] \\
&+ \sum_{i,x} q(i, x) \left[h_{k_+, \ell_+}(i, x) \left[r_{k_+, \ell_+}(x, k_+) U_{k_+} + r_{k_+, \ell_+}(x, \ell_+) \hat{W}_{k_+, i}(w) - \hat{W}_{k_+, \ell_+}(w) \right] \right] \\
&+ \left. \sum_y \lambda_e p(y) \left[\max\{ \hat{W}_{k_+, \ell_+}(w), \min\{ v(k_+, y), \gamma v(k_+, y) + (1 - \gamma)v(k_+, x) \} \} - \hat{W}_{k_+, \ell_+}(w) \right] \right\}
\end{aligned} \tag{35}$$

The above expression is similar to the previous one. Therefore, we shall only point out the differences between (35) and (34). First, the coworker may exit the labor market or move into unemployment. In this case, the worker's continuation value is $\hat{W}_{k_+,0}(w)$. Second, the coworker may be hired by a poaching firm. In this case, the worker's continuation value is also $\hat{W}_{k_+,0}(w)$. Third, when a firm hires a new employee, the continuation value of the worker depends on whom the firm replaces. If the firm replaces the coworker, the worker's continuation value is $\hat{W}_{k_+,i}(w)$. If the firm replaces the worker, his continuation value is U_{k_+} .

C Planner's Problem

The problem of the social planner is

$$\begin{aligned}
&S(u_p, e_p) \\
&= \max_{h,r,d} \sum_k [u_p(k)b(k) + e_p(k, 0)f(k, 0) + \sum_\ell e_p(k, \ell)f(k, \ell)/2] + \beta S(u_{p+}, e_{p+}),
\end{aligned} \tag{36}$$

subject to the law of motions (25)-(30). The planner's choice variables are: $h_y(k, x)$ which is the probability that a firm of type $y \in Y$ hires a worker of type $k \in K$ in employment state $x \in X$; $r_y(k, i)$ which is the probability that a firm of type $y = (i, j) \in K \times K$ replaces the employee i with a worker of type k ; $d_y(i)$ which is the probability that a firm of type $y = (i, j)$ fires an employee of type i .

Let U_k^* denote the derivative of S with respect to $u_p(k)$, evaluated at the steady state $\{u_p^*, e_p^*\}$. Let $V_{k,0}^-$ denote the derivative of S with respect to $e_p(k, 0)$, evaluated at

the steady state. Let $V_{k,\ell}^-/2$ denote the derivative of S with respect to $e_p(k, \ell)$. Since $e_p(k, \ell) = e_p(\ell, k)$, $V_{k,\ell}^-/2 + V_{\ell,k}^-/2 = V_{k,\ell}^-$, which is the marginal value to the planner of a production unit with workers (k, ℓ) .

Let $\hat{V}_{k,0}^-$ and $\hat{V}_{k,\ell}^-$ be respectively defined as

$$\hat{V}_{k,0}^- = (1 - d_{k,0}(k))V_{k,0}^- + d_{k,0}(k)U_k, \quad (37)$$

and

$$\begin{aligned} \hat{V}_{k,\ell}^- &= (1 - d_{k,\ell}(k))(1 - d_{k,\ell}(\ell))V_{k,\ell}^- + (1 - d_{k,\ell}(k))d_{k,\ell}(\ell) \left[U_\ell^* + \hat{V}_{k,0}^- \right] \\ &\quad + (1 - d_{k,\ell}(\ell))d_{k,\ell}(k) \left[U_k^* + \hat{V}_{\ell,0}^- \right] + d_{k,\ell}(k)d_{k,\ell}(\ell) \left[U_k^* + U_\ell^* \right]. \end{aligned} \quad (38)$$

The optimality conditions for d imply that we can rewrite $\hat{V}_{k,0}^-$ and $\hat{V}_{k,\ell}^-$ as

$$\hat{V}_{k,0}^- = \max\{V_{k,0}^-, U_k\}, \quad (39)$$

and

$$\hat{V}_{k,\ell}^- = \max\{V_{k,\ell}^-, \hat{V}_{k,0}^- + U_\ell^*, \hat{V}_{\ell,0}^- + U_k^*\}. \quad (40)$$

Let $\hat{V}_{k,0}^* = \hat{V}_{k,0}^- + V_{0,0}^*$, $\hat{V}_{k,\ell}^* = \hat{V}_{k,\ell}^- + V_{0,0}^*$, $V_{k,0}^* = V_{k,0}^- + V_{0,0}^*$ and $V_{k,\ell}^* = V_{k,\ell}^- + V_{0,0}^*$, where $V_{0,0}^*$ represents the marginal value of an idle firm and is defined as

$$V_{0,0}^* = \beta \left\{ \left[\sum_{i,x} q^*(i, x) \max\{\hat{V}_{i,0}^* - V_{0,0}^* - v^*(i, x), 0\} \right] + V_{0,0}^* \right\}. \quad (41)$$

Given the above notation, we can write the derivative of S with respect to $u_p(k)$ as

$$\begin{aligned} U_k^* &= b_k + \beta \mathbb{E} \left\{ U_{k_+}^* + \sigma \left[0 - U_{k_+}^* \right] \right. \\ &\quad + \lambda_u p(0, 0) h_{0,0}(k, u) \left[\hat{V}_{k_+,0}^* - U_{k_+}^* - V_{0,0}^* \right] \\ &\quad + \lambda_u \sum_i p(i, 0) h_{i,0}(k, u) \left[\hat{V}_{k_+,i}^* - U_{k_+}^* - \hat{V}_{i,0}^* \right] \\ &\quad + \lambda_u \sum_{i,j} p(i, j) h_{i,j}(k, u) r_{i,j}(k, i) \left[\hat{V}_{k_+,j}^* + U_i^* - U_{k_+}^* - \hat{V}_{i,j}^* \right] \\ &\quad \left. + \lambda_u \sum_{i,j} p(i, j) h_{i,j}(k, u) r_{i,j}(k, j) \left[\hat{V}_{k_+,i}^* + U_j^* - U_{k_+}^* - \hat{V}_{i,j}^* \right] \right\}. \end{aligned} \quad (42)$$

The optimality conditions for h and r imply that we can rewrite U_k^* as

$$\begin{aligned}
U_k^* = b_k + \beta \mathbb{E} & \left\{ U_{k_+}^* + \sigma \left[0 - U_{k_+}^* \right] \right. \\
& + \lambda_u p(0, 0) \max\{\hat{V}_{k_+,0}^* - U_{k_+}^* - V_{0,0}^*, 0\} \\
& + \lambda_u \sum_i p(i, 0) \max\{\hat{V}_{k_+,i}^* - U_{k_+}^* - \hat{V}_{i,0}^*, 0\} \\
& \left. + \lambda_u \sum_{i,j} p(i, j) \max\{\max\{\hat{V}_{k_+,j}^* + U_i^*, \hat{V}_{k_+,i}^* + U_j^*\} - U_{k_+}^* - \hat{V}_{i,j}^*, 0\}. \right.
\end{aligned} \tag{43}$$

The sum of the derivative of S with respect to $e_p(k, 0)$ and $V_{0,0}^*$ can be written as

$$\begin{aligned}
V_{k,0}^* = f(k, 0) & + \beta \mathbb{E} \left\{ \hat{V}_{k_+,0}^* + \sigma \left[V_{0,0}^* - \hat{V}_{k_+,0}^* \right] + \delta \left[V_{0,0}^* + U_{k_+}^* - \hat{V}_{k_+,0}^* \right] \right. \\
& + \lambda_u p(0, 0) h_{0,0}(k, 0) \left[\hat{V}_{k_+,0}^* + V_{0,0}^* - \hat{V}_{k_+,0}^* - V_{0,0}^* \right] \\
& + \lambda_u \sum_i p(i, 0) h_{i,0}(k, 0) \left[\hat{V}_{k_+,i}^* + V_{0,0}^* - \hat{V}_{k_+,0}^* - \hat{V}_{i,0}^* \right] \\
& + \lambda_u \sum_{i,j} p(i, j) h_{i,j}(k, 0) r_{i,j}(k, i) \left[\hat{V}_{k_+,j}^* + U_i^* + V_{0,0}^* - \hat{V}_{k_+,0}^* - \hat{V}_{i,j}^* \right] \\
& + \lambda_u \sum_{i,j} p(i, j) h_{i,j}(k, 0) r_{i,j}(k, j) \left[\hat{V}_{k_+,i}^* + U_j^* + V_{0,0}^* - \hat{V}_{k_+,0}^* - \hat{V}_{i,j}^* \right] \\
& \left. + \sum_{i,x} q(i, x) h_{k_+,0}(i, x) \left[\hat{V}_{k_+,i}^* - \hat{V}_{k_+,0}^* - v^*(i, x) \right] \right\}
\end{aligned} \tag{44}$$

The optimality conditions for h and r imply that we can rewrite $V_{k,0}^*$ as

$$\begin{aligned}
V_{k,0}^* = f(k, 0) & + \beta \mathbb{E} \left\{ \hat{V}_{k_+,0}^* + \sigma \left[V_{0,0}^* - \hat{V}_{k_+,0}^* \right] + \delta \left[V_{0,0}^* + U_{k_+}^* - \hat{V}_{k_+,0}^* \right] \right. \\
& + \lambda_u p(0, 0) \max\{\hat{V}_{k_+,0}^* + V_{0,0}^* - \hat{V}_{k_+,0}^* - V_{0,0}^*, 0\} \\
& + \lambda_u \sum_i p(i, 0) \max\{\hat{V}_{k_+,i}^* + V_{0,0}^* - \hat{V}_{k_+,0}^* - \hat{V}_{i,0}^*, 0\} \\
& + \lambda_u \sum_{i,j} p(i, j) \max\{\max\{\hat{V}_{k_+,j}^* + U_i^*, \hat{V}_{k_+,i}^* + U_j^*\} - (\hat{V}_{k_+,0}^* - V_{0,0}^*) - \hat{V}_{i,j}^*, 0\} \\
& \left. + \sum_{i,x} q(i, x) \max\{\hat{V}_{k_+,i}^* - \hat{V}_{k_+,0}^* - v^*(i, x), 0\} \right\}.
\end{aligned} \tag{45}$$

The expressions in (43) and (45) coincide with those in (18) and (14). Similarly, one can derive (15) from the derivative of the planner's problem with respect to $e_p(k, \ell)$.

D Additional Tables and Figures

Table 4: EUE Sample: Summary Statistics

| Variable | Mean | SD |
|------------------------------|-------------|-----------|
| Age | 39.7 | 10.03 |
| Years of Education (imputed) | 13.15 | 2.932 |
| Tenure | 3.081 | 2.399 |
| Individual Wage (log) | 10.57 | 0.666 |
| Coworker Wage (log) | 10.13 | 0.809 |
| Number of observations | 61000 | |

Table 5: EE Sample: Summary Statistics

| Variable | Mean | SD |
|------------------------------|-------------|-----------|
| Age | 41.95 | 10.39 |
| Years of Education (imputed) | 13.54 | 2.904 |
| Tenure | 4.137 | 2.994 |
| Individual Wage (log) | 10.77 | 0.721 |
| Coworker Wage (log) | 10.27 | 1.27 |
| Wage Distance | 0.624 | 0.604 |
| Number of Observations | 9,648,000 | |

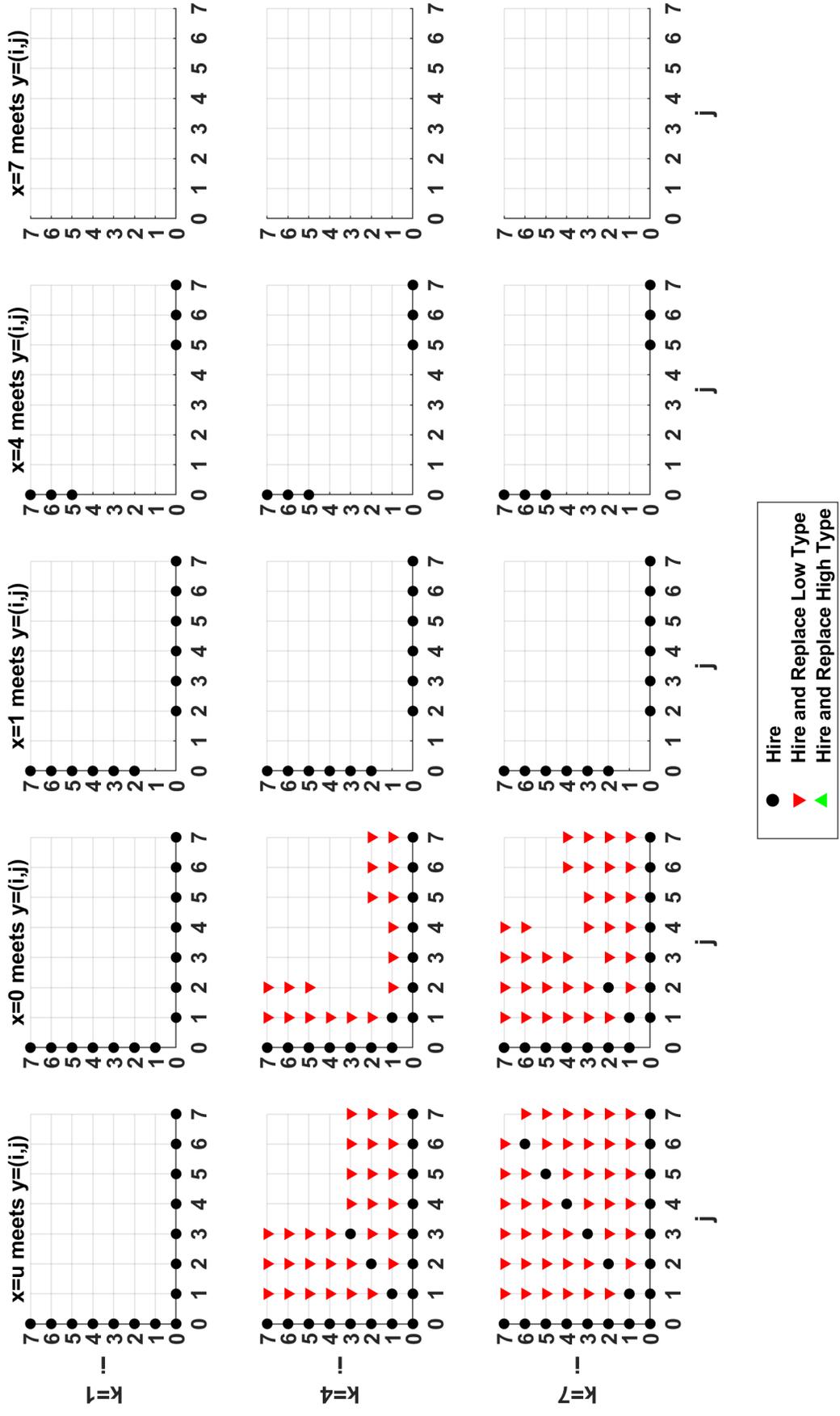


Figure 14: NLC Policy Functions

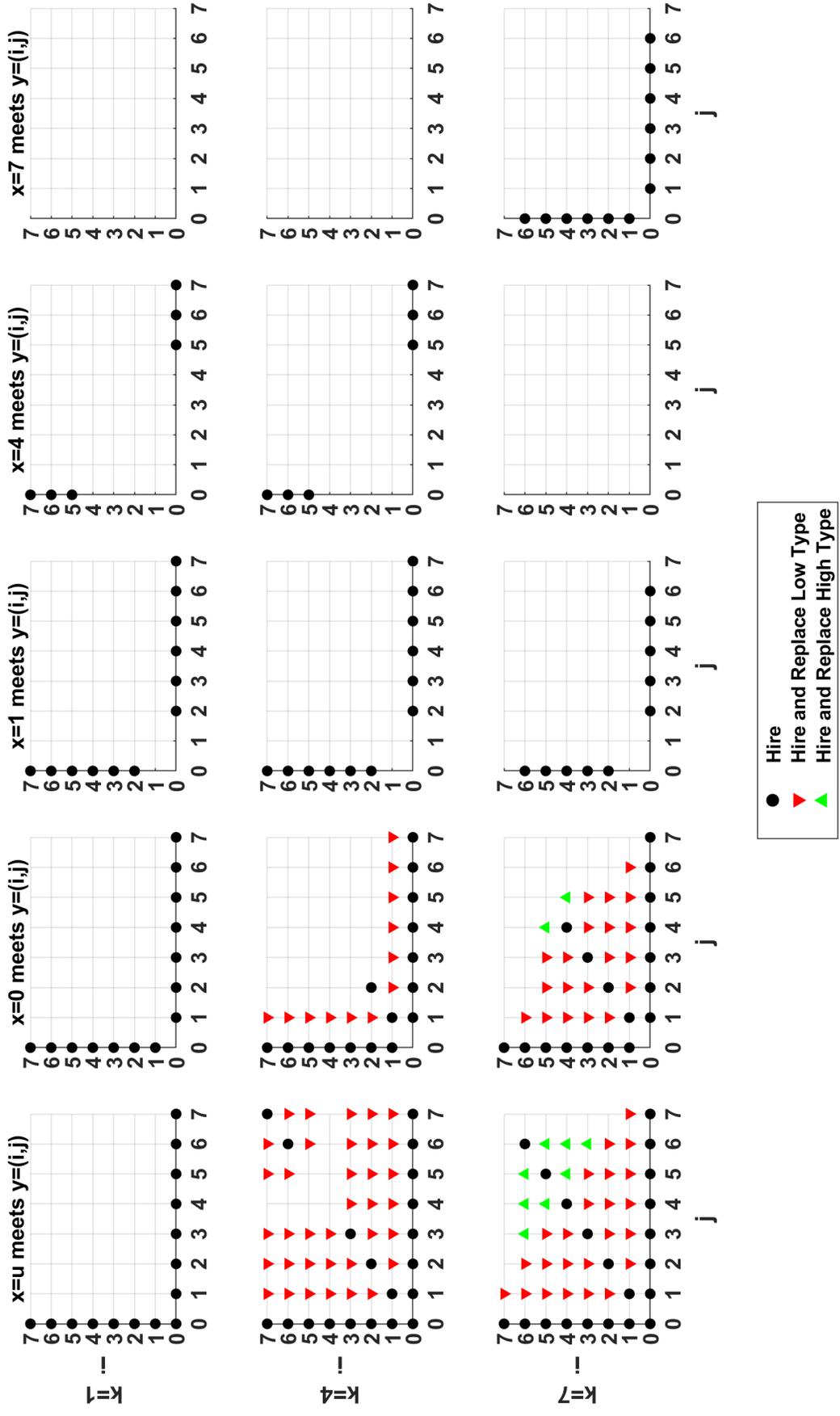


Figure 15: MP Policy Functions

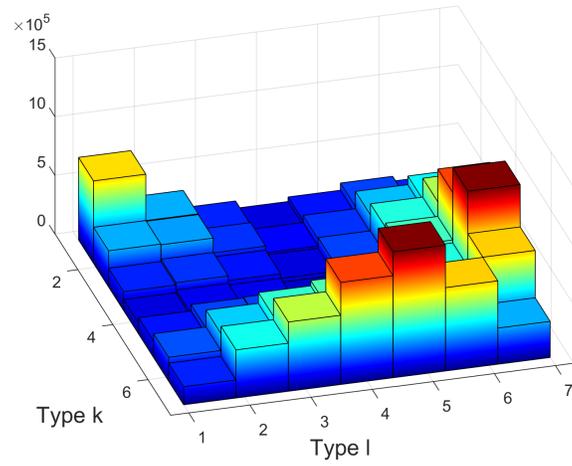


Figure 16: Equilibrium Pattern of Sorting: Baseline

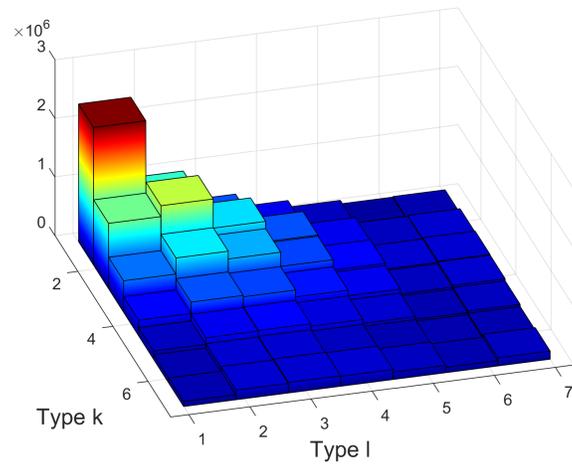


Figure 17: Equilibrium Pattern of Sorting: NLC

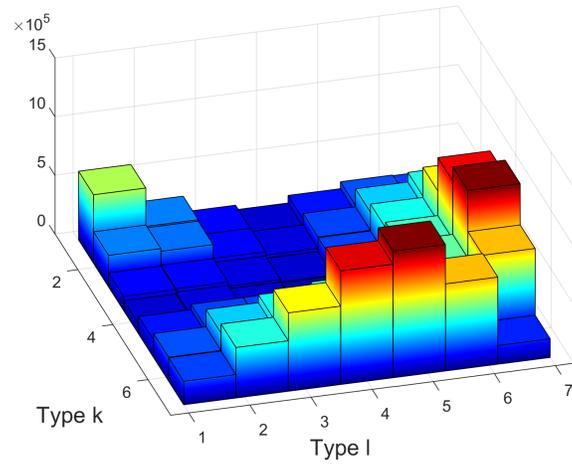


Figure 18: Equilibrium Pattern of Sorting: MP

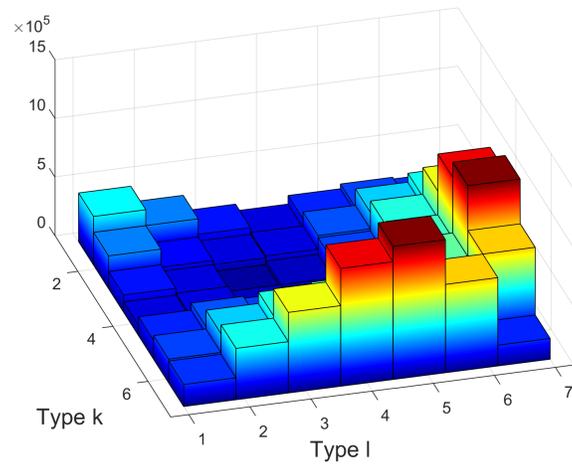


Figure 19: Equilibrium Pattern of Sorting: SP