Knowledge Diffusion in the Workplace*

Kyle Herkenhoff Jeremy Lise Guido Menzio Gordon Phillips

July 14, 2018

Preliminary

Abstract

We develop a theory of teams to measure the way knowledge diffuses across workers. We build a frictional sorting framework with production complementarities which allows for workers to influence each other's knowledge. We estimate the model using matched employer-employee data for the U.S. Our estimates imply strong peer effects. With strong peer effects, both the decentralized economy and planner optimally pair low human capital workers with high human capital workers. We show that at least 16% of measured "mismatch" (pairing high and low types) in the U.S. economy is due to peer effects. Lastly, peer effects and worker mobility are equally important determinants of output, with each factor accounting for roughly 1/6 of U.S. GDP.

*Herkenhoff: University of Minnesota. Lise: University of Minnesota. Menzio: University of Pennsylvania, & NBER. Phillips: Dartmouth College, University of Maryland, & NBER. We thank seminar participants at Aarhus, Atlanta Fed, Carnegie-Mellon, FGV, IIES, Oslo, SAET, SaM, San Francisco Fed, and the SED. We are especially indebted to Carlos Henrique Leite Corseuil for his help and comments. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation.

1 Introduction

How does knowledge diffuse across workers? In this paper, we develop a model which allows us to estimate the importance of peer effects and worker mobility for knowledge diffusion. By doing so, we contribute to a fast and growing literature on knowledge diffusion (*inter alia* Jovanovic and Rob [1989], Eeckhout and Jovanovic [2002], Luttmer [2014], Lucas and Moll [2014], Perla and Tonetti [2014], Heggedal, Moen, and Preugschat [2017]), and the relatively large existing literature on peer effects (*inter alia* Mas and Moretti [2009], Nix [2015], Cornelissen, Dustmann, and Schönberg [2016]).

Our theoretic contribution is to develop a model of teams with peer effects and on-the-job search. We build on a relatively small class of existing sorting models with dynamic types (e.g. Anderson and Smith [2010], Chade and Eeckhout [2013], Lentz and Roys [2015], Lise and Postel-Vinay [2015], and Herkenhoff, Phillips, and Cohen-Cole [2016]) by introducing dynamic types that evolve as a function of coworker characteristics in a search environment. The firm 'type' is no longer an exogenous draw from a distribution as it is commonly modeled in the frictional assignment literature (*inter alia* Shimer and Smith [2000], Hagedorn, Law, and Manovskii [2012]), but instead it is an evolving function of the set of coworkers. In our framework, workers spread knowledge through two channels: (i) through interactions with existing team members, or (ii) job transitions in which a worker leaves for another firm, either directly through on-the-job search or indirectly through a spell of unemployment, and the worker transfers whatever knowledge they have to new coworkers.

Our quantitative contribution is to estimate the sources of knowledge diffusion using the structure of our model and matched employer-employee data from the Longitudinal Employer-Household Dynamics (LEHD) database. The main challenge when estimating peer effects in our model is that existing empirical methods, such as those in Nix [2015] and Cornelissen et al. [2016], impose fixed-type assumptions that are inconsistent with our model assumptions. Likewise, many methods used to estimate the degree of sorting such as Abowd, Kramarz, and Margolis [1999], Hagedorn et al. [2012], Bonhomme, Lamadon, and Manresa [2014], and Borovickova and Shimer [2017] rely on fixed type assumptions.

We estimate the degree of learning and sorting by using the structure of our model in conjunction with the LEHD. Our approach is to indirectly infer the model's sorting and learning parameters from three moments in the LEHD, that, when viewed through the lens of our model, directly inform sorting and learning. The first moment exploits the fact that learning has unique predictions about job mobility and the wages of coworkers. Through the lens of the model, the worker's wage is a noisy proxy of an individual's type and the coworker's wage is a noisy proxy of the coworker's type. With strong production complementarity and no learning, the further a worker's wage is from their coworker's wage, the more likely they are to switch employers in search of a better match. With strong learning and no production complementarity, the further a worker's wage is from their coworker's wage, the *less* likely they are to switch employers because they are either learning or teaching their coworker. The second moment we use to disentangle learning and sorting is based on the correlation between an individual's wage and their previous coworkers' wages. To parse out sticky-wages, built-up rents, and other forces which drive a wedge between a worker's type and their true productivity, we focus on individuals who separate, experience a spell of unemployment, and then find a new job. We refer to this as an 'EUE' transition. As sorting improves, coworker wages become a stronger predictor of the worker's type, and thus coworker wages become better predictors of future individual wages. As learning improves, the same is true. Our last moment is the share of wage variance that is between firms. Since there are multiple workers at a firm in our model, our framework has a concept of within-firm and between-firm wage variance. As production complementarities increase, between-firm wage variance grows relative to within-firm wage variance. Conversely, as learning becomes more important, the within-firm wage variance share grows. The fact that learning and sorting move two of our moments in opposing directions and one of our moments in the same direction allows us to pin down both the level and relative importance of peer effects and worker complementarity.

Our estimated parameters imply very strong degrees of both skill complementarity and peer effects. We then use the estimated model to measure the relative importance of peer effects and worker mobility for U.S. output. Our first finding is that eliminating peer effects would lower output by 16.3%, even though standard measures of sorting improve. Without learning, fewer workers reach the highest skill level since low-type workers can no longer learn from their more-skilled peers. When learning is shutdown there are no incentives to generate *schools* (a school is a pairing of a low and high type worker), and so more same-type matches arise in equilibrium. As a side-effect, sorting improves. Sorting, measured by the Spearman rank correlation coefficient, increases by 25% when learning is eliminated, going from .4 to .5. Our estimates imply that learning accounts for at least 16% of measured 'mismatch.' This is not to say that learning is actually causing any mismatch by incentivizing pairings of high and low-type workers (e.g. forming schools). We demonstrate that the planner's solution features 38% of individuals in schools, 4 percentage points more than in the decentralized economy.

Eliminating endogenous reallocation of workers between jobs, but still allowing workers to learn from each other, also reduces output by 15.7%. There are two mechanisms generating this result: (i) fewer teams are formed since workers can no longer search for productive partners on-the-job, and (ii) with fewer teams, human capital diffuses more slowly. The net effect is that worker mobility explains roughly 1/6 of U.S. GDP, while shutting down both worker mobility and learning simultaneously reduces output by nearly 1/3.

Lastly, the decentralized equilibrium is inefficient. Firms are not fully compensated for educating workers. Because wages are determined via Nash-Bargaining and Bertrand competition as in Cahuc et al. [2006], firms only receive a fraction of the social surplus they generate by educating their workers. As a result, there are too few schools in the decentralized equilibrium relative to the social planner's problem. By generating more schools, we find that the social planner's allocation would increase output by 4 percentage points.

The paper proceeds as follows. Section 2 discusses the related literature in more detail. Section 3 includes the model. Section 4 describes the calibration. Section 5 illustrates the main knowledge diffusion decomposition, and Section 6 compares the decentralized economy the to planner's solution. Section 7 concludes.

2 Related Literature

Our model contributes to the literature on sorting with dynamic types (e.g. see the survey of sorting models in Chade, Eeckhout, and Smith [2017]). Anderson and Smith [2010] and Anderson [2015] consider frictionless assignment models with partner-dependent dynamic types. Related work by Jovanovic [2014] extends Anderson and Smith [2010] to allow for growth. In frictional sorting models, an equally sparse set of papers allow for one-sided dynamic types, including Chade and Eeckhout [2013], Lentz and Roys [2015] and Lise and Postel-Vinay [2015]. Lentz and Roys [2015] develop a search and matching model in which firms can invest in worker training; thus, the worker type fluctuates over time and is determined by the firm's choices. Lise and Postel-Vinay [2015] allow firms to influence the vector of worker skills. Recent work by Herkenhoff et al. [2016] allows worker human capital and firm capital to fluctuate over the course of a match, but the firm investment choice does not affect the worker's human capital. Lastly, Heggedal et al. [2017] model firm innovation choices in the presence of worker mobility in a two-period economy. Their focus is on theoretically characterizing welfare gains and losses from various innovation-related interventions in the decentralized economy. Our paper also relates to the theoretic literature on knowledge diffusion through interaction, (e.g. Jovanovic and Rob [1989], Eeckhout and Jovanovic [2002], Luttmer [2014], Lucas and Moll [2014], Perla and Tonetti [2014], Chiu et al. [2017]) and a growing literature that models knowledge diffusion through trade, (e.g. Monge-Naranjo [2012], Buera and Oberfield [2016]). Relative to these existing frameworks, our contribution is to build a frictional sorting model with peer affects and job-to-job flows.

In terms of empirics, Stoyanov and Zubanov [2012] and Serafinelli [2015] provide recent summaries of the empirical literature on worker mobility and firm productivity. These studies, as well as the majority of other papers in this literature, find that poaching high-wage workers has a large positive impact on firm productivity, although the mechanisms are not well understood.[1] In the peer-effects literature, recent work by Cornelissen et al. [2016] and Nix [2015] rely on long panel dimensions of administrative data to disentangle worker traits, such as ability, from coworker peer-effects.[2] Both Cornelissen et al. [2016] and Nix [2015] find significant, but relatively small, coworker influence on future wages. Lastly, recent empirical work by Brooks, Donovan, and Johnson [2017] measured large degrees of knowledge diffusion through a randomized controlled trial in which they matched experienced entrepreneurs with inexperienced entrepreneurs. Our contribution complements this existing body of empirical work by measuring peer effects in a framework with dynamic types.

3 Theoretical Framework

In this section, we develop a model of worker transitions and diffusion of knowledge. The model is an extension of Postel-Vinay and Robin [2002], Cahuc, Postel-Vinay, and Robin [2006], and Lise and Postel-Vinay [2015].

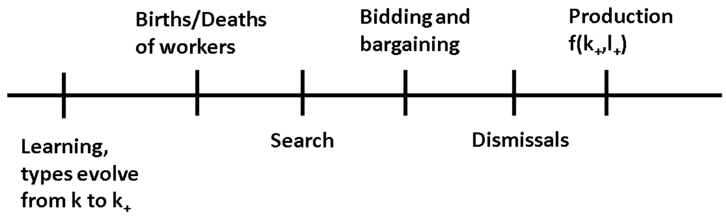
3.1 Environment

We consider a labor market populated by a continuum of workers of measure 1 and by a continuum of firms of measure $F > 0$. Every worker maximizes the present value of their

[1] A growing finance literature that measures the determinants of startups, including the role of non-compete clauses which limit the knowledge that can be transferred between employers (e.g. Babina [2015]). Our work is also related to the largely theoretic work of Chatterjee and Rossi-Hansberg [2012] on knowledge diffusion through spinoffs.

[2] A large literature summarized in Mas and Moretti [2009] measures peer effects in very specific settings, such as checkout employees at grocery stores.

Figure 1: Timing of events.



income, discounted at the factor $\beta \in (0,1)$. Workers are heterogeneous with respect to their human capital. Let $H = \{h_1, \dots, h_N\}$ denote human capital, where $0 < h_1 < h_2 < \dots h_N$, and let $k \in \mathcal{N} = \{1, \dots, N\}$ be the corresponding index of the worker's human capital. Human capital determines the worker's contribution to output when employed and the worker's income when unemployed. Moreover, the human capital of a worker determines how much can be learned from a coworker of type $l \in H$, as well as how much the coworker may learn from the worker.

Every firm maximizes the present value of its profits, discounted at the factor β. Every firm operates the same production function. If the firm has no workers, it produces no output. If the firm employs only one worker of type k, it produces $f(k,0)$ units of output. If the firm employs a worker of type k and a worker of type l, it produces $f(k,l)$ units of output, with $f(k,l) = f(l,k)$. For the sake of simplicity, we assume that the firm can employ at most two workers. For the quantitative analysis in Section 4, we specialize the production function f to be CES, e.g. $f(k,l) = (h_k^{\frac{1}{\rho}} + h_l^{\frac{1}{\rho}})^\rho$. If $\rho > 1$ then $f(k,l)$ is supermodular, if $\rho = 1$ then workers are perfect substitutes, and if $\rho < 1$ the production function is submodular.

Time is discrete and continues forever. Each period is divided into five stages: learning, births/deaths, search, bidding and bargaining, dismissal and production. Figure 1 illustrates the timing of events within a period.

At the learning stage, the human capital of a worker of type k evolves according to a law of motion that depends on the employment status of the worker as well as the human capital of coworkers. Specifically, if the worker is unemployed, the worker's human capital transitions from k to k_+ with probability $g_u(k_+ \mid k)$, with $g_u(k_+ \mid k) \in (0,1)$ and $\sum_{k_+ \in H} g_u(k_+ \mid k) = 1$. If the worker is employed without a coworker, the worker's human capital transitions from k to k_+ with probability $g_e(k_+ \mid k)$, with $g_e(k_+ \mid k) \in (0,1)$ and $\sum_{k_+ \in H} g_e(k_+ \mid k) = 1$. Finally, if the worker is employed with a coworker of type l, the worker's human capital transitions from k to k_+ with probability $g_e(k_+ \mid k, l)$, with $g_e(k_+ \mid k, l) \in (0,1)$ and $\sum_{k_+ \in H} g_e(k_+ \mid k, l) = 1$.

For the quantitative analysis, we posit a simple structure for the evolution of human capital. Unemployed workers lose human capital, and employed workers gain human capital, where the gain is faster with better coworkers. For an unemployed worker of type k, we assume that $g_u(k-1|k) = \alpha_u$ if $i > 1$ and $g_u(k-1|k) = 0$ if $k = 1$, and $g_u(k|k) = 1 - g_u(k-1|k)$. This process assumes that human capital can only depreciate while unemployed. For a worker of type k who is employed without coworkers, we assume that $g_e(k+1|k,0) = \alpha_0$ if $k < N$ and $g_e(k+1|k,0) = 0$ if $k = N$, and $g_e(k|k,0) = 1 - g_e(k+1|k,0)$. This process captures the idea that the human capital of a worker may increase over time due to accumulation of working experience, not via peer effects. For a worker of type k employed with a coworker of type $l > k$, we assume that $g_e(k+1|k,l) = \alpha_0 + \alpha_1 \frac{h_l - h_k}{h_N - h_1}$ if $k < N$ and $g_e(k+1|k,l) = 0$ if $k = N$, and $g_e(k|k,l) = 1 - g_e(k+1|k,l)$. If $l < k$, we assume that $g_e(k+1|k,l) = \alpha_0$ if $k < N$ and $g_e(k+1|k,l) = 0$ if $k = N$, and $g_e(k|k,l) = 1 - g_e(k+1|k,l)$. We refer to α_1 as a *peer effect*. This process captures the idea that the human capital of a worker may increase faster when producing in the company of a coworker with more human capital. Furthermore, these processes assume that the human capital of solo workers and team workers does not depreciate.

At the birth/death stage, some workers exit and some other workers enter the labor market. Specifically, a worker of type k exits the labor market with probability $\chi \in (0,1)$. Simultaneously, a measure χ of new workers enters the labor market. A new worker enters the market with human capital $k \in H$ with probability π_k, where $\pi_k \in [0,1]$ and $\sum_{k=1}^{N} \pi_k = 1$. A new worker enters the labor market unemployed.

At the search stage, some employed workers lose their job for exogenous reasons and some other workers come in contact with firms. Specifically, an employed worker of type k faces a probability $\delta \in (0,1)$ of losing their job and moving into unemployment. An unemployed worker of type i meets a randomly-selected firm with probability $\lambda_u \in (0,1]$. An employed worker of type i meets a randomly-selected firm with probability $\lambda_e \in [0,1]$. Clearly, a worker may meet a firm without workers, a firm employing only one worker, or a firm employing two workers. Conversely, a firm may meet an unemployed worker, a worker employed on their own, or a worker employed with a coworker.

At the bidding and bargaining stage, surplus is divided using a variant of the bargaining protocol in Cahuc et al. [2006]. Workers and firms engage in a mix of both Bertrand competition and Nash bargaining. Let σ denote the bargaining power of the worker, and let $1 - \sigma$ denote the bargaining power of the firm. If the firm contacts an unemployed worker and there is positive surplus from the match, the worker and firm Nash-bargain over surplus. If the firm (henceforth, the poacher) contacts an employed worker, the poacher and the current employer (henceforth, the incumbent) engage in Bertrand competition. If the poacher's value of hiring the employee is greater than the incumbent's value of hiring the employee, the worker moves to the poacher. The poacher and employee then Nash bargain over surplus using the highest bid from Bertrand competition as the outside option of the worker. Otherwise, the worker stays with the incumbent, and potentially renegotiates the split of surplus. Note that if the poacher has two employees, it must dismiss one of them if it succeeds in hiring the worker. During Bertrand competition, we assume that both the incumbent and poacher observe each other's fundamental states (i.e. how many workers each employs and of what type), but not the offers. Therefore, bids can only be contingent on the fundamental state and not on offers.

At the dismissal stage, the firm may separate from any of its employees. At the production stage, firms and workers produce, workers receive their income and firms receive their profits. Specifically, an employed worker of type k who does not have a coworker produces $f(k, 0)$ units of output and receives the income that is specified by the employment contract. An employed worker of type k and a coworker of type l produce $f(k, l)$ units of output. The worker of type k receives the income that is specified by the employment contract. An unemployed worker of type k receives an income of $b(k)$, which the reader may interpret as the value of home production, as an unemployment benefit, or as a combination of the two. For the quantitative analysis, we assume that $b(k) = \phi h_k$, with $\phi \in (0, 1)$. That is, the income of an unemployed worker of type k is a fraction ϕ of the worker's human capital.

We assume that employment contracts are complete. That is, an employment contract specifies the worker's employment probability contingent on the entire history of the firm-worker match (e.g., the worker's type, the worker's contact with a poaching firm, the coworker's type, the coworker's contact with another firm, and the firm's contact with another worker), the worker's wage conditional on the history of the match, as well as the offers that the firm makes to other workers. We assume that an employment contract must satisfy the worker's participation constraint. The constraint requires that, in all contingencies in which the contract specifies the worker to be employed at the production stage, the lifetime utility of the contract to the worker must be greater than the value of unemployment. Furthermore, in all contingencies in which the worker receives an offer from a poacher and the contract specifies that the worker stays with the incumbent, the lifetime utility of the contract to the worker must be greater than the expected value of the poacher's offer. The contract need not satisfy the firm's participation constraint because the firm can commit to participate.

3.2 Definition of Equilibrium

In order to define an equilibrium, we need to introduce some notation. We denote as $U(k)$ the lifetime utility of a worker of type k who is unemployed. We denote as Π_0 the lifetime profit of a firm that has no employees. We denote as $V_1(k)$ the sum of the lifetime profit of a firm and the lifetime utility of a worker of type k who are matched together. We refer to $V_1(k)$ as the joint value of a match between a firm and a worker of type k. We denote as $V_2(k,l)$ the sum of the lifetime profit of a firm, the lifetime utility of a worker of type k, and the lifetime utility of a worker of type l who are matched together. We refer to $V_2(k,l)$ as the joint value of a match between a firm and a team of workers of type (k,l). The value functions $U(k)$, $V_1(k)$ and $V_2(k,l)$ are all evaluated at the beginning of the production stage. We also find it useful to denote as $\widehat{V}_1(k)$ and as $\widehat{V}_2(k,l)$ the joint values of firm-worker matches evaluated at the beginning of the dismissal stage.

We measure the distribution of workers across types and employment states at the beginning of the search stage. We denote as u_k the measure of workers of type k who are unemployed at the beginning of the search stage. We denote as $e_{k,0}$ the measure of workers of type k who are employed without a coworker at the beginning of the search stage. Finally, we denote as $e_{k,l}$ the measure of workers of type k who are employed with a coworker of type l at the beginning of the search stage.

We restrict attention to equilibria with the following bidding and bargaining protocol. Consider a poaching firm contacting an unemployed worker of type i. Let v_P denote the marginal value of the worker at the firm, where v_P is equal to $\widehat{V}_1(k) - \Pi_0$ if the firm has no employees, $\widehat{V}_2(k,i) - \widehat{V}_1(k)$ if the firm has an employee of type k, $\max\{\widehat{V}_2(k,i), \widehat{V}_2(i,l)\} - \widehat{V}_2(k,l)$ if the firm has two employees of type k and l. If $v_P > U(i)$, the firm and worker Nash-bargain over surplus. If $v_P \leq U(i)$, no match is formed.

Now, consider a firm contacting a worker of type i who is already employed. The marginal value of the worker at the poaching firm is v_P. Let v_I denote the marginal value of the worker at the incumbent, where v_I is equal to $\widehat{V}_1(i) - \Pi_0$ if the worker is the only employee of the incumbent and $\widehat{V}_2(i,j) - \widehat{V}_1(j)$ if the incumbent also employs a worker of type j. There are three cases. If $v_P > v_I$, the worker moves the poaching firm, and they Nash-bargain over the division of surplus. The worker's outside option is v_I which is the most the incumbent firm is willing to bid for the worker under Bertrand competition. Therefore, Nash bargaining yields a value to the worker of $v_I + \sigma[v_P - v_I]$. If $v_P = v_I$, the poacher and the incumbent both bid v_I and the worker stays with the incumbent. If $v_P < v_I$, the poacher bids v_P, the incumbent offers a continuation value greater than v_P, and the worker stays with the incumbent.

At the dismissal stage, the firm decides which workers to dismiss. Consider the case of a firm with one type k employee. If the firm retains the employee, the joint continuation value is $V_1(k)$. If the firm dismisses the employee, the continuation value to the firm is Π_0 and the continuation value to the worker is $U(k)$. Therefore, the joint value $\widehat{V}_1(k)$ of a firm and an employee of type k at the beginning of the dismissal stage is such that

$$\widehat{V}_1(k) = \max\{V_1(k), \Pi_0 + U(k)\}. \tag{1}$$

Consider a firm with two employees, one of type k and one of type l. If the firm retains both employees, the joint continuation value is $V_2(k,l)$. If the firm retains the k employee and dismisses the l employee, the joint continuation value of the firm and the k employee is $V_1(k)$ and the continuation value of the l employee is $U(l)$. Similarly, if the firm retains the l employee and dismisses the k employee, the joint continuation value to the firm and the l employee is $V_1(l)$ and the continuation value of the k employee is $U(k)$. Finally, if the firm dismisses both employees, the continuation value of the firm is Π_0, the continuation value of the k worker is $U(k)$ and the continuation value of the l worker is $U(l)$. The joint value $\widehat{V}_2(k,l)$ of a firm, an employee of type k and an employee of type l at the beginning of the dismissal stage is such that

$$\widehat{V}_2(k,l) = \max\{V_2(k,l), V_1(k) + U(l), V_1(l) + U(k), \Pi_0 + U(k) + U(l)\}. \tag{2}$$

The value Π_0 of a firm without employees at the beginning of the production stage is such that

$$\begin{aligned} \Pi_0 =& f(0,0) + \beta\Big[\sum_{i=1}^{N} \frac{\lambda_u u_i}{F}(1-\sigma)\max\{\widehat{V}_1(i) - U(i) - \Pi_0, 0\} \\ &+ \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\lambda_e e_{i,j}}{F}(1-\sigma)\max\{\widehat{V}_1(i) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] - \Pi_0, 0\} + \Pi_0\Big]. \end{aligned} \tag{3}$$

The flow profit of the firm is 0. In the next period, the firm contacts an unemployed worker of type i with probability $\lambda_u u_i/F$. If there is positive surplus from the match, the firm hires the worker and the firm's continuation profit is a share $(1-\sigma)$ of surplus, $\widehat{V}_1(i) - U(i) - \Pi_0$. With probability $\lambda_e e_{i,0}/F$, the firm contacts a worker of type i who is currently employed without a coworker. In this case, the incumbent keeps the worker since the employee has the same marginal product at both the incumbent and poaching firm. Finally, with probability $\lambda_e e_{i,j}/F$, the firm contacts a worker of type i who is employed with a worker of type j. Suppose the worker's marginal product at the poaching firm, $\widehat{V}_1(i) - \Pi_0$, is greater than the worker's marginal product at the incumbent, $\widehat{V}_2(i,j) - \widehat{V}_1(j)$. Then through Bertrand competition, the worker's outside option is their marginal product at the incumbent firm, $\widehat{V}_2(i,j) - \widehat{V}_1(j)$. The firm hires the worker and they Nash-bargain over the division of surplus. The firm's continuation profit is a share $(1-\sigma)$ of the surplus, $\widehat{V}_2(k_+, i) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_1(k_+)$. If the worker's marginal product is smaller at the poaching firm, then the poacher does not hire the worker.

The joint value $V_1(k)$ of a match between a firm and an employee of type k at the beginning of the production stage is such that

$$\begin{aligned}
V_1(k) = f(k,0) &+ \beta \mathbb{E}_{k_+} \Big[\delta[U(k_+) + \Pi_0 - \widehat{V}_1(k_+)] + \chi[\Pi_0 - \widehat{V}_1(k_+)] \\
&+ \sum_{i=1}^{N} \frac{\lambda_u u_i}{F} (1-\sigma) \max\{\widehat{V}_2(k_+, i) - U(i) - \widehat{V}_1(k_+), 0\} \\
&+ \sum_{i=1}^{N} \frac{\lambda_e e_{i,0}}{F} (1-\sigma) \max\{\widehat{V}_2(k_+, i) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_1(k_+), 0\} \\
&+ \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\lambda_e e_{i,j}}{F} (1-\sigma) \max\{\widehat{V}_2(k_+, i) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] - \widehat{V}_1(k_+), 0\} \\
&+ \lambda_e \frac{F_0}{F} \sigma \max\{\widehat{V}_1(k_+) - \Pi_0 - (\widehat{V}_1(k_+) - \Pi_0), 0\} \\
&+ \lambda_e \sum_i \frac{e_{i,0}}{F} \sigma \max\{\widehat{V}_2(k_+, i) - \widehat{V}_1(i) - (\widehat{V}_1(k_+) - \Pi_0), 0\} \\
&+ \lambda_e \sum_i \sum_j \frac{e_{i,j}}{2F} \sigma \max\{\widehat{V}_2(k_+, i) + U(j) - \widehat{V}_2(i,j) - (\widehat{V}_1(k_+) - \Pi_0), \\
&\qquad \widehat{V}_2(k_+, j) + U(i) - \widehat{V}_2(i,j) - (\widehat{V}_1(k_+) - \Pi_0), 0\} \\
&+ \widehat{V}_1(k_+) \Big].
\end{aligned} \tag{4}$$

In the current period, the joint income of the firm-employee match is $f(k,0)$. At the learning stage of next period, the employee's type becomes k_+ with probability $g_e(k_+|k,0)$. At the search stage of next period, the firm and the employee separate for exogenous reasons with probability δ. In this case, the continuation value of the match is $U(k_+) + \Pi_0$. Likewise, with probability χ, the worker dies. With probability $\lambda_u u_i/F$, the firm contacts an unemployed worker of type i. In this case, the joint value of matching is $\widehat{V}_2(k_+, i)$, the worker's outside option is $U(i)$, and the firm's outside option is $\widehat{V}_1(k)$. Thus, the surplus of the match is $\widehat{V}_2(k_+, i) - U(i) - \widehat{V}_1(k_+)$. If positive, the firm obtains a share $(1-\sigma)$ of the surplus and the worker is hired out of unemployment. With probability $\lambda_e e_{i,0}/F$, the firm contacts a worker of type i who is employed without a coworker. In this case, the joint value of matching is $\widehat{V}_2(k_+, i)$, the outside option of the worker is $(\widehat{V}_1(i) - \Pi_0)$, and the outside option of the firm is Π_0. Therefore, the surplus of the match is $\widehat{V}_2(k_+, i) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_1(k_+)$, and if its positive the firm adds the type i worker to form a team. Lastly, with probability $\lambda_e e_{i,j}/F$, the firm contacts a worker of type i who is employed with a coworker of type j. In this case, the joint value of matching is $\widehat{V}_2(k_+, i)$, the outside option of the worker is $\widehat{V}_2(i, j) - \widehat{V}_1(j)$, and the outside option of the firm is Π_0. If surplus is positive, the firm poaches i and forms a team.

At the search stage, the employee may come into contact with a poacher. The employee contacts a poacher without workers with probability $\lambda_e F_0/F$, a poacher with a worker of type i with probability $\lambda_e e_{i,0}/F$, and a poacher with a team of workers of type (i, j) with probability $\lambda_e e_{i,j}/(2F)$. When the employee comes into contact with a poacher, the continuation value of the match is unuchanged if the marginal value of the employee is higher at the incumbent than at the poacher. If the marginal value of the employee is higher at the poacher than at the incumbent, the match dissolves and the employee is hired by the poacher. When the worker is poached, the outside option of the worker is their current marginal product, $(\widehat{V}_1(k_+) - \Pi_0)$. The marginal product of the worker at the poacher is $\widehat{V}_1(k_+) - \Pi_0$ at the poacher with zero employees, $\widehat{V}_2(k_+, i) - \widehat{V}_1(i)$ at the poacher with one type i employee, and $\max\{\widehat{V}_2(k_+, i) + U(j) - \widehat{V}_2(i, j), \widehat{V}_2(k_+, j) + U(i) - \widehat{V}_2(i, j)\}$ at the poacher with an (i, j) team. The worker and poacher then Nash-bargain. Thus, the joint worker-firm team capture σ, the worker's share, of the surplus generated from the move.

The joint value $V_2(k, l)$ of a match between a firm, an employee of type k and an employee of type l at the beginning of the production stage is such that

$$
\begin{aligned}
V_2(k,l) =& f(k,l) + \beta \mathbb{E}_{k_+,l_+}\Big[\delta(U(k_+) + \widehat{V}_1(l_+) - \widehat{V}_2(k_+,l_+)) + \delta(U(l_+) + \widehat{V}_1(k_+) - \widehat{V}_2(k_+,l_+)) \\
&+ \chi(\widehat{V}_1(l_+) - \widehat{V}_2(k_+,l_+)) + \chi(\widehat{V}_1(k_+) - \widehat{V}_2(k_+,l_+)) \\
&+ \sum_{i=1}^{N} \frac{\lambda_u u_i}{F}(1-\sigma)\max\{\widehat{V}_2(k_+,i) + U(l_+) - U(i) - \widehat{V}_2(k_+,l_+), \\
&\qquad \widehat{V}_2(i,l_+) + U(k_+) - U(i) - \widehat{V}_2(k_+,l_+), 0\} \\
&+ \sum_{i=1}^{N} \frac{\lambda_e e_{i,0}}{F}(1-\sigma)\max\{\widehat{V}_2(k_+,i) + U(l_+) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k_+,l_+), \\
&\qquad \widehat{V}_2(i,l_+) + U(k_+) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k_+,l_+), 0\} \\
&+ \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\lambda_e e_{i,j}}{F}(1-\sigma)\max\{\widehat{V}_2(k_+,i) + U(l_+) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] - \widehat{V}_2(k_+,l_+), \\
&\qquad \widehat{V}_2(i,l_+) + U(k_+) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] - \widehat{V}_2(k_+,l_+), 0\} \\
&+ \lambda_e \frac{F_0}{F}\sigma \max\{\widehat{V}_1(k_+) - \Pi_0 - (\widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)), 0\} \\
&+ \lambda_e \sum_i \frac{e_{i,0}}{F}\sigma \max\{\widehat{V}_2(k_+,i) - \widehat{V}_1(i) - (\widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)), 0\} \\
&+ \lambda_e \sum_i \sum_j \frac{e_{i,j}}{2F}\sigma \max\{\widehat{V}_2(k_+,i) + U(j) - \widehat{V}_2(i,j) - (\widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)), \\
&\qquad \widehat{V}_2(k_+,j) + U(i) - \widehat{V}_2(i,j) - (\widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)), 0\} \\
&+ \lambda_e \frac{F_0}{F}\sigma \max\{\widehat{V}_1(l_+) - \Pi_0 - (\widehat{V}_2(k_+,l_+) - \widehat{V}_1(k_+)), 0\} \\
&+ \lambda_e \sum_i \frac{e_{i,0}}{F}\sigma \max\{\widehat{V}_2(l_+,i) - \widehat{V}_1(i) - (\widehat{V}_2(k_+,l_+) - \widehat{V}_1(k_+)), 0\} \\
&+ \lambda_e \sum_i \sum_j \frac{e_{i,j}}{2F}\sigma \max\{\widehat{V}_2(l_+,i) + U(j) - \widehat{V}_2(i,j) - (\widehat{V}_2(k_+,l_+) - \widehat{V}_1(k_+)), \\
&\qquad \widehat{V}_2(l_+,j) + U(i) - \widehat{V}_2(i,j) - (\widehat{V}_2(k_+,l_+) - \widehat{V}_1(k_+)), 0\} \\
&+ \widehat{V}_2(k_+,l_+)\Big].
\end{aligned} \tag{5}
$$

In the current period, the joint income of the match is $f(k,l)$. At the learning stage of next period, the first employee's type becomes k_+ with probability $g_e(k_+|k,l)$ and the second employee's type becomes l_+ with probability $g_e(l_+|l,k)$. At the search stage of next period, the first employee moves into unemployment with probability δ. In this case, the continuation value of the match is $U(k_+) + \widehat{V}_1(l_+)$. The second employee moves into unemployment with probability δ, as well. In this case, the continuation value of the match is $U(l_+) + \widehat{V}_1(k_+)$.

Likewise, either employee may die with probability χ.

With probability $\lambda_u u_i/F$, the firm contacts an unemployed worker of type i. The joint value of matching is the maximum of replacing employee k_+, $\widehat{V}_2(i,l_+) + U(k_+)$, or replacing employee l_+, $\widehat{V}_2(k_+,i) + U(l_+)$. The outside option of the worker is $U(i)$ and the outside option of the firm is $\widehat{V}_2(k_+,l_+)$. The surplus of matching is $\max\{\widehat{V}_2(i,l_+)+U(k_+), \widehat{V}_2(k_+,i)+U(l_+)\} - U(i) - \widehat{V}_2(k_+,l_+)$. If surplus is positive, the firm captures $(1-\sigma)$ of the surplus, replaces one of its workers, and hires the unemployed agent.

With probability $\lambda_e e_{i,0}/F$, the firm contacts a worker of type i who is employed without a coworker. The joint value of matching and the outside option of the firm remain the same, however, the worker's outside option is now equal to their marginal product $\widehat{V}_1(i) - \Pi_0$. With probability $\lambda_e e_{i,j}/F$, the firm contacts a worker of type i who is employed with a worker of type j. Again, the joint value of producing is the same, except now the worker's outside option is $\widehat{V}_2(i,j) - \widehat{V}_1(j)$.

The employees may also come into contact with poachers. With probability $\lambda_e F_0/F$ the k_+ employee meets a poacher with no employees. The joint value of forming the match is $\widehat{V}_1(k_+)$, the outside option of the poacher is Π_0, and the outside option of the worker is their current marginal product at the incumbent, $\widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)$. If surplus is positive, the worker moves and captures σ of the surplus. With probability $\lambda_e e_{i,0}/F$ the k_+ employee meets a poacher with a type i worker, and with probability $\lambda_e e_{i,j}/2F$ the k_+ employee meets a poacher with an (i,j) team. Likewise, an identical set of contacts can occur for the l_+ employee.

Finally, the lifetime utility $U(k)$ of an unemployed worker of type k at the beginning of the production stage is given by

$$\begin{aligned} U(k) =& b(k) + \beta\mathbb{E}_{k+}[(1-\chi)U(k_+) + \frac{\lambda_u F_0}{F}\sigma \max\{\widehat{V}_1(k_+) - \Pi_0 - U(k_+), 0\} \\ &+ \sum_i \frac{\lambda_u e_{i,0}}{F}\sigma \max\{\widehat{V}_2(k_+,i) - \widehat{V}_1(i) - U(k_+), 0\} \\ &+ \sum_i \sum_j \frac{\lambda_u e_{i,j}}{2F}\sigma \max\{\widehat{V}_2(k_+,i) + U(j) - \widehat{V}_2(i,j) - U(k_+), \\ &\qquad \widehat{V}_2(k_+,j) + U(i) - \widehat{V}_2(i,j) - U(k_+), 0\}]. \end{aligned} \tag{6}$$

The worker receives an income of $b(k)$ in the current period. At the learning stage of next period, the worker's type becomes k_+ with probability $g_u(k_+|k)$. With probability χ the employee dies. At the search stage, the worker does not contact any firm with probability $1-\lambda_u$. In this case, the worker's continuation value is $U(k_+)$. With probability $\lambda_u F_0/F$, the unemployed worker contacts a firm with zero employees, with probability $\lambda_u e_{i,0}/F$ the unemployed worker contacts a firm with a type i employee, and with $\lambda_u e_{i,j}/2F$ the unemployed worker contacts a firm with an (i,j) team. In each case, the worker obtains a share σ of the surplus if a match is consummated.

3.2.1 Distributions of Unemployed and Employed Workers

In this section we describe how the distribution of unemployed agents, u_k, evolves. Appendix A describes the distributions of workers at single worker firms and two worker firms, $(e_{k,0}, e_{k,l})$.

We split the distribution of type k unemployed workers within the period into four subperiods: (1) the initial distribution u_k^-, (2) the distribution after the learning stage u_k^{learn}, (2) the distribution after births and deaths u_k (the start of the search stage), (3) the distribution after search outcomes are realized u_k^{search}, and (4) the distribution after dismissals occur, u_k^+. u_k^+ becomes the initial distribution for the next period.

The distribution of unemployed agents after the learning stage is given by

$$u_k^{learn} = u_k^- + \sum_{j\neq k} u_j^- g_u(k \mid j) - \sum_{j\neq k} u_k^- g_u(j \mid k). \tag{7}$$

Agents then die at a rate of χ, and newborns enter the economy initially unemployed. The newborns draw their type from the discrete pdf $\pi(k)$. Therefore, the distribution of unemployed agents after the birth stage (at the *start* of the search stage) is given by

$$u_k = (1-\chi)u_k^{learn} + \chi\Big(\sum_k u_k^{learn} + \sum_k e_{k,0}^{learn} + \sum_k \sum_l e_{k,l}^{learn}\Big)\pi(k). \tag{8}$$

There are several events that result in the flow of a type k worker into unemployment during the search stage. The first three terms in equation (9) account for workers flowing out of unemployment by meeting vacant, single worker firms, or two worker firms. The next three terms in equation (9) account for workers flowing into unemployment by being replaced when their firm meets an agent that is unemployed, working at a single worker firm, or working at a two worker firm. The last line of equation (9) accounts for exogenous layoffs.

$$
\begin{aligned}
u_k^{search} &= u_k - u_k \lambda_u \frac{F_0}{F} \mathbb{I}(\widehat{V}_1(k) - \Pi_0 - U(k) > 0) \\
&- u_k \sum_i \frac{\lambda_u e_{i,0}}{F} \mathbb{I}(\widehat{V}_2(k,i) - U(k) - \widehat{V}_1(i) > 0) \\
&- u_k \sum_i \sum_j (\lambda_u \frac{e_{i,j}}{2F} \mathbb{I}(\max\{\widehat{V}_2(k,j) + U(i) - U(k) - \widehat{V}_2(i,j), \widehat{V}_2(i,k) + U(j) - U(k) - \widehat{V}_2(i,j)\} > 0)) \\
&+ \sum_i e_{k,l} \frac{\lambda_u u_i}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l)\} > 0) \\
&\qquad \times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l)) \\
&+ \sum_i e_{k,l} \frac{\lambda_e e_{i,0}}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l)\} > 0) \\
&\qquad \times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l)) \\
&+ \sum_i \sum_j e_{k,l} \frac{\lambda_e e_{i,j}}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l)\} > 0) \\
&\qquad \times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l)) \\
&+ \delta \Big(e_{k,0} + \sum_i e_{k,i} \Big)
\end{aligned}
\tag{9}
$$

Equation (10) describes the distribution of unemployment in the dismissal stage. The first line of equation (10) captures single worker firm dismissals, and the second and third lines capture two worker firm dismissals. For two worker firms, any combination of workers may be dismissed.

$$
\begin{aligned}
&u_k^+ = u_k^{search} + e_{k,0}^{search} \mathbb{I}(\widehat{V}_1(k) < \Pi_0 + U(k)) \\
&+ e_{k,l}^{search} \mathbb{I}[max\{U(k) + U(l) + \Pi_0 - V_2(k,l), V_1(k) + U(l) - V_2(k,l), V_1(l) + U(k) - V_2(k,l), 0\} = U(k) + U(l) + \Pi_0 - V_2(k,l)] \\
&+ e_{k,l}^{search} \mathbb{I}[max\{U(k) + U(l) + \Pi_0 - V_2(k,l), V_1(k) + U(l) - V_2(k,l), V_1(l) + U(k) - V_2(k,l), 0\} = V_1(l) + U(k) - V_2(k,l)]
\end{aligned}
\tag{10}
$$

3.2.2 Equilibrium Definition

A stationary equilibrium is a tuple of value functions $\{\Pi_0, V_1, V_2, \widehat{V}_1, \widehat{V}_2, U\}$ together with a distribution of workers across employment states $\{u, e\}$ such that: (i) The value functions satisfy the Bellman Equations (1) to (6); (ii) The distribution $\{u, e\}$ is stationary given the transitions implied by the value functions.

3.3 Wage determination

We assume promised values are delivered using a fixed wage. The wage changes only if the worker receives an outside offer, a coworker is added, a coworker is lost, or the promised value of the worker exceeds the worker's marginal product.

Let v_W denote the current promised value to the worker. There are three cases to consider. Suppose the worker contacts a poaching firm where $v_P < v_W$. Then the contact does not generate a wage change. Suppose the worker meets a poaching firm where $v_W < v_P < v_I$. The worker stays with the incumbent and triggers Bertrand competition. The worker receives a new constant wage which delivers a promised value equal to v_P. Suppose the worker meets a poaching firm where $v_I < v_P$. The worker switches employers. The worker's outside option is to trigger Bertrand competition which yields, v_I. The worker then Nash-bargains over the remaining surplus generated with the poaching firm. The worker receives a new constant wage which delivers the promised value of $v_I + \sigma[v_P - v_I]$.

To facilitate exposition, we introduce some additional notation. Let the value of obtaining an outside offer to an employee from a poacher with marginal product v_P be

$$A(v_P, v_I) = \min\{v_P, v_I + \sigma \max\{v_P - v_I, 0\}\}.$$

The value function of a type k worker being paid w is given by,

$$
\begin{aligned}
&W_1(k,w) = w + \beta\mathbb{E}_{k+}[\delta U(k_+) \\
&+ \lambda_e \frac{F_0}{F}\max\{\min\{W_1(k_+,w),\widehat{V}_1(k_+)-\Pi_0\},U(k_+),A\big(\widehat{V}_1(k_+)-\Pi_0,\widehat{V}_1(k_+)-\Pi_0\big)\} \\
&+ \sum_i \lambda_e \frac{e_{i,0}}{F}\max\Big\{\min\{W_1(k_+,w),\widehat{V}_1(k_+)-\Pi_0\},U(k_+),A\big(\widehat{V}_2(k_+,i)-\widehat{V}_1(i),\widehat{V}_1(k_+)-\Pi_0\big)\Big\} \\
&+ \sum_i\sum_j \lambda_e \frac{e_{i,j}}{2F}\max\Big\{\min\{W_1(k_+,w),\widehat{V}_1(k_+)-\Pi_0\},U(k_+), \\
&\qquad A\Big(\max\{\widehat{V}_2(k_+,i)+U(j)-\widehat{V}_2(i,j),\widehat{V}_2(k_+,j)+U(i)-\widehat{V}_2(i,j)\},\widehat{V}_1(k_+)-\Pi_0\Big)\Big\} \\
&+ \sum_{i=1}^N \frac{\lambda_u u_i}{F}\mathbb{I}(\widehat{V}_2(k_+,i)-U(i)>\widehat{V}_1(k_+))\max\{\min\{W_2(k_+,i,w),\widehat{V}_2(k_+,i)-\widehat{V}_1(i)\},U(k_+)\} \\
&+ \sum_{i=1}^N \frac{\lambda_e e_{i,0}}{F}\mathbb{I}(\widehat{V}_2(k_+,i)-(\widehat{V}_1(i)-\Pi_0)>\widehat{V}_1(k_+))\max\{\min\{W_2(k_+,i,w),\widehat{V}_2(k_+,i)-\widehat{V}_1(i)\},U(k_+)\} \\
&+ \sum_{i=1}^N\sum_{j=1}^N \frac{\lambda_e e_{i,j}}{F}\mathbb{I}(\widehat{V}_2(k_+,i)-[\widehat{V}_2(i,j)-\widehat{V}_1(j)]>\widehat{V}_1(k_+))\max\{\min\{W_2(k_+,i,w),\widehat{V}_2(k_+,i)-\widehat{V}_1(i)\},U(k_+)\}] \\
&+ (1-\delta-\lambda_e-\chi-H(k))\max\{\min\{W_1(k_+,w),\widehat{V}_1(k_+)-\Pi_0\},U(k_+)\}
\end{aligned}
$$

With probability $H(k)$, a coworker is added:

$$
\begin{aligned}
H(k) = &\sum_{i=1}^N \frac{\lambda_u u_i}{F}\mathbb{I}(\widehat{V}_2(k_+,i)-U(i)>\widehat{V}_1(k_+)) \\
&+ \sum_{i=1}^N \frac{\lambda_e e_{i,0}}{F}\mathbb{I}(\widehat{V}_2(k_+,i)-(\widehat{V}_1(i)-\Pi_0)>\widehat{V}_1(k_+)) \\
&+ \sum_{i=1}^N\sum_{j=1}^N \frac{\lambda_e e_{i,j}}{F}\mathbb{I}(\widehat{V}_2(k_+,i)-[\widehat{V}_2(i,j)-\widehat{V}_1(j)]>\widehat{V}_1(k_+))
\end{aligned}
$$

The value function of a type k worker being paid w while matched with a type l coworker is given by,

$$
\begin{aligned}
W_2(k,l,w) &= w + \beta \mathbb{E}_{k_+,l_+}[(\delta + R(k_+,l_+))U(k_+) + P(k_+,l_+)\max\{\min\{W_1(k_+,w),\widehat{V}_1(k_+) - \Pi_0\}, U(k_+)\} \\
&+ \lambda_e \frac{F_0}{F} \max\Big\{ \min\{W_2(k_+,l_+,w), \widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)\}, U(k_+), A\big(\widehat{V}_1(k_+) - \Pi_0, \widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)\big)\Big\} \\
&+ \sum_i \lambda_e \frac{e_{i,0}}{F} \max\Bigg\{ \min\{W_2(k_+,l_+,w), \widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)\}, U(k_+), A\big(\widehat{V}_2(k_+,i) - \widehat{V}_1(i), \widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)\big)\Bigg\} \\
&+ \sum_i \sum_j \lambda_e \frac{e_{i,j}}{2F} \max\Bigg\{ \min\{W_2(k_+,l_+,w), \widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)\}, U(k_+), \\
&\qquad A\Big(\max\{\widehat{V}_2(k_+,i) + U(j) - \widehat{V}_2(i,j), \widehat{V}_2(k_+,j) + U(i) - \widehat{V}_2(i,j)\}, \widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)\Big)\Bigg\} \\
&+ \sum_{i=1}^N \frac{\lambda_u u_i}{F} \mathbb{I}(\widehat{V}_2(k_+,i) + U(l_+) - U(i) > \widehat{V}_2(k_+,l_+)) \max\{\min\{W_2(k_+,i,w), \widehat{V}_2(k_+,i) - \widehat{V}_1(i)\}, U(k_+)\} \\
&+ \sum_{i=1}^N \frac{\lambda_e e_{i,0}}{F} \mathbb{I}(\widehat{V}_2(k_+,i) + U(l_+) - (\widehat{V}_1(i) - \Pi_0) > \widehat{V}_2(k_+,l_+)) \max\{\min\{W_2(k_+,i,w), \widehat{V}_2(k_+,i) - \widehat{V}_1(i)\}, U(k_+)\} \\
&+ \sum_{i=1}^N \sum_{j=1}^N \frac{\lambda_e e_{i,j}}{F} \mathbb{I}(\widehat{V}_2(k_+,i) + U(l_+) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] > \widehat{V}_2(k_+,l_+)) \max\{\min\{W_2(k_+,i,w), \widehat{V}_2(k_+,i) - \widehat{V}_1(i)\}, U(k_+)\}] \\
&+ (1 - \delta - \lambda_e - \chi - R(k_+,l_+) - P(k_+,l_+) - Q(k_+,l_+)) \max\{\min\{W_2(k_+,l_+,w), \widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+)\}, U(k_+)\}
\end{aligned}
$$

The probability a type k worker is replaced in a (k,l) team is given by,

$$
\begin{aligned}
R(k,l) &= \sum_i \frac{\lambda_u u_i}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l)\} > 0) \\
&\qquad \times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l)) \\
&+ \sum_i \frac{\lambda_e e_{i,0}}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l)\} > 0) \\
&\qquad \times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l)) \\
&+ \sum_i \sum_j \frac{\lambda_e e_{i,j}}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l)\} > 0) \\
&\qquad \times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l))
\end{aligned}
$$

With probability $P(k,l)$ the type l coworker leaves the firm:

$$P(k,l) = \lambda_e \frac{F_0}{F}\mathbb{I}(\widehat{V}_1(l) - (\widehat{V}_2(k,l) - \widehat{V}_1(k)) - \Pi_0 > 0) + \sum_i \lambda_e \frac{e_{i,0}}{F}\mathbb{I}(\widehat{V}_2(l,i) - (\widehat{V}_2(k,l) - \widehat{V}_1(k)) - \widehat{V}_1(i) > 0)$$
$$+ \sum_i \sum_j \lambda_e \frac{e_{i,j}}{2F}\mathbb{I}(\max\{\widehat{V}_2(l,i) + U(j) - (\widehat{V}_2(k,l) - \widehat{V}_1(k)) - \widehat{V}_2(i,j),$$
$$\widehat{V}_2(l,j) + U(i) - (\widehat{V}_2(k,l) - \widehat{V}_1(k)) - \widehat{V}_2(i,j)\} > 0)$$

With probability $Q(k,l)$ the type l coworker is replaced:

$$Q(k,l) = \sum_{i=1}^{N} \frac{\lambda_u u_i}{F}\mathbb{I}(\widehat{V}_2(k,i) + U(l) - U(i) > \widehat{V}_2(k,l)) + \sum_{i=1}^{N} \frac{\lambda_e e_{i,0}}{F}\mathbb{I}(\widehat{V}_2(k,i) + U(l_+) - (\widehat{V}_1(i) - \Pi_0) > \widehat{V}_2(k,l))$$
$$+ \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\lambda_e e_{i,j}}{F}\mathbb{I}(\widehat{V}_2(k,i) + U(l) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] > \widehat{V}_2(k,l))$$

4 Calibration

The model is calibrated so that one period is one month. The death rate, χ, is .003% per month, corresponding to a 30 year working life. The discount factor is set to $\beta = .9816$, which corresponds to a discount rate of 25% per annum.[3] The production function is given by $f(k,l) = (h_k^{\frac{1}{\rho}} + h_l^{\frac{1}{\rho}})^{\rho}$. ρ controls the degree of complementarity between types. We discuss the estimation of ρ in more detail below.

We assume there are $N = 7$ types. We assume productive abilities are evenly spaced between h_1 and h_7. We normalize h_1 to equal 1 and we estimate h_7 to match the p90/p10 wage ratio in the pooled 2000-2016 Current Population Survey Merged Outgoing Rotation Groups (henceforth, the CPS).

The bargaining weight of the unemployed workers, σ, is estimated to match the ratio of the average wage of job finders to the average wage of employed individuals in the CPS. We estimate home production ϕ to match the 16% decline in consumption after 6 months of unemployment reported by Browning and Crossley [2001].

For newborns, we assume that the type-distribution of newborns is a discrete truncated normal distribution over the productive abilities, $\{h_1, \ldots, h_7\}$, with mean m_{new} and variance v_{new}. We parameterize the mean and variance parameters to target (i) the mean wage of 22-24 year old individuals to the average wage of employed individuals in the CPS, and (ii) the p90 to p10 wage ratio for this subgroup in the CPS.

We normalize the measure of firms, F, to a unit mass, and the parameters that govern the unemployed job contact rate, the employed job contact rate, and the job destruction rate, $\{\lambda_u, \lambda_e, \delta\}$, are set to target the job finding rate, the job-to-job transition rate, and the unemployment rate as measured in the CPS.

We estimate the rate of human capital depreciation among the unemployed, α_u, to match the negative relationship between the job finding hazard and unemployment in the CPS. We estimate the rate of human capital appreciation for solo workers, α_0, to match the lifecycle profile of wages in the CPS.

The production complementarity parameter, ρ, and the learning parameter for teams, α_1, are estimated jointly to match three sets of moments: (i) job flows as function of the coworker-worker wage gap, (ii) the relationship between future wages and prior coworker

[3] This large discount rate allows us to avoid negative wages.

wage for EUE transitioners, and (iii) the share of wage variance that is between firms.

The first set of moments exploit the fact that learning has unique predictions about job-to-job flows and the relative wage of a worker and their coworker. Through the lens of the model, the worker's wage is a noisy proxy of an individual's type and the coworker's wage is a noisy proxy of the coworker's type. With strong production complementarity and no learning, the further a worker's wage is from their coworker's wage, the more likely they are to switch employers in search of a better match. With strong learning and no production complementarity, the further a worker's wage is from their coworker's wage, the *less* likely they are to switch employers because they are either learning or teaching their coworker.

The second set of moments we use to disentangle learning and sorting are based on the correlation between an individual's wage and their old coworkers' wages. To parse out sticky-wages, built-up rents, and other forces which drive a wedge between a worker's type and their true productivity, we focus on individuals who separate, experience a spell of unemployment, and then find a new job. We refer to this as an 'EUE' transition. As sorting improves, coworker wages become a stronger predictor of the worker's type, and thus coworker wages become better predictors of future individual wages. As learning improves, coworkers also become a stronger determinanet of future individual wages. The fact that both learning and sorting move this moment in the same direction allows us to pin down the level of sorting and learning.

And our last moment is the share of wage variance that is between firms. Since there are multiple workers at a firm in our model, our framework has a concept of within-firm and between-firm wage variance. As production complementarities increase, between-firm wage variance grows relative to within-firm wage variance. As learning becomes more important, within-firm wage variance share grows. Let $V(w_{i,t})$ denote total wage variance, let $w_{i,t}$ denote individual i's wage at date t, let $\bar{w}^f_{i,t}$ denote the average wage of firm f where individual i works at time t, and let $\bar{w}_t$ denote the economy wide average wage at date t. We decompose wage variance as follows,

$$V(w_{i,t}) = \underbrace{\frac{1}{N}\sum_{i=1}^{N}(w_{i,t} - \bar{w}^f_{i,t})^2}_{within} + \underbrace{\frac{1}{N}\sum_{i=1}^{N}(\bar{w}^f_{i,t} - \bar{w}_t)^2}_{between}$$

According to Spletzer [2014], 50.3% of cross-sectional earnings variance is across firms in the LEHD.

We describe our estimation of the job-to-job flow moments in Section 4.1 and the worker-coworker wage moments in 4.2.

4.1 Job mobility and the coworker-worker wage gap

To estimate job mobility as a function of coworker and worker wages, we use a 10% random sample from the Longitudinal Employer-Household Dynamics database between 2001 and 2008.[4] We restrict our sample to prime-age (24-65) males at single-unit firms with between 2 and 250 employees. We also require individuals in our sample to be employed throughout the previous year at the same primary employer.[5] We define a job-to-job transition from year t to $t+1$ if the individual switches primary employers without a spell of non-employment (earning \$1k or less in a quarter).

Let i index individuals and let t index years. Let $JJ_{i,t+1}$ be a dummy for a job-to-job transition between t and $t+1$. Let $w_{i,t}$ denote log real annual earnings of individual i (which we will refer to as an individual's wage in the LEHD) at his/her primary employer in year t.[6] Let $\bar{w}_{-i,t}$ denote the log real average annual wage of the coworkers of an individual at his/her primary employer in year t, excluding individual i. We refer to $\bar{w}_{-i,t} - w_{i,t}$) as the 'wage gap.' We estimate the following specifications separately for those above the average wage at their firm ($w_{i,t} > \bar{w}_{-i,t}$) and those below ($w_{i,t} < \bar{w}_{-i,t}$):

$$JJ_{i,t+1} = \beta_0 + \gamma_t + \beta_1 w_{i,t} + \beta_2(\bar{w}_{-i,t} - w_{i,t}) + \Gamma X_{i,t} + \epsilon_{i,t} \tag{11}$$

Our regression controls, $X_{i,t}$, include firm size, state dummies, 1-digit sic dummies, race dummies, education dummies, quadratics in age and tenure, as well as year fixed effects. All standard errors are clustered at the SEIN level.

Table 1 provides summary statistics in our main sample.[7] The annual job-to-job transition rate is 3.6% per annum. The average age of an individual in our sample is 41, and they

[4]The LEHD database is a matched employer-employee dataset that covers 95% of U.S. private sector jobs. The LEHD includes data on earnings, worker demographic characteristics, firm identifiers, and firm characteristics. Our data covers 2001 through 2008 for 11 states: California, Illinois, Indiana, Maryland, Nevada, New Jersey, Oregon, Rhode Island, Texas, Virginia, and Washington.

[5]We count individuals as employed if they earn \$1k or more from an employer in a given quarter. We count individuals as non-employed if they earn less than \$1k in a given quarter. The primary employer is the employer that pays the worker the most.

[6]All variables which are not bounded above are winsorized at the 1% level. Nominal variables are deflated using the CPI.

[7]For disclosure purposes, we were required to round the sample size to the nearest thousand.

have, on average, one-and-a-half years of college education. Average tenure at their previous employer is approximately 4 years. They earn roughly \$61k per annum, and their coworkers earned approximately \$42k per annum. The discrepancy in earnings is due to the fact that we have tenure requirements on the individuals in our sample, but not on their coworkers.

Table 1: Sample of Employed Workers (Source: LEHD)

Variable	**Mean**	**SD**
Job to Job Transition Rate (Annual)	3.60%	0.186
Age	41.95	10.39
Imputed Years of Education	13.54	2.904
Tenure	4.137	2.994
Log Average Firm Wage, Excluding Individual i ($\bar{w}_{-i,t}$)	10.27	1.27
Log Individual Wage ($w_{i,t}$)	10.77	0.721
Average Firm Wage, Excluding Individual i	41,670	29,320
Individual Wage	60,920	46,020
Abs. Value of Log Avg. Firm Wage Minus Log Individual Wage	0.624	0.604
Number of Observations	9,648,000	

Table 2 reports the results for the estimated equation (11). Columns (1) through (4) use the sample of individuals who earn less than their coworkers. Columns (5) through (8) use the sample of individuals who earn more than their coworkers. Column (1) reveals that for individuals below the average wage at their employer (excluding the individual, 'ex-i'), the larger the gap between the individual and their coworkers (henceforth, the wage gap), the less likely the individual will transition between employers. Column (2) reruns the same specification with controls. The coefficient of -.0141 on the wage gap ($\bar{w}_{-i,t} - w_{i,t}$) can be interpreted as follows: if the coworkers of agent i earn 10 percent more, agent i is .14% less likely to transition between employers next year. Relative to the sample average job-to-job transition rate of 3.6% per annum, this represents a 5.3% reduction. Column (3) adds in additional lags of the worker's earnings to control for the worker type. Column (4) separates the wage gap into its components. Column (4) reveals that for workers below the mean wage of their employer, they are *less* likely to transition if their coworkers earn more. Columns (5) through (8) reveal that the opposite is true for individuals who earn above the mean. In these instances, the wage gap, ($\bar{w}_{-i,t} - w_{i,t}$), takes on negative values. Thus, they are less likely to move if they are earning more than their coworkers.

Table 2: Dependent Variable is Job-to-Job Transition Rate from t to t+1 (Source: LEHD)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Dependent variable	J to J Transition, t to t+1	J to J Transition, t to t+1	J to J Transition, t to t+1	J to J Transition, t to t+1	J to J Transition, t to t+1	J to J Transition, t to t+1	J to J Transition, t to t+1	J to J Transition, t to t+1
Sample	Below Coworker	Below Coworker	Below Coworker	Below Coworker	Above Coworker	Above Coworker	Above Coworker	Above Coworker
Wage Gap ($\bar{w}_{-i,t} - w_{i,t}$)	-0.0121***	-0.0141***	-0.0141***		0.000146	0.00186***	0.00172***	
	(0.000674)	(0.000703)	(0.000704)		(0.000286)	(0.000205)	(0.000205)	
Log Individual Wage ($w_{i,t}$)	0.00243***	0.00437***	0.00404***	0.0162***	-0.00760***	-0.00115***	-0.00307***	-0.00586***
	(0.000658)	(0.000615)	(0.000614)	(0.000409)	(0.000391)	(0.000316)	(0.000313)	(0.000271)
Size of Employer		-5.66e-07***	-5.67e-07***	-5.67e-07***		-5.78e-07***	-5.77e-07***	-5.82e-07***
		(9.47e-08)	(9.48e-08)	(9.48e-08)		(1.23e-07)	(1.24e-07)	(1.25e-07)
1Yr Lag Log Indiv. Wage ($w_{i,t-1}$)			0.000703***	0.000703***			0.00172***	0.00172***
			(5.56e-05)	(5.56e-05)			(4.82e-05)	(4.82e-05)
2Yr Lag Log Indiv. Wage ($w_{i,t-2}$)			-0.000369***	-0.000369***			0.000149***	0.000128***
			(4.46e-05)	(4.46e-05)			(3.71e-05)	(3.72e-05)
Log Avg. Firm Wage, ex-i ($\bar{w}_{-i,t}$)				-0.0122***				0.00281***
				(0.000622)				(6.92e-05)
Controls	N	Y	Y	Y	N	Y	Y	Y
R2	0.000750	0.00805	0.00812	0.00808	0.000755	0.0108	0.0110	0.0113
Round N	2,441,000	2,441,000	2,441,000	2,441,000	7,203,000	7,203,000	7,203,000	7,203,000

*Notes: SE clustered at SEIN level. *** p<0.01, ** p<0.05, * p<0.1. Controls include: firm size, state dummies, 1-digit sic dummies, race dummies, education dummies, quadratics in age and tenure, as well as year fixed effects.*

4.2 Wages and past coworkers

We measure the correlation of future wages and past coworker wages by focusing on individuals who transition between employers through a spell of unemployment ('EUE' transitions). Through the lens of our model, this type of job transition parses out wage stickiness and built-up rents, and so wages out of unemployment are a better proxy of the worker's human capital. The relationship between prior coworker wages and future individual wages reflect both sorting and learning. As sorting increases, coworker wages become a better proxy of worker type and therefore become more correlated with future individual wages. As learning increases, coworker wages become more correlated with future individual wages via knowledge diffusion.

We identify EUE transitions in the data as those who have at least 1 year of tenure in year t at their primary employer, spend at least one quarter non-employed in year t+1 (defined as earning less than $1k in the quarter), and then obtain a job at a different primary employer in year t+2. Let i index people who make an EUE transition and let t index years. The dependent variable of interest is individual i's log annual earnings in year t+2, $w_{i,t+2}$. We regress that on their own log wage prior to the unemployment spell, $w_{i,t}$, as well as the log real average annual wage of the coworkers, $\bar{w}_{-i,t}$. We estimate the following specifications:

$$w_{i,t+2} = \beta_0 + \gamma_t + \beta_1 w_{i,t} + \beta_2 \bar{w}_{-i,t} + \Gamma X_{i,t} + \epsilon_{i,t} \tag{12}$$

Our regression controls, $X_{i,t}$, include firm size, state dummies, 1-digit sic dummies, race dummies, education dummies, quadratics in age and tenure, as well as year fixed effects.

Table 3 provides summary statistics for our sample of EUE job transitioners prior to their unemployment spell (date t). The average age of an individual in our sample is 38, and they have, on average, one year of college education. Average tenure at their previous employer is approximately 3 years. They earn roughly $39k per annum, and their coworkers earn approximately $30k per annum. As before, the discrepancy in earnings is due to the fact that we have a tenure requirement on the individuals in our sample, but not on their coworkers.

Table 4 includes the results from estimating equation 12 for our sample of EUE transitioners. Columns (1) through (3) restrict the sample to those who earn less than their coworkers at their previous firm in year t. Column (1) includes no controls. The point estimate implies that across all workers who experienced an EUE transition between t and t+2, a 10%

Table 3: Summary Statistics of EUE Transitioners (Source: LEHD)

Variable	**Mean**	**SD**
Age	38.2	9.5
Imputed Years of Education	13.0	2.9
Tenure	2.9	2.3
Log Average Firm Wage, Excluding Individual i ($\bar{w}_{-i,t}$)	10.0	0.8
Log Individual Wage ($w_{i,t}$)	10.4	0.7
Average Firm Wage, Excluding Individual i	30,436	21,743
Individual Wage	39,365	28,267
Number of observations	55000	

greater coworker wage prior to the unemployment spell is associated with a 1.34% greater individual wage coming out of the unemployment spell. Column (2) includes controls, and the point estimate implies that a 10% greater coworker wage prior to the unemployment spell is associated with a 1.25% greater individual wage coming out of the unemployment spell. Column (3) adds additional lags of the worker's wage as a better proxy for the worker's type, and the point estimate does not change. Columns (4) through (6) restrict the sample to those who earn more than their coworkers at their previous firm in year t, and we see a much more muted relationship. The point estimates in Column (5) imply that a 10% greater coworker wage prior to the unemployment spell is associated with a .16% greater individual wage coming out of the unemployment spell.

What is consistent across all specifications is that prior firm size is negatively related to future wages. We view this as suggestive evidence that our results are not being driven by network effects.

Table 4: Future Individual wages in year t+2 as a function of coworker wages in year t, EUE Transitioners (Source: LEHD)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Individual Wage, t+2	Individual Wage, t+2	Individual Wage, t+2	Individual Wage, t+2	Individual Wage, t+2	Individual Wage, t+2
Sample	Below Coworker	Below Coworker	Below Coworker	Above Coworker	Above Coworker	Above Coworker
Log Average Firm Wage, Excluding Individual ($\bar{w}_{-i,t}$)	0.134***	0.125***	0.122***	0.00959*	0.0159***	0.0150***
	(0.0127)	(0.0127)	(0.0127)	(0.00514)	(0.00519)	(0.00518)
Log Individual Wage ($w_{i,t}$)	0.299***	0.289***	0.284***	0.587***	0.557***	0.545***
	(0.0118)	(0.0118)	(0.0119)	(0.00664)	(0.00691)	(0.00710)
Size of Employer		-0.000145**	-0.000151**		-0.000282***	-0.000280***
		(6.34e-05)	(6.34e-05)		(4.53e-05)	(4.53e-05)
1Yr Lag Log Individual Wage ($w_{i,t-1}$)			0.0100***			0.0146***
			(0.00231)			(0.00196)
2Yr Lag Log Individual Wage ($w_{i,t-2}$)			-2.83e-05			-0.000518
			(0.00151)			(0.00106)
Controls	N	Y	Y	N	Y	Y
R-squared	0.159	0.194	0.195	0.324	0.341	0.342
Observations	18000	18000	18000	38000	38000	38000

*Notes: SE clustered at SEIN level. *** p<0.01, ** p<0.05, * p<0.1. Controls include: firm size, state dummies, 1-digit sic dummies, race dummies, education dummies, quadratics in age and tenure, as well as year fixed effects.*

Appendix D includes additional robustness checks and reruns our specifications on Brazilian data.[8] We conduct several exercises: (i) we specify coworker wages within a firm and 2-digit occupation, (ii) we follow workers over a 4 year horizon and show that the effects remain stable, (iii) we isolate industry switchers and we find point estimates that are similar for industry switchers, (iv) we compare our results to other regressions in the peer-effects literature including Nix [2015], and (v) we look at different summary measures of the firm's workforce, including the different deciles of the coworker wage distribution.

4.3 Model Fit

Table 5 summarizes the model's fit relative to the targeted moments. The model is over-identified, and manages to perform well at matching almost all of the targeted moments. While all of the parameters were estimated jointly, the first 7 rows include the moments that discipline the learning and sorting parameters most directly, and the model does well at replicating those 7 moments. The model requires a large degree worker complementarity, $\rho = 1.96$, to match the share of total wage variance that is between-firms. The model also requires reasonably strong learning parameters to match the correlation of job-to-job mobility and the coworker-worker wage gap as well as match the correlation of lagged worker wages and future wages for EUE transitioners. The learning parameter $\alpha_1 = .035$ governs the rate at which less-skilled workers learn from more skilled coworkers. Our estimates imply that if the lowest skilled worker is paired with the highest skilled worker, the lowest skilled worker will move up to the next rung of the human capital ladder once every 2 years.

Our estimates for the rate at which solo employed workers learn, $\alpha_0 = .007$, implies that once every 12 years, the worker will move up to the next rung of the human capital ladder. The rate of dislearning among the unemployed, $\alpha_u = .15$, is much faster. The unemployed fall one rung on the human capital ladder once every 6 months. While this parameter affects several moments, it is primarily identified from the job finding hazard. Even with such a large dislearning parameter, the model struggles to generate the observed decline in job finding rates.

The model does well at matching flows into and out of unemployment, as well as between jobs. The model is also capable of generating the levels of wage dispersion observed in the data, and it is also capable of matching relative wages of the young. However, it understates the wages of job finders.

[8] All RAIS results contained in this paper were run on IPEA servers in accordance with MTE guidelines.

Table 5: Calibration: Model Moments vs. Data Moments (Source: CPS and LEHD)

Param.	Value	Description	Model	Data	Moment Description	Source
ρ	1.966	Productive complementarity	0.482	0.503	Between firm wage variance share	Splezter (2017)
α_1	0.035	Coworker Learning	-0.018	-0.014	Elast. of EE rate to Coworker-Worker wage gap, below firm wage	LEHD (2002-2007)
			0.001	0.002	Elast. of EE rate to Coworker-Worker wage gap, above firm wage	LEHD (2002-2007)
			0.120	0.125	Elast. of own wage to prior coworker wage, EUE below firm wage	LEHD (2002-2007)
			0.018	0.016	Elast. of own wage to prior coworker wage, EUE above firm wage	LEHD (2002-2007)
			0.100	0.289	Elast. of own wage to prior own wage, EUE below firm wage	LEHD (2002-2007)
			0.704	0.557	Elast. of own wage to prior own wage, EUE above firm wage	LEHD (2002-2007)
α_0	0.007	Solo Employed	1.741	1.717	Mean wage 54/Mean wage 22	CPS (2000-2016)
α_u	0.150	Unemployed	0.934	0.654	Ratio of Job Finding Rate at 3mo Duration to 1mo Duration	CPS (2000-2016)
h_N	2.412	Top productive ability	4.093	4.546	p90/p10 wage ratio	CPS (2000-2016)
v_{new}	0.338	Variance of newborn types	2.639	2.975	22-24yo p90/p10 wage ratio	CPS (2000-2016)
m_{new}	0.745	Mean newborn type	0.591	0.620	22-24yo mean wage to avg wage ratio	CPS (2000-2016)
$1-\sigma$	0.118	Bargaining weight of firm	0.486	0.795	Wage of job finders to avg wage ratio	CPS (2000-2016)
λ_u	0.293	Unemployed Job Contact Rate	0.197	0.212	UE Flow Rate, Monthly	CPS (2000-2016)
λ_e	0.352	Employed Job Contact Rate	0.015	0.017	EE Flow Rate, Monthly	CPS (2000-2016)
δ	0.007	Layoff Rate	0.011	0.010	EU Flow Rate, Monthly	CPS (2000-2016)
ϕ	1.135	Home production	0.900	0.840	Consumption drop after 6 mo of unemployment	Browning and Crossley (2001)

4.4 Steady State Distribution of Workers

Figure 2 plots the estimated newborn type distribution and the population type distribution. Roughly 60% of workers are born as the lowest type, but via learning, as solo employed or team workers, the fraction of the population that is the lowest type drops to 21.1%. Almost no workers are born as the highest type, however, through learning, 18.3% of the population reaches the highest type.

Figure 4 is a contour map of the team distribution. The distribution features many high-low pairings (schools), many (1,1) teams, and many (7,7) teams. Figure 5 is the corresponding joint pdf. Since there are disproportionately more low-type workers in the economy, most pairs are formed between low-type workers. Roughly 4.4% of all teams are composed solely of the lowest type, i.e. (1,1) teams. To understand these sorting patterns, Figure 9 illustrates the hiring policy functions for teams. Type 1 agents are not hired out of unemployment by any team. Type 2 agents displace type 1 agents from teams. Type 3 agents displace type 1 and type 2 agents from teams. A consequence of learning is that the highest-productivity team, (7,7), may actually be dissolved in order to pair a lower type agent with one of the type 7s. This is the case with type 3 agents who are hired by (7,7) teams. The relative strength of learning and sorting determines how likely a lower-type agent will split up a pairing of the highest types. Type 4 agents similarly displace type 1, 2, and 3 agents from teams. Due to learning, type 4 agents will also break up (6,6), (6,7) and (7,7) teams.

Figure 3 illustrates the job flow patterns that arise in steady state. Lowest type individuals have the hardest time obtaining employment, are the most likely to be laid off, and are also the most likely to switch employers. The highest type individuals are the exact opposite. To provide more detail on which teams are generating job flows, Figure 8 is a contour map of the team separation rate. Similar to Figure 9, the greatest separation rates occur among all teams that have a type 1 worker. These teams are split whenever their firm contacts a non-type-1 worker. Along the diagonal, separation rates are high, which simply reflects the fact that learning is a very strong force in the model. Learning is strong enough to disband (6,6) and (7,7) teams.

Figure 6 plots the distribution of workers across teams. Figure 6 reveals that the lowest type workers are disproportionately unemployed and working by themselves. As the skill level rises, the fraction of workers unemployed drops, and the fraction of workers on high-skill teams rises non-monotonically. The highest type 7 individuals are much more likely to be paired with a type 5 or type 6 than a fellow type 7. Again, this simply reflects the fact that there are large returns to learning.

To formalize these graphical measures of sorting, Table 6 shows that the Spearman rank correlation coefficient among worker types is .477, which is consistent with strong worker complementarity estimates from studies which use different frameworks (e.g. Bonhomme et al. [2014] and Borovickova and Shimer [2017]). While there is strong sorting, there are still a large number of workers in 'schools.' If we classify a school as a pairing of two workers who differ by 1 type or more, then approximately 34% of the population is in a school. Figure 7 plots the measure of schools, and nearly 50% of type 1 individuals are being 'mentored' by a higher-ranking team member. As we discuss in Section 6, the wage setting mechanism disincentivizes firms from educating workers by generating a gap between the private and social returns to learning.

Figure 2: Newborn Type Distribution

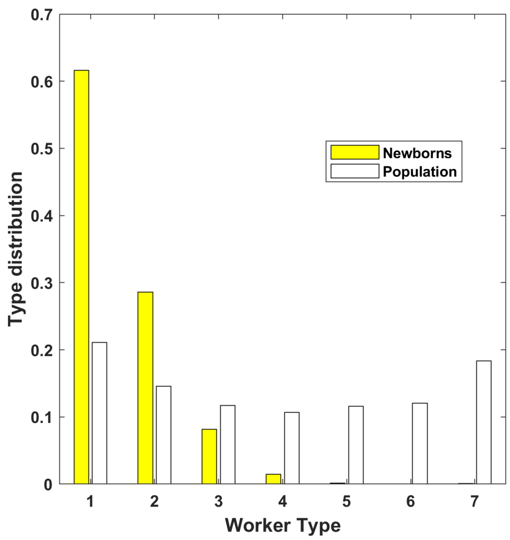


Figure 3: Transition Rates by Type

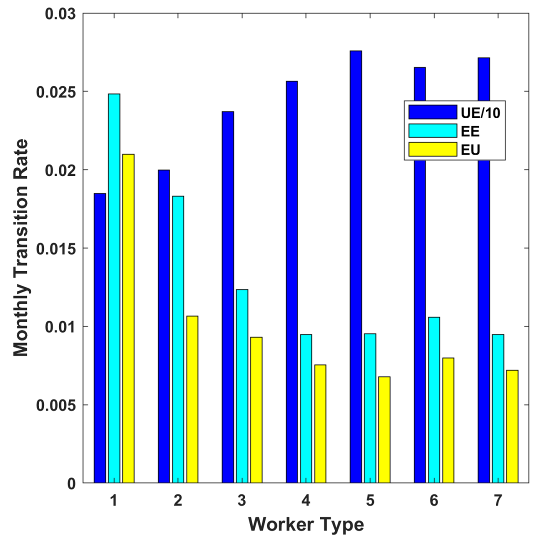


Table 6: Sorting Moments

Sorting Moments	**Value**
Spearman rank correlation coefficient	0.39707
Fraction in school (Coworker type higher than own type)	0.34456

Figure 4: Contour map of team distribution

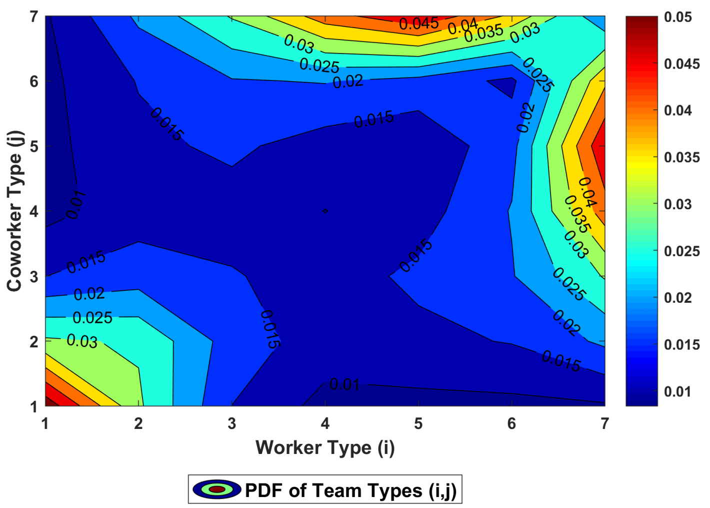


Figure 5: PDF of team distribution

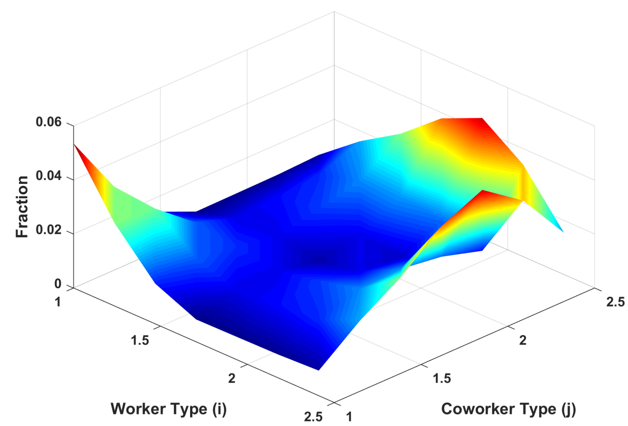


Figure 6: Distribution of Workers Across Teams

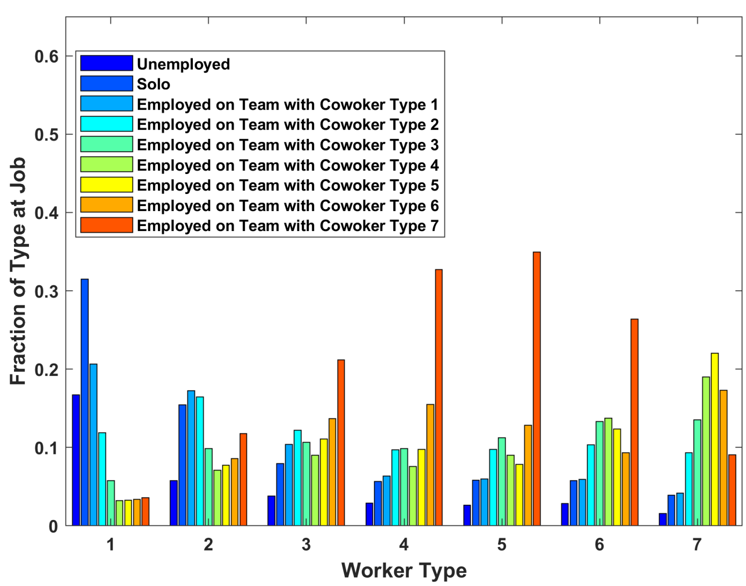


Figure 7: Fraction of types in 'school'

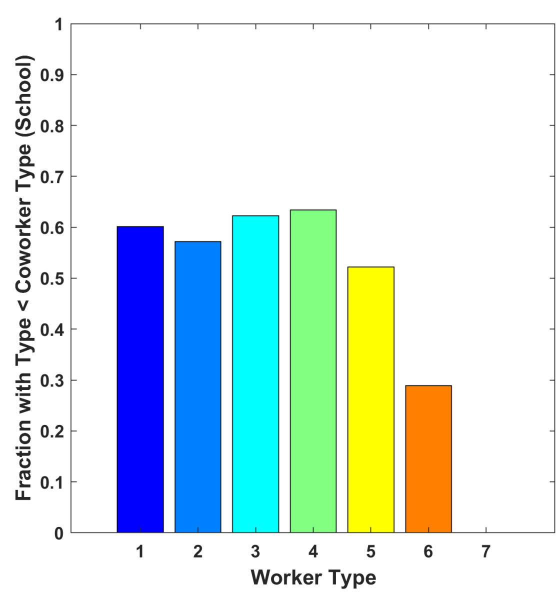


Figure 8: Separation Rate by Teams

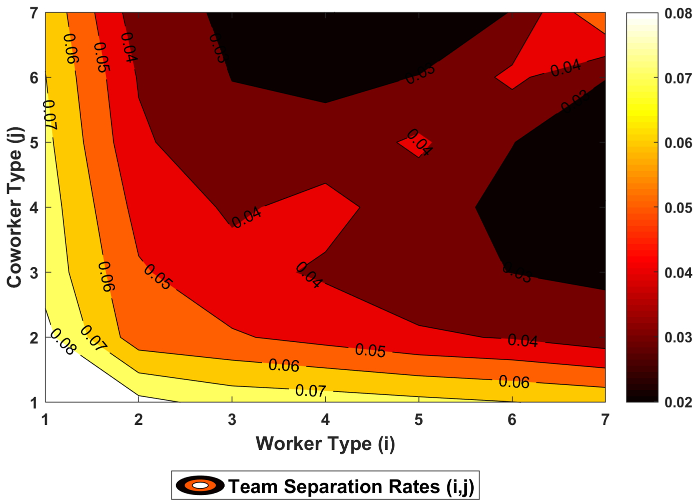
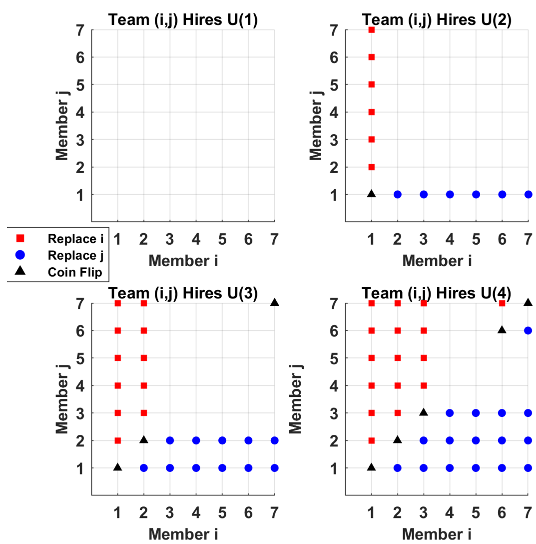


Figure 9: Team hiring policy function



5 Sources of Knowledge Diffusion

In this section, we use the calibrated model to understand the determinants of knowledge diffusion. Table 7 includes results from three counterfactuals, (i) inhibiting all forms of coworker learning ($\alpha_1 = 0$), but still allowing for worker mobility ($\lambda_e > 0$), (ii) inhibiting worker mobility ($\lambda_e = 0$), but still allowing for coworker learning ($\alpha_1 > 0$), and (iii) shutting down both worker mobility and coworker learning ($\lambda_e = 0,\ \alpha_1 > 0$).

Column (1) of Table 7 includes output, sorting, and key moments related to learning for the baseline U.S. economy. Column (2) of Table 7 shows that if we inhibit all forms of coworker learning ($\alpha_1 = 0$), but still allowing for worker mobility ($\lambda_e > 0$), output drops by 16.3 percentage points. This implies that coworker learning accounts for 1/6 of U.S. GDP.

When coworker learning is eliminated in Column (2), inequality, measured by the p90 to p10 wage ratio, decreases enormously as the fraction of individuals of the highest type drops from 18% to 6%. The fraction of high types plummets because high types are no longer disseminating their knowledge and training lower types. As a consequence, sorting, measured by the Spearman rank correlation coefficient, increases by 25% when learning is eliminated, going from .4 to .5. If one minus the spearman rank is interpreted as an index of 'mismatch,' our estimates imply that learning accounts for at least 16% (=.1/(1-.4)) of measured 'mismatch.' This is not to say that learning is actually generating any mismatch. We will demonstrate in the Planner's solution that the U.S. allocation of workers to firms is close to efficient, given the estimated search frictions and learning parameters.

The final three rows of Column (2) include three moments we use to measure sorting and knowledge diffusion. The counterfactual elimination of coworker learning in Column (2) implies both (i) an improvement of sorting (an equilibrium result), and (ii) a decline of learning (by construction, $\alpha_1 = 0$). Regarding the first moment, increased sorting and decreased learning both serve to augment the share of between-firm wage variance, which increases by 4 percentage points. The second moment is the correlation of prior coworker wages and future individual wages among EUE transitioners. Increased sorting and decreased learning push this moment in opposite directions. Better sorting makes prior coworker wages a better predictor of future wages, whereas reduced learning trends to decrease the importance of prior coworker wages for individual wages. Ultimately, increased sorting dominates and prior coworker wages become better predictors of future wages of EUE transitioners. Lastly, low-wage individuals who are paired with high wage individuals are now more likely to switch employers since they are not learning. This drives the increased sorting.

Column (3) shows that relative to the baseline economy, disallowing on-the-job mobility lowers output by 15.7 percentage points. This implies that mobility is also accountable for approximately 1/6 of U.S. GDP. With a very strong degree of worker complementarity, much of the losses from worker mobility are generated by weaker sorting. The Spearman rank correlation coefficient drops by close to 25%. The share of between firm variance increases because there are simply twice as many solo workers (with only one wage at the firm), increasing from 12% to 22%.

Lastly, Column (4) illustrates the turning off both learning and worker mobility would reduce output by 29%. Despite strong interactions between learning and mobility in the theory, the impact of jointly eliminating peer effects and mobility is roughly additive. These three counterfactuals imply that learning and worker mobility both account for 1/3 of U.S. GDP and are approximately equally important determinants of U.S. output.

Table 7: Relative Importance of Mobility vs. Learning for Knowledge Diffusion and Output

	(1) Baseline	**(2)** No Learning ($\alpha_1 = 0$)	**(3)** No Mobility ($\lambda_e = 0$)	**(4)** No Mobility and No Learning
Output	5.81	4.86	4.90	4.08
Percent output loss relative to baseline		-16.3%	-15.7%	-29.8%
Wage Dispersion	4.09	3.35	1.65	1.26
Spearman rank correlation coefficient	0.40	0.50	0.31	0.28
Measure of Solo Employed Workers	0.13	0.12	0.22	0.08
Share of wage variance between firms	0.48	0.52	0.61	0.57
Elasticity of own wage with respect to past coworker wage, below coworker	0.12	0.20	0.19	0.30
Elasticity of Job to Job transition rate with respect to wage gap (Coworker wage minus individual), below coworker	-0.02	-0.01	0.00	0.00

5.1 Role of Skill Complementarity for Learning and School Formation

Figure 10 illustrates the team distribution under the assumption of no production complementarities between workers, i.e. $\rho = 1$. The value of a high-type worker is now greater if that high-type worker is matched with a low-type worker; the high-type worker's production is undistorted by their partner's type, and the the high-type worker will educate the low-type worker. This discrepancy of types within a team is what we call a school. Once the low-type worker learns, the low-type worker will leave and form their own school.

Figure 11 illustrates the percentage of team-workers by type in schools under the baseline calibration.[9] Low-type workers are predominantly in schools, and higher type workers are still occasionally matched with an even higher type (i.e. a 'mentor').

Figure 13 illustrates team hiring patterns with perfect substitutes. Type 1's are still not hired by any existing team. However, type 2 agents are capable of splitting a (7,7) team. Learning is the only reason teams are formed, and so it maximizes joint surplus to pair the lowest types with the highest types. Type 3 agents are now capable of dissolving (5,6), (6,6), (6,7) and (7,7) teams. Figure 12 illustrates the separation rate by worker types. Along the diagonal, mobility is highest as same-type teams do not benefit from learning. Thus, with perfect substitutability, dissorting is privately optimal.

This exercise highlights an important point, which is that depending on the strength of learning, standard measures of mismatch under supermodular production functions, e.g. Spearman rank correlations on true types, may actually mistakenly identify an efficient allocation of workers to 'mentors' and 'schools' as misallocation.

Figure 10: Contour map of team distribution under perfect substitutes, $\rho = 1$.

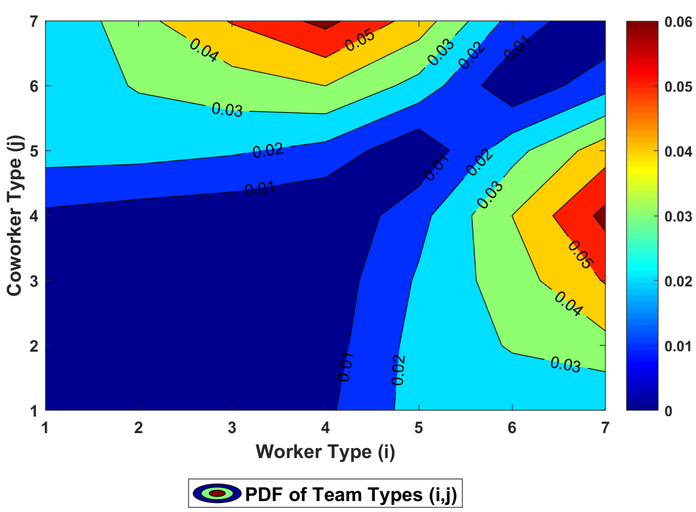
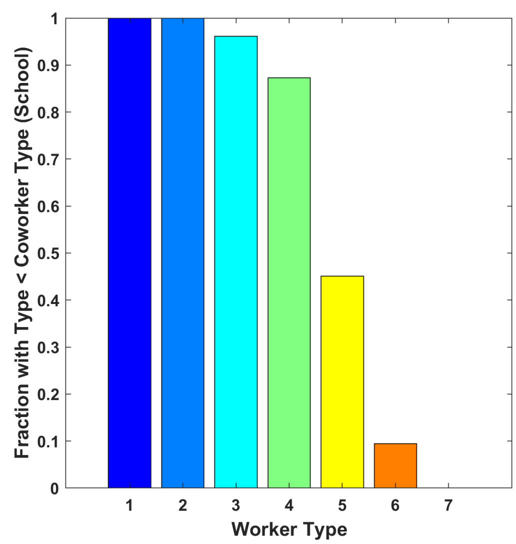


Figure 11: Perfect substitutes ($\rho = 1$), fraction of schools.



[9]This graph is constructed by summing the measure of type i individuals in schools, and then dividing by the total measure of type i individuals on teams.

Figure 12: Contour map of separation rates with perfect substitutes, $\rho = 1$.

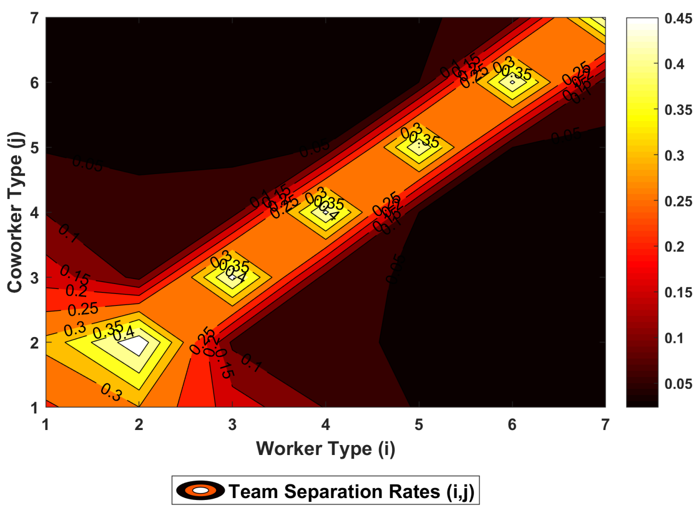
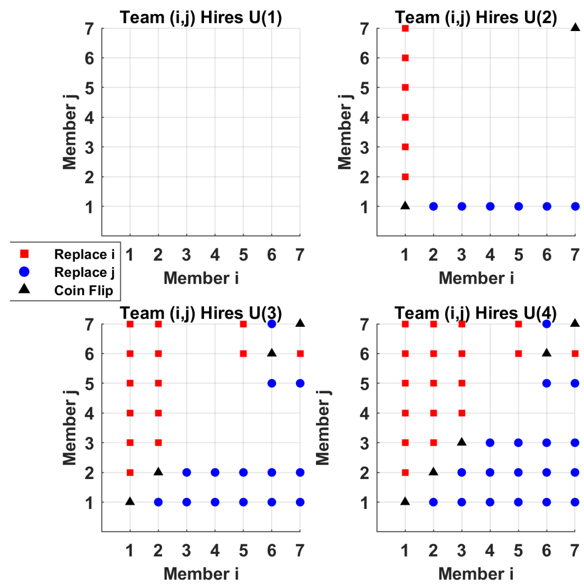


Figure 13: Team hiring policy function under perfect substitutes, $\rho = 1$.



6 Planner's Problem

In this section we define the planner's problem. The main source of inefficiency in the decentralized version of the economy is the fact that teams are not fully compensated for educating workers due to the wage bargaining protocol (e.g. Cahuc et al. [2006]). There is a social value to altering a workers' type, but due to the fact that the worker can switch jobs, and the team is only compensated for a fraction of what the worker will produce in future matches via the bargaining protocol, not enough education takes place from a planner's perspective.

The planner collects home production from the unemployed, as well as production from single worker firms and two worker firms. In the planner's problem, we use n to denote the measure of firms. n_0 is the measure of firms with no employees, n_1 is the measure of firms with one employee, and n_2 is the measure of firms with two employees. The planner's problem is to maximize the following objective function,

$$P(u, e_{i,0}, e_{i,j}) = \max_{\hat{u}, \hat{e}_{i,0}, \hat{e}_{i,j}} \sum_k b(k)u_k + \sum_k f(k,0)e_{k,0} + \sum_{k,j} \frac{f(k,j)e_{k,j}}{2} + \beta P(\hat{u}, \hat{e}_{i,0}, \hat{e}_{i,j})$$

subject to the law of motion for distributions given by equations (6) to (25) (evaluated using the planner's value functions and measures of firms), and taking the mass of firms as given

$$n_0 = n - n_1 - n_2 \tag{13}$$

The general strategy is to take partial derivatives of the planner's problem, taking into account that when the planner adds an additional single worker or two worker firm, the stock of vacant firms (which is a residual) declines linearly according to equation (13). This is the opportunity cost to the planner of match formation. We define the value of an additional single worker firm to the planner as $V_1^P(i)$, where relative to the decentralized version, $V_1^P(i)$ is redefined to reflect the fact that the planner will lose a vacant firm and the contacts that the vacant firm would generate. Conditional on a given set of value functions, the equations that govern the flow of workers across states is the same in both the decentralized economy and the planner's problem. What differs is the value functions themselves. Appendix B includes the value functions for the planner.

6.1 Planners Allocation vs. Decentralized Allocation

Table 8 shows that the planner's allocation generates roughly 4% more output relative to the decentralized economy. The main source of gains come from the fact that the planner generates more schools, and thus generates a population with greater human capital. The fraction of individuals in school is 38% in the planner's allocation, and 34% in the baseline economy. As a result, Figure 15 illustrates that the planner generates significantly more high-skilled agents. In the decentralized economy 18.3% of the population is the highest type, whereas in the planner's economy, 21.2% of the population is the highest type.

To better understand where the output gain originates, Figure 14 plots the difference in the joint team distribution between the planner's allocation and the decentralized allocation. Positive numbers indicate that the planner has increased the number of teams in a given region. Figure 14 shows that the planner makes more high-end schools at the expense of low-level sorting. The planner pairs more type-6 and type-7 workers with lower-type workers, breaks up low-skill (1,1) teams and high-skill (7,7) teams, and generates more (1,7), (2,7), (3,7), and (4,7) schools. The planner's desire to create more schools reflects the inefficiencies present in the decentralized economy. Since there is a social value to altering a workers' type not captured by the bargaining mechanism in the decentralized economy, not enough schools are generated from the planner's perspective. The net effect is a 4% output gain from the planner's allocation.

Table 8: Social Planner vs. Baseline Decentralized Economy

	Decentralized	Planner
Output	5.81	6.03
Percent Gain Relative to Decentralized Output		3.86%
Measure of schools	0.34	0.38

Figure 14: Contour map of the *difference* in team distribution across planner and decentralized economies: Social planner's PDF of teams minus decentralized PDF of teams.

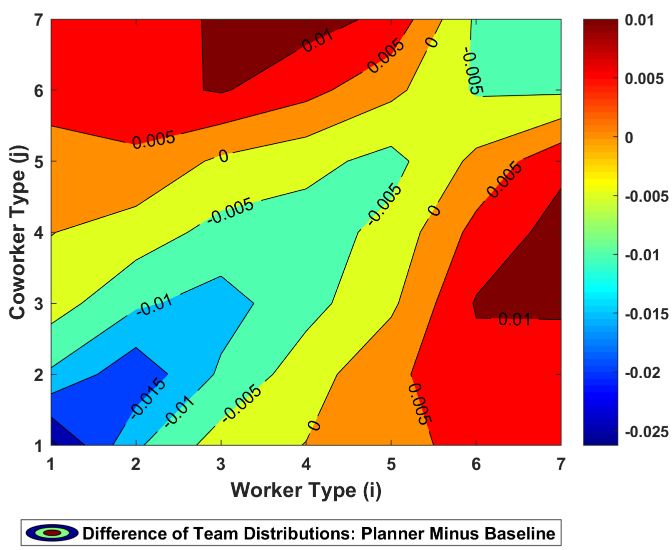
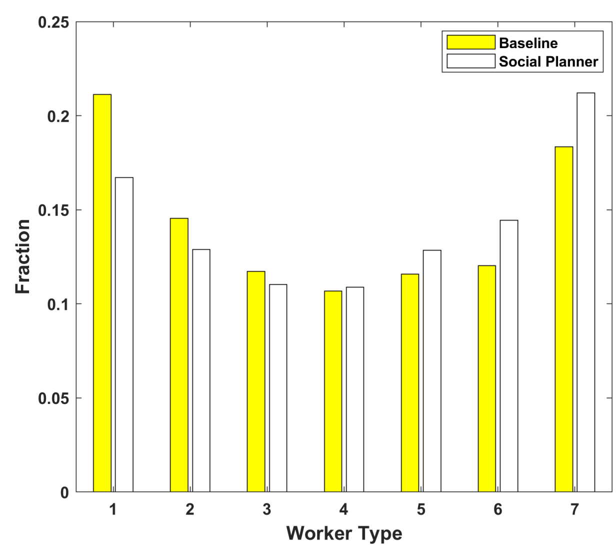


Figure 15: Type distribution of Social Planner vs. Baseline Decentralized Economy.



7 Conclusion

What role do peer effects, job transitions, and worker complementarities play in the diffusion of knowledge throughout the US economy? To answer this question, we develop a theory of teams to measure the way knowledge diffuses within and across firms. Our theoretic contribution is to extend the frictional sorting framework of Shimer and Smith [2000] to allow for workers within a firm to influence each other's knowledge. Workers can search on-the-job and leave their team to start a new team, carrying some of their knowledge with them. In contrast to standard sorting models, a firm's type is no longer exogenous; it is coworker human capital.

Our quantitative contribution is to estimate peer effects and worker complementarity using the structure of our model and data from the Longitudinal Employer-Household Dynamics (LEHD) database. We indirectly infer the model's sorting and learning parameters from three moments in the LEHD, that, when viewed through the lens of our model, directly inform sorting and learning. Using this methodology, we estimate strong skill complementarity and strong peer effects.

We then use the estimated model to measure the relative importance of peer effects and worker mobility for U.S. output. We do so by running counterfactuals that shutdown learning from coworkers and shutdown job-to-job mobility. Our main finding is that peer-effects explain roughly 1/6 of GDP in the U.S. while worker mobility explains roughly 1/6 of GDP in the U.S. Thus worker mobility and learning are equally important determinants of output in the U.S.

Our estimates also imply that learning accounts for at least 16% of measured 'mismatch,' which we define be a pairing of a high-type and low-type worker. However, this is not to say that learning is actually causing any true mismatch; in fact, the planner's problem features even more pairings of high and low-type workers (what we also call *schools*). The planner's solution features 38% of individuals in schools, 4% more than in the decentralized economy, and the planner also manages to produce 4% more than the baseline economy. With learning, standard measures of sorting are misleading because they do not take into account the dynamic benefits of coworker interaction.

A Expressions for Distributions of Unemployed and Employed Workers

Similar to the Bellman equations, we split the distribution of type k unemployed workers within the period into four subperiods: (1) the initial distribution u_k^-, (2) the distribution after the learning stage u_k^{learn}, (2) the distribution after births and deaths u_k (the start of the search stage), (3) the distribution after search outcomes are realized u_k^{search}, and (4) the distribution after dismissals occur, u_k^+. u_k^+ becomes the initial distribution for the next period, u_k^-. An equilibrium is a fixed point of this mapping.

A.1 Learning

After dismissal, production occurs, and then types are realized according to the function g, which takes as inputs the agent's type and the coworker's type (if an agent is unemployed, the coworker type is denoted 'u', if an agent is employed alone, the coworker type is denoted '0', and if an agent is employed with a coworker of type j, then the coworker type is simply 'j'). Let u_j^- denote the mass of unemployed agents at the start of the period. Let $e_{k_0}^-$ and $e_{i,j}^-$ be similarly defined. The evolution of types for the unemployed is therefore given by,

$$u_k^{learn} = u_k^- + \sum_{j \neq k} u_j^- g_u(k \mid j) - \sum_{j \neq k} u_k^- g_u(j \mid k) \tag{14}$$

The evolution of types for agents at single worker firms is therefore given by,

$$e_{k,0}^{learn} = e_{k,0}^- + \sum_{j \neq k} e_{j,0}^- g_0(k \mid j) - \sum_{j \neq k} e_{k,0}^- g_0(j \mid k) \tag{15}$$

The evolution of types for workers at two worker firms is therefore given by,

$$e_{k,l}^{learn} = e_{k,l}^- + \sum_{(i,j) \neq (k,l)} e_{i,j}^- g_j(k \mid i) g_i(l \mid j) - \sum_{(i,j) \neq (k,l)} e_{k,l}^- g_l(i \mid k) g_k(j \mid l) \tag{16}$$

A.2 Birth and Deaths

$$u_k = (1-\chi)u_k^{learn} + \chi\Big(\sum_k u_k^{learn} + \sum_k e_{k,0}^{learn} + \sum_k \sum_l e_{k,l}^{learn}\Big)\Gamma(k) \tag{17}$$

$$e_{k,0} = (1-\chi)e_{k,0}^{learn} + \chi \sum_l e_{k,l}^{learn} \tag{18}$$

$$e_{k,l} = (1-2\chi)e_{k,l}^{learn} \tag{19}$$

A.3 Unemployed

There are several events that result in the flow of a type k worker into unemployment prior to the dismissal stage. The type k individual could be exogenously fired, or the firm could meet another worker, and then replace the type k worker before the dismissal stage. Workers flow out of unemployment by meeting and consolidating matches with vacant firms, single worker firms, and two worker firms.

$$
\begin{aligned}
u_k^{search} =& u_k - u_k \lambda_u \frac{F_0}{F} \mathbb{I}(\widehat{V}_1(k) - \Pi_0 - U(k) > 0) \\
& - u_k \sum_i \frac{\lambda_u e_{i,0}}{F} \mathbb{I}(\widehat{V}_2(k,i) - U(k) - \widehat{V}_1(i) > 0) \\
& - u_k \underbrace{\sum_i \sum_j (\lambda_u \frac{e_{i,j}}{2F} \mathbb{I}(\max\{\widehat{V}_2(k,j) + U(i) - U(k) - \widehat{V}_2(i,j), \widehat{V}_2(i,k) + U(j) - U(k) - \widehat{V}_2(i,j)\} > 0))}_{\text{worker meets firm (i,j) and is hired}} \\
& + \sum_i e_{k,l} \frac{\lambda_u u_i}{F} \underbrace{\mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l)\} > 0)}_{\text{firm (k,l) meets unemployed and hires}} \\
& \qquad \times \underbrace{\mathbb{I}(\widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l))}_{\text{firm replaces k}} \\
& + \sum_i e_{k,l} \frac{\lambda_e e_{i,0}}{F} \underbrace{\mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l)\} > 0)}_{\text{firm (k,l) meets solo and hires}} \\
& \qquad \times \underbrace{\mathbb{I}(\widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l))}_{\text{firm replaces k}} \\
& + \sum_i \sum_j e_{k,l} \frac{\lambda_e e_{i,j}}{F} \underbrace{\mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l)\} > 0)}_{\text{firm (k,l) meets agent i in (i,j) match}} \\
& \qquad \times \underbrace{\mathbb{I}(\widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l))}_{\text{firm replaces k}} \\
& + \delta\Big(e_{k,0} + \sum_i e_{k,i}\Big)
\end{aligned} \tag{20}
$$

A.4 Solo Employed

The flow equation for single worker firms incorporates many events; flows out include normal exogenous layoffs, team formation, and on-the-job-search. Flows in include exogenous layoffs (of one member of a two worker firm), new hires, and poaching from teams.

$$
\begin{aligned}
e_{k,0}^{search} =& e_{k,0} + \delta \sum_i e_{k,i} - \delta e_{k,0} + \underbrace{u_k \lambda_u \frac{F_0}{F} \mathbb{I}(\widehat{V}_1(k) - \Pi_0 - U(k) > 0)}_{unempl\ hired} \\
& - \underbrace{e_{k,0} \frac{\lambda_e F_0}{F} \mathbb{I}(\widehat{V}_1(k) - (\widehat{V}_1(k) - \Pi_0) - \widehat{V}_1(i) > 0)}_{=0\ worker\ meets\ idle\ firm} \\
& - \underbrace{\sum_i e_{k,0} \frac{\lambda_e e_{i,0}}{F} \mathbb{I}(\widehat{V}_2(k,i) - (\widehat{V}_1(k) - \Pi_0) - \widehat{V}_1(i) > 0)}_{worker\ meets\ single\ worker\ firm} \\
& - \underbrace{\sum_i \sum_j e_{k,0} \lambda_e \frac{e_{i,j}}{2F} \mathbb{I}(\max\{\widehat{V}_2(k,j) + U(i) - [\widehat{V}_1(k) - \Pi_0] - \widehat{V}_2(i,j), \widehat{V}_2(k,i) + U(j) - [\widehat{V}_1(k) - \Pi_0] - \widehat{V}_2(i,j)\} > 0)}_{worker\ meets\ two\ worker\ firm} \\
& - \underbrace{\sum_i e_{k,0} \frac{\lambda_u u_i}{F} \mathbb{I}(\widehat{V}_2(k,i) - U(i) - \widehat{V}_1(k) > 0)}_{firm\ meets\ unemployed} \\
& - \underbrace{\sum_i e_{k,0} \frac{\lambda_e e_{i,0}}{F} \mathbb{I}[\widehat{V}_2(k,i) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_1(k) > 0]}_{firm\ meets\ solo\ worker} \\
& - \underbrace{\sum_i \sum_j e_{k,0} \frac{\lambda_e e_{i,j}}{F} \mathbb{I}[\widehat{V}_2(k,i) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_1(k) > 0]}_{firm\ meets\ team\ member\ i} \\
& + \sum_i \underbrace{e_{k,i} \frac{\lambda_e F_0}{F} \mathbb{I}[\widehat{V}_1(k) - (\widehat{V}_2(k,i) - \hat{V}_1(i)) - \Pi_0 > 0]}_{team\ member\ k\ meets\ vacant\ firm\ and\ k\ gets\ poached} \\
& + \underbrace{e_{k,l} \lambda_e \frac{F_0}{F} \mathbb{I}(\widehat{V}_1(l) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \Pi_0 > 0)}_{coworker\ poached\ by\ idle} \\
& + \underbrace{\sum_i e_{k,l} \lambda_e \frac{e_{i,0}}{F} \mathbb{I}(\widehat{V}_2(i,l) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \hat{V}_1(i) > 0)}_{coworker\ poached\ by\ solo} \\
& + \underbrace{\sum_i \sum_j e_{k,l} \lambda_e \frac{e_{i,j}}{2F} \mathbb{I}(\max\{\widehat{V}_2(i,l) + U(j) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \hat{V}_2(i,j), \widehat{V}_2(l,j) + U(i) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \hat{V}_2(i,j)\} > 0)}_{coworker\ poached\ by\ team}
\end{aligned} \tag{21}
$$

A.5 Teams

The distribution across teams is given below. There are several events that can alter the measure of (k, l) teams in an economy. Team formation and replacement hiring (occurring through the firm meeting an unemployed agent, an agent at a single worker firm, or an agent at a two worker firm) generate flows into new (k, l) teams. Natural job separation (the first two terms) reduces the number of teams, as well as exits from teams for the same set of causes (being poached, having a coworker poached, or replacement hiring).

$$
\begin{aligned}
e_{k,l}^{search} =& e_{k,l} - \underbrace{\delta e_{k,l}}_{lose\ k} - \underbrace{\delta e_{k,l}}_{lose\ l} + \underbrace{u_k \frac{\lambda_u e_{l,0}}{F} \mathbb{I}[\widehat{V}_2(k,l) - U(k) - \widehat{V}_1(l) > 0]}_{unemployed\ meets\ solo\ l} \\
& + \sum_i \sum_j \mathbb{I}(j = l) * u_k \lambda_u \frac{e_{i,j}}{2F} \underbrace{\mathbb{I}(\max\{\widehat{V}_2(k,j) + U(i) - U(k) - \widehat{V}_2(i,j), \widehat{V}_2(i,k) + U(j) - U(k) - \widehat{V}_2(i,j)\} > 0)}_{hire} \\
& * \underbrace{\mathbb{I}(\widehat{V}_2(k,j) + U(i) - U(k) - \widehat{V}_2(i,j) > \widehat{V}_2(i,k) + U(j) - U(k) - \widehat{V}_2(i,j))}_{unempl\ meets\ team\ and\ displaces\ member\ i} \\
& + \sum_i \sum_j \mathbb{I}(i = l) * u_k \lambda_u \frac{e_{i,j}}{2F} \mathbb{I}(\max\{\widehat{V}_2(k,j) + U(i) - U(k) - \widehat{V}_2(i,j), \widehat{V}_2(i,k) + U(j) - U(k) - \widehat{V}_2(i,j)\} > 0) \\
& * \underbrace{\mathbb{I}(\widehat{V}_2(i,k) + U(j) - U(k) - \widehat{V}_2(i,j) > \widehat{V}_2(k,j) + U(i) - U(k) - \widehat{V}_2(i,j))}_{unempl\ meets\ team\ and\ displaces\ member\ j} \\
& + \underbrace{e_{k,0} \frac{\lambda_e e_{l,0}}{F_1} \mathbb{I}(\widehat{V}_2(k,l) - (\widehat{V}_1(k) - \Pi_0) - \widehat{V}_1(l) > 0)}_{solo\ worker\ meets\ solo\ l\ firm} \\
& + \sum_i \sum_j \mathbb{I}(j = l) e_{k,0} \lambda_e \frac{e_{i,j}}{2F} \mathbb{I}(\max\{\widehat{V}_2(k,j) + U(i) - [\widehat{V}_1(k) - \Pi_0] - \widehat{V}_2(i,j), \widehat{V}_2(k,i) + U(j) - [\widehat{V}_1(k) - \Pi_0] - \widehat{V}_2(i,j)\} > 0) \\
& * \underbrace{\mathbb{I}(\widehat{V}_2(k,j) + U(i) - [\widehat{V}_1(k) - \Pi_0] - \widehat{V}_2(i,j) > \widehat{V}_2(k,i) + U(j) - [\widehat{V}_1(k) - \Pi_0] - \widehat{V}_2(i,j))}_{solo\ meets\ firm\ team\ and\ displaces\ member\ i} \\
& + \sum_i \sum_j \mathbb{I}(i = l) e_{k,0} \lambda_e \frac{e_{i,j}}{2F} \mathbb{I}(\max\{\widehat{V}_2(k,j) + U(i) - [\widehat{V}_1(k) - \Pi_0] - \widehat{V}_2(i,j), \widehat{V}_2(k,i) + U(j) - [\widehat{V}_1(k) - \Pi_0] - \widehat{V}_2(i,j)\} > 0) \\
& * \underbrace{\mathbb{I}(\widehat{V}_2(k,i) + U(j) - [\widehat{V}_1(k) - \Pi_0] - \widehat{V}_2(i,j) > \widehat{V}_2(k,j) + U(i) - [\widehat{V}_1(k) - \Pi_0] - \widehat{V}_2(i,j))}_{solo\ meets\ firm\ team\ and\ displaces\ member\ j} \qquad (22) \\
& + \underbrace{e_{k,0} \frac{\lambda_u u_l}{F} \mathbb{I}(\widehat{V}_2(k,l) - U(l) - \widehat{V}_1(k) > 0)}_{firm\ meets\ unempl\ l} \\
& + \underbrace{e_{k,0} \frac{\lambda_e e_{l,0}}{F} \mathbb{I}[\widehat{V}_2(k,l) - (\widehat{V}_1(l) - \Pi_0) - \widehat{V}_1(k) > 0]}_{firm\ meets\ solo\ worker} \\
& + \underbrace{\sum_i \sum_j \mathbb{I}(i = l) e_{k,0} \frac{\lambda_e e_{i,j}}{F} \mathbb{I}(\widehat{V}_2(k,i) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_1(k) > 0)}_{firm\ meets\ indiv\ i\ from\ team\ (i,j)} \\
& - \underbrace{e_{k,l} \frac{\lambda_e F_0}{F} \mathbb{I}[\widehat{V}_1(k) - (\widehat{V}_2(k,l) - \hat{V}_1(l)) - \Pi_0 > 0]}_{worker\ k\ meets\ idle\ firm} \\
& - \underbrace{\sum_i e_{k,l} \frac{\lambda_e e_{i,0}}{F} \mathbb{I}(\widehat{V}_2(k,i) - (\widehat{V}_2(k,l) - \hat{V}_1(l)) - \widehat{V}_1(i) > 0)}_{worker\ k\ meets\ solo\ firm} \\
& - \underbrace{\sum_i \sum_j e_{k,l} \lambda_e \frac{e_{i,j}}{2F} \mathbb{I}(\max\{\widehat{V}_2(k,j) + U(i) - (\widehat{V}_2(k,l) - \hat{V}_1(l)) - \widehat{V}_2(i,j), \widehat{V}_2(k,i) + U(j) - (\widehat{V}_2(k,l) - \hat{V}_1(l)) - \widehat{V}_2(i,j)\} > 0)}_{worker\ k\ meets\ firm\ team}
\end{aligned}
$$

. . . **(Continued on next page)**

(from previous page) ...

$$
\begin{aligned}
&- \underbrace{e_{k,l}\lambda_e \frac{F_0}{F}\mathbb{I}(\widehat{V}_1(l) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \Pi_0 > 0)}_{coworker\ poached\ by\ idle} \\
&- \underbrace{\sum_i e_{k,l}\lambda_e \frac{e_{i,0}}{F}\mathbb{I}(\widehat{V}_2(i,l) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \hat{V}_1(i) > 0)}_{coworker\ poached\ by\ solo\ firm} \\
&- \underbrace{\sum_i \sum_j e_{k,l}\lambda_e \frac{e_{i,j}}{2F}\mathbb{I}(\max\{\widehat{V}_2(i,l) + U(j) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \hat{V}_2(i,j), \widehat{V}_2(l,j) + U(i) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \widehat{V}_2(i,j)\} > 0)}_{coworker\ poached\ by\ firm\ team} \\
&- \underbrace{\sum_i e_{k,l} \frac{\lambda_u u_i}{F}\mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l)\} > 0)}_{firm\ meets\ unempl} \\
&- \underbrace{\sum_i e_{k,l} \frac{\lambda_e e_{i,0}}{F}\mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l), 0\} > 0)}_{my\ firm\ meets\ solo} \\
&- \sum_i \sum_j \underbrace{e_{k,l} \frac{\lambda_e e_{i,j}}{F}\mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \hat{V}_1(j)) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \hat{V}_1(j)) - \widehat{V}_2(k,l)\} > 0)}_{my\ firm\ meets\ i\ from\ team\ (i,j)}
\end{aligned}
$$

A.6 Dismissal Stage

In the dismissal stage, for two worker firms, either both workers can be dismissed or one worker can be dismissed. For single worker firms, there is only one choice of whether or not to dismiss the sole employee.

$$
\begin{aligned}
u_k^+ &= u_k^{search} + e_{k,0}^{search}\mathbb{I}(\widehat{V}_1(k) < \Pi_0 + U(k)) \\
&+ \underbrace{e_{k,l}^{search}\mathbb{I}[max\{U(k) + U(l) + \Pi_0 - V_2(k,l), V_1(k) + U(l) - V_2(k,l), V_1(l) + U(k) - V_2(k,l), 0\} = U(k) + U(l) + \Pi_0 - V_2(k,l)]}_{dimiss\ both} \\
&+ \underbrace{e_{k,l}^{search}\mathbb{I}[max\{U(k) + U(l) + \Pi_0 - V_2(k,l), V_1(k) + U(l) - V_2(k,l), V_1(l) + U(k) - V_2(k,l), 0\} = V_1(l) + U(k) - V_2(k,l)]}_{dismiss\ k}
\end{aligned} \tag{23}
$$

The flows into single worker firms come from teams that dismiss one of their employees,

$$
e_{k,0}^+ = e_{k,0}^{search} + \underbrace{e_{k,l}^{search}\mathbb{I}[max\{U(k) + U(l) + \Pi_0 - V_2(k,l), V_1(k) + U(l) - V_2(k,l), V_1(l) + U(k) - V_2(k,l), 0\} = V_1(k) + U(l) - V_2(k,l)]}_{dismiss\ l} \tag{24}
$$

Only teams that do not dimiss will remain as teams at the start of the next period,

$$e^{+}_{k,l} = e^{search}_{k,l} + \underbrace{e^{search}_{k,l} \mathbb{I}[max\{U(k) + U(l) + \Pi_0 - V_2(k,l), V_1(k) + U(l) - V_2(k,l), V_1(l) + U(k) - V_2(k,l), 0\} = 0]}_{no\ dismissal} \tag{25}$$

The values u^+_k, $e^+_{k,0}$, and $e^+_{k,l}$ become next period's starting values u^-_k, $e^-_{k,0}$, and $e^-_{k,l}$, respectively. An equilibrium is a fixed point of this mapping.

B Planner's Problem

This section includes the planner's continuation values for both single-worker firms and two-worker firms net of the opportunity cost of creating those jobs (the opportunity cost is the value of the contacts forfeited by losing a vacant firm), as well as the planner's continuation value for the unemployed. The opportunity cost of creating a match includes the forfeited contacts of a vacant firm:

$$Planners\ Opportunity\ Cost\ of\ Match = \sum_i \lambda_u \frac{u_i}{n} \max\{\widehat{V}_1^P(i) - U^P(i), 0\}$$
$$+ \sum_i \lambda_e \frac{e_{i,0}}{n} \max\{\widehat{V}_1^P(i) - \widehat{V}_1^P(i), 0\} + \sum_i \sum_j \lambda_e \frac{e_{i,j}}{n} \max\{\widehat{V}_1^P(i) + \widehat{V}_1^P(j) - \widehat{V}_2^P(i,j), 0\}$$

The value of an unemployed worker to the planner is,

$$\begin{aligned}
&U^P(k) = b(k) + \beta \mathbb{E}_{k_+}\big[U^P(k_+) \\
&+ \lambda_u \frac{n_0}{n} \max\{\widehat{V}_1^P(k_+) - U^P(k_+), 0\} \\
&+ \lambda_u \sum_i \frac{e_{i,0}}{n} \max\{\widehat{V}_2^P(k_+, i) - \widehat{V}_1^P(i) - U^P(k_+), 0\} \\
&+ \lambda_u \sum_i \sum_j \frac{e_{i,j}}{2n} \max\{\widehat{V}_2^P(k_+, i) + U(j) - \widehat{V}_2^P(i,j) - U^P(k_+), \\
&\qquad\qquad \widehat{V}_2^P(k_+, j) + U^P(i) - \widehat{V}_2^P(i,j) - U^P(k_+), 0\}\big]
\end{aligned}$$

The joint value of an additional single worker firm to the planner is,

$$
\begin{aligned}
&V_1^P(k) = f(k,0) + \beta \mathbb{E}_{k+}\{\widehat{V}_1^P(k+) \\
&+ \lambda_e \frac{n_0}{n} \max\{\widehat{V}_1^P(k_+) - \widehat{V}_1^P(k_+), 0\} \\
&+ \lambda_e \sum_i \frac{e_{i,0}}{n} \max\{\widehat{V}_2^P(k_+, i) - \widehat{V}_1^P(i) - \widehat{V}_1^P(k_+), 0\} \\
&+ \lambda_e \sum_i \sum_j \frac{e_{i,j}}{2n} \max\{\widehat{V}_2^P(k_+, i) + U^P(j) - \widehat{V}_2^P(i,j) - \widehat{V}_1^P(k_+), \widehat{V}_2^P(k_+, j) + U^P(i) - \widehat{V}_2^P(i,j) - \widehat{V}_1^P(k_+), 0\} \\
&+ \sum_i \lambda_u \frac{u_i}{n} \max\{\widehat{V}_2^P(i, k_+) - U^P(i) - \widehat{V}_1^P(k_+), 0\} \\
&+ \sum_i \lambda_e \frac{e_{i,0}}{n} \max\{\widehat{V}_2^P(i, k_+) - \widehat{V}_1^P(i) - \widehat{V}_1^P(k_+), 0\} \\
&+ \sum_i \sum_j \lambda_e \frac{e_{i,j}}{n} \max\{\widehat{V}_2^P(i, k_+) + \widehat{V}_1^P(j) - \widehat{V}_2^P(i,j) - \widehat{V}_1^P(k_+), 0\} \\
&- \underbrace{\sum_i \lambda_u \frac{u_i}{n} \max\{\widehat{V}_1^P(i) - U^P(i), 0\} - \sum_i \lambda_e \frac{e_{i,0}}{n} \max\{\widehat{V}_1^P(i) - \widehat{V}_1^P(i), 0\} - \sum_i \sum_j \lambda_e \frac{e_{i,j}}{n} \max\{\widehat{V}_1^P(i) + \widehat{V}_1^P(j) - \widehat{V}_2^P(i,j), 0\}}_{\textit{opportunity cost to planner}} \\
&+ \delta(U^P(k) - \widehat{V}_1^P(k))\}
\end{aligned}
$$

The joint value of a two worker firm to the planner is,

$$
\begin{aligned}
&\widehat{V}_2^P(k,l) = f(k,l) + \beta\mathbb{E}_{k_+,l_+}\{\widehat{V}_1^P(k_+,l_+) \\
&+ \lambda_e \frac{n_0}{n}\max\{\widehat{V}_1^P(k_+) + \widehat{V}_1^P(l_+) - \widehat{V}_2^P(k_+,l_+),0\} \\
&+ \lambda_e \sum_i \frac{e_{i,0}}{n}\max\{\widehat{V}_2^P(k_+,i) + \widehat{V}_1^P(l_+) - \widehat{V}_2^P(k_+,l_+) - \widehat{V}_1^P(i),0\} \\
&\underbrace{+ \lambda_e \sum_i\sum_j \frac{e_{i,j}}{2n}\max\{\widehat{V}_2^P(k_+,i) + \widehat{V}_1^P(l_+) + U^P(j) - \widehat{V}_2^P(k_+,l_+) - \widehat{V}_2^P(i,j), \widehat{V}_2^P(k_+,j) + \widehat{V}_1^P(l_+) + U^P(i) - \widehat{V}_2^P(k_+,l_+) - \widehat{V}_2^P(i,j),0\}}_{k\ meets\ (i,j)\ team} \\
&+ \lambda_e \frac{n_0}{n}\max\{\widehat{V}_1^P(l_+) + \widehat{V}_1^P(k_+) - \widehat{V}_2^P(k_+,l_+),0\} \\
&+ \sum_i \lambda_e \frac{e_{i,0}}{n}\max\{\widehat{V}_2^P(l_+,i) + \widehat{V}_1^P(k_+) - \widehat{V}_2^P(k_+,l_+) - \widehat{V}_1^P(i),0\} \\
&\underbrace{+ \sum_i\sum_j \lambda_e \frac{e_{i,j}}{2n}\max\{\widehat{V}_2^P(l_+,i) + \widehat{V}_1^P(k_+) + U^P(j) - \widehat{V}_2^P(k_+,l_+) - \widehat{V}_2^P(i,j), \widehat{V}_2^P(l_+,j) + \widehat{V}_1^P(k_+) + U^P(i) - \widehat{V}_2^P(k_+,l_+) - \widehat{V}_2^P(i,j),0\}}_{l\ meets\ (i,j)\ team} \\
&+ \sum_i \lambda_u \frac{u_i}{n}\max\{\widehat{V}_2^P(k_+,i) + U^P(l_+) - U^P(i) - \widehat{V}_2^P(k_+,l_+), \widehat{V}_2^P(i,l_+) + U^P(k_+) - U^P(i) - \widehat{V}_2^P(k_+,l_+),0\} \\
&+ \sum_i \lambda_e \frac{e_{i,0}}{n}\max\{\widehat{V}_2^P(k_+,i) + U^P(l_+) - \widehat{V}_1^P(i) - \widehat{V}_2^P(k_+,l_+), \widehat{V}_2^P(i,l_+) + U^P(k_+) - \widehat{V}_1^P(i) - \widehat{V}_2^P(k_+,l_+),0\} \\
&+ \sum_i\sum_j \lambda_e \frac{e_{i,j}}{n}\max\{\widehat{V}_2^P(k_+,i) + U^P(l_+) - (\widehat{V}_2^P(i,j) - \widehat{V}_1^P(j)) - \widehat{V}_2^P(k_+,l_+), \widehat{V}_2^P(i,l_+) + U^P(k_+) - (\widehat{V}_2^P(i,j) - \widehat{V}_1^P(j)) - \widehat{V}_2^P(k_+,l_+),0\} \\
&- \sum_i \lambda_u \frac{u_i}{n}\max\{\widehat{V}_1^P(i) - U^P(i),0\} \\
&- \sum_i \lambda_e \frac{e_{i,0}}{n}\max\{\widehat{V}_1^P(i) - \widehat{V}_1^P(i),0\} \\
&- \sum_i\sum_j \lambda_e \frac{e_{i,j}}{n}\max\{\widehat{V}_1^P(i) + \widehat{V}_1^P(j) - \widehat{V}_2^P(i,j),0\} \\
&+ \delta(\widehat{V}_1^P(k_+) + U^P(l_+) - V_2^P(k_+,l_+)) \\
&+ \delta(\widehat{V}_1^P(l_+) + U^P(k_+) - V_2^P(k_+,l_+))\}
\end{aligned}
$$

C Data Appendix

Our unit of observation is a State Employment Identification Number (SEIN), and we focus on single-unit firms within a state, meaning that the firm only had one physical plant location in that state and it thus corresponds to the SEIN in the LEHD database. We identify primary employers as the SEINs which pay the individual the most in a given year. Coworker wages are measured as the quarter 4 wage bill net of an individual's own wage divided by the number of employees in quarter 4 at the SEIN.

To be in our base dataset, one must have worked at least 260 hours (one quarter, parttime) at least once between 1998 and 2008. These are similar restrictions used in Guvenen et al. [2015]. Our variables are deflated by the BLS CPI, and we winsorize the top 1% of all continuous variables. We define an individual to be full-year employed if they earn above $1,000 in each quarter of the year. We define an individual to be non-employed in a given quarter if they earn less than $1,000. Our results are not sensitive to this cutoff.

Individuals are in the sample if they work at single-unit firm (meaning all the workers are working in the same physical area) with between 2 and 250 employees, are prime-age, and are male. The sample covers 2001 to 2008, but based on the forward lags of the variables used, the last cohort used for job transitions is 2006, with outcomes measured in 2008. An EUE transition is defined as a full-year of employment in year t, one quarter of non-employment in year t+1, and full-year employment in year t+2.

D Data Description: Brazilian Data

In this section, we verify our results on Brazilian data. We conduct several important robustness checks: (i) we specify coworker wages within an occupation, (ii) we follow workers over a 4 year horizon, (iii) we isolate industry switchers, (iv) we compare our results to standard methods, and (v) we look at learning based on different summary measures of the firm's workforce, including the different deciles of the coworker wage distribution.[10]

We use the RAIS database from Brazil, which is an annual census of the universe of formal firms, including the establishments of those firms, both public and private, in Brazil.[11] We use a sample of this data from the state of Bahia in Brazil which includes approximately 37 million person-year observations. Bahia is the fourth largest state in Brazil with a population of approximately 15 million individuals. Our sample of data spans 1998 to 2010.

Unlike the US administrative data (such as the LEHD), RAIS includes occupation data.[12] We are therefore able to isolate coworkers within a given firm and 2-digit occupation code. We

[10] All RAIS results contained in this paper were run on IPEA servers in accordance with MTE guidelines.

[11] In English RAIS roughly stands for Annual Relation of Social Information (RAIS). For more discussion of the data, see Araujo [2014] and Engbom and Moser [2016].

[12] Brazil's occupational classification system is called the CBO (Classificacao Brasileira de Ocupacoes), and is discussed at length in Muendler et al. [2004]. The CBO is more detailed than international occupation classifications, and Muendler et al. [2004] maps the CBO into the international standard classification ISCO (International Standard Classification of Occupations). As Muendler et al. [2004] describe, "At the finest level, CBO-94 defines an individual occupation as a category that unifies jobs which are fundamentally similar with regard to their 'content' and 'skill requirements."'

believe the jobs that are being carried out within these establishment level 2-digit occupation bins have similar skill requirements.

Another advantage of the Brazilian data is that we observe the reason for job transition (retirement, transfer, fired, quit). We are therefore able to condition on the worker being fired.

D.1 Empirical Approach: Brazilian Data

Our empirical approach follows Section 4.2. We therefore focus on the following sample: Prime age (24 to 65) males with at least 1 year of tenure who must have positive earnings in year t and t+2 from different primary employers (the primary employer is simply the establishment that paid the worker the most in a given year); moreover, they must be fired and switch primary employers in year t+1 and have at least one quarter of non-employment in year t+1. We require that they have at least one coworker within the establishment, within the same 2-digit occupation. We further condition on the individuals earning at least 100 real per month (in 2010 real), in order to get rid of workers who are earning below the minimum wage and likely to be in the informal sector.

All variables are winsorized at the top 1%, and nominal variables are deflated using the Brazilian CPI. Let $\tilde{w}_{i,t}$ denote an individual's raw *average monthly wage* at his/her primary employer at date t. Our main outcome of interest is the log of this wage, $w_{i,t} = ln(\tilde{w}_{i,t})$. Furthermore, define $\tilde{w}_{-i,t}$ as the raw *average monthly wage* of the coworkers of an individual at his/her primary employer at date t within the same 2-digit occupation. Let $\bar{w}_{-i,t} = ln(\tilde{w}_{-i,t})$ denote the log of this variable.

D.2 Estimation Equation: Brazilian Data

Let i index people and t index years. Let $Y_{i,t+2}$ be the dependent variable of interest, e.g. the individual i's log *average monthly wage* year t+2 or the *log average monthly wage of workers* in a given occupation at individual i's primary employer in year t+2. The primary employer is defined as the employer that paid the worker the most in a given year. We estimate the following specifications on the Brazilian data:

$$Y_{i,t+2} = \beta_0 + \gamma_t + \beta_1 w_{i,t} + \beta_2 \bar{w}_{-i,t} + \Gamma X_{i,t} + \epsilon_{i,t}$$

D.3 Baseline Specifications: Brazilian Data

Table 9 illustrates the impact of coworker wages, within the same establishment and same 2-digit occupation, on the subsequent wages of an individual after a spell of non-employment. The independent covariates are measured as of date t, and the outcome variable is an individual's log average monthly wage at their new employer at date t+2. Each specification includes fixed effects for contract type, year, sector (1 digit), occupation (2 digit), as well as quadratics in age and tenure.

Column (1) reveals a strong correlation between coworker wages and an individual's future wages for EUE transitioners. This coefficient is about two and a half times smaller than the same coefficient in the LEHD data. If an individual's coworkers earn 10% more, the individual will earn, on average, a .16% greater wage upon reemployment. Column (2) restricts the sample to EUE transitioners who were initially below the mean wage, within their occupation at their establishment. Consistent with the LEHD, the effects of coworker wages are roughly two times as large as the population average effect. Column (4) focuses on workers above the mean wage, and the effect is indiscernible from zero.

To check whether we are measuring general versus industry-specific or establishment-specific forms of human capital, Column (4) focuses on those who switch 1-digit sectors in their EUE transition.[13] Column (4) reveals that the elasticity of future wage with respect to coworker wages is quite stable, even among industry switchers. This suggests that our measures of peer-effects are capturing general human capital accumulation, as opposed to other forms of human capital accumulation.

As another test of whether we are measuring human capital or some other short-term characteristic of the prior employer, Table 10 repeats our regressions with different horizons ranging from 1 year after job loss (reemployed at t+2, lost job at t+1), to 4 years after job loss, (reemployed at t+5, lost job at t+1). Column (1) includes the 1 year horizon, Column (2) is the 2 year horizon, Column (3) is the 3 year horizon, and Column (4) is the 4 year horizon. Table 10 reveals a fairly stable impact of coworker wages on future individual wages. The impact of coworker wages on future individual wages remains significant 4 years after

[13] These sectors include: 1 "Metal & Mineral Manufacturing"; 2 "Machinery & Eletronics Manufacturing" 3 "Others Manufacturing"; 4 "Chemical Manufacturing"; 5 "Textile Manufacturing"; 6 "Food & Drink Manufacturing" 7 "Construction" ; 8 "Retail Trade"; 9 "Wholesale Trade" ; 10 "Finance & Insurance Services" ; 11 "Real State Services" ; 12 "Traffic & Technical Services" ; 13 "Accommodation & Food Services"; 14 "Health & Social Services" ; 15 "Educational Services" ; 16 "Government" ; 17 "Agriculture, etc"

the initial job loss.

Lastly, Table 11 reruns our regression with additional controls. We include 3-years worth of lagged individual wages in Column (1), and our point estimate declines slightly. Column (2) uses coworker education attainment instead of wages as the main independent variable of interest, and we effects in line with Nix [2015]; we discuss this in more detail below. Column (3) reveals that both coworker wages and coworker educational attainment have a positive impact on future individual wages if they are both included at the same time. Column (4) shows that wage dispersion within an occupation and establishment is associated with lower learning. Lastly Column (5) includes other measures of the wage distribution, in addition to the average. The results are difficult to interpret due to the magnitudes of the coefficients, but other moments of the wage distribution do not appear as important as the average.

Table 9: Baseline Job-Transition Specification (EUE). Dependent variable is Log Avg. Monthly Wage of Individual at Date t+2. (RAIS:2000-2008)

Sample:	**(1)** Full	**(2)** Below Mean	**(3)** Above Mean	**(4)** Industry Switcher (1-digit)
Log of Average Coworker Wages date t (Same Estab & 2-Dig OCC)	0.016***	0.032***	0.003	0.046***
	(4.06)	(5.13)	(0.34)	(2.88)
Log of Avg Monthly Indiv Wages date t	0.563***	0.543***	0.572***	0.573***
	(133.10)	(70.14)	(67.55)	(34.03)
Number of Coworkers (Same Estab and 2-Dig Occ) date t	-0.000	-0.000	-0.000***	0.000
	(-0.92)	(-0.14)	(-3.27)	(0.71)
Observations	133,706	72,270	61,436	7,633
R-squared	0.511	0.455	0.553	0.522
Demographic Controls	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
OCC FE	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y
Contract type FE	Y	Y	Y	Y

*Notes: SE clustered at SEIN level. *** p<0.01, ** p<0.05, * p<0.1. Demographic controls include: education dummies and quadratics in age and tenure.*

D.4 Replicating Nix in the Brazilian Data

To benchmark the data, we estimate specifications similar to Nix [2015]. Rather than look at the wages of coworkers, Nix [2015] focuses on the impact of educational attainment of coworkers, $h_{-i,t}$, on the individual's log wage in the following year, $w_{i,t+1}$. This empirical

Table 10: Time Series Impact of Coworker Wages on Future Individual Wages (EUE). Dependent variable is Log Avg. Monthly Wage of Individual at Date t+2, t+3, t+4, and t+5. (RAIS:2000-2008)

Sample:	**(1)** EUE, t+2	**(2)** EUE, t+3	**(3)** EUE, t+4	**(4)** EUE, t+5
Log of Average Coworker Wages date t (Same Estab & 2-Dig OCC)	0.016***	0.025***	0.018***	0.014**
	(4.06)	(5.15)	(3.39)	(2.38)
Log of Avg Monthly Indiv Wages date t	0.563***	0.560***	0.552***	0.531***
	(133.10)	(110.63)	(96.41)	(81.65)
Number of Coworkers (Same Estab and 2-Dig Occ) date t	-0.000	0.000	0.000	0.000
	(-0.92)	(0.37)	(0.19)	(0.08)
Observations	133,706	99,036	75,778	58,839
R-squared	0.511	0.486	0.467	0.450
Demographic Controls	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
OCC FE	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y
Contract type FE	Y	Y	Y	Y

*Notes: SE in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Demographic controls include: education dummies and quadratics in age and tenure.*

strategy uses a host of fixed effects to disentangle sorting from peer effects.

In this section, our sample imposes minimal restriction: we include all prime age males with non-missing educational attainment. In our specifications, we include worker fixed effects (α_i), establishment fixed effects (ψ_j), industry, contract type, and occupation fixed effects (which we do not separately write out but include in $X_{i,t}$), and dynamic controls such as quadractics in age and tenure. In particular we estimate regressions of the following form,

$$w_{i,t+1} = \beta_0 + \alpha_i + \psi_j + \gamma_t + \beta_1 h_{i,t} + \beta_2 h_{-i,t} + \Gamma X_{i,t} + \epsilon_{i,t}$$

We define coworker human capital, $h_{-i,t}$, as the fraction of coworkers with a college degree of an individual at his/her primary employer at date t.

Table 12 replicates Nix [2015]'s regressions. Column (1) includes no controls or fixed effects, Column (2) only includes worker and establishment fixed effects, and Column (3) includes all fixed effects. The coefficient on coworker wages in Column (3) implies that if 10% more of an individual's coworkers have a college degree, an individual earns 1.6% more. Nix [2015] reports that "increasing average education of a given worker's colleagues by 10 percentage points increases that worker's wages in the following year by 0.3%, which is significant at the 1% level." The discrepancy between the Brazilian and Swedish data may

Table 11: Other moments of the establishment wage distribution. DV: Log Avg. Monthly Wage of Individual at Date t+2 (After Transition). (RAIS:2000-2008, Sample 1)

Sample:	**(1)** Full	**(2)** Full	**(3)** Full	**(4)** Full	**(5)** Full
Log of Average Coworker Wages date t (Same Estab & 2-Dig OCC)	0.011**		0.009*	0.030***	-0.037***
	(2.35)		(1.86)	(6.59)	(-7.15)
Log of Avg Monthly Indiv Wages date t	0.494***	0.500***	0.494***	0.557***	0.546***
	(66.42)	(74.36)	(66.47)	(128.89)	(121.20)
Number of Coworkers (Same Estab and 2-Dig Occ) date t	-0.000	-0.000	-0.000	-0.000	-0.000
	(-0.49)	(-0.56)	(-0.61)	(-0.27)	(-0.47)
Log of Avg Monthly Indiv Wages date t-1	0.029***	0.029***	0.029***		
	(5.55)	(5.55)	(5.54)		
Log of Avg Monthly Indiv Wages date t-2	0.057***	0.057***	0.057***		
	(13.59)	(13.60)	(13.57)		
Fraction of Coworkers with College Degree		0.057***	0.052**		
		(2.79)	(2.51)		
Log of Average Coworker Wages date t (Not within occupation)					
Coefficient of Variation of Wages (Same Estab & 2-Dig OCC)				-0.029***	
				(-8.04)	
25th Percentiles of Wages (Same Estab & 2-Dig OCC)					0.000***
					(5.89)
50th Percentile of Wages (Same Estab & 2-Dig OCC)					0.000
					(0.27)
75th Percentile of Wages (Same Estab & 2-Dig OCC)					0.000***
					(4.26)
Observations	93,100	93,100	93,100	133,706	133,706
R-squared	0.528	0.528	0.528	0.511	0.512
Demographic Controls	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
OCC FE	Y	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y	Y
Contract type FE	Y	Y	Y	Y	Y

*Notes: SE in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Demographic controls include: education dummies and quadratics in age and tenure.*

be explained by the fact that so few individuals in Brazil have a college education. In our Nix-replication sample only 8% of coworkers have a college degree.

Table 12: Nix (2015) Replication. DV: Individual's log avg. monthly wage at primary employer. (Source: RAIS 1998-2010, 10% Random Sample)

	(1)	**(2)**	**(3)**
Fraction of Coworkers with College Degree	0.340***	0.203***	0.166***
	(12.45)	(7.56)	(5.39)
Observations	778,021	778,021	738,147
R-squared	0.000	0.000	0.008
Demographic Controls	N	N	Y
Year FE	N	N	Y
OCC FE	N	N	Y
Sector FE	N	N	Y
Indiv FE	N	Y	Y
Establishment FE	N	Y	Y

*Notes: SE in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Demographic controls include: quadratics in age and tenure.*

References

John M Abowd, Francis Kramarz, and David N Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

Axel Anderson. A dynamic generalization of becker's assortative matching result. *Journal of Economic Theory*, 159:290–310, 2015.

Axel Anderson and Lones Smith. Dynamic matching and evolving reputations. *The Review of Economic Studies*, 77(1):3–29, 2010.

Ana Luisa Pessoa Araujo. Wage inequality and job stability. 2014.

Tania Babina. Destructive creation at work: How financial distress spurs entrepreneurship. 2015.

Stéphane Bonhomme, Thibaut Lamadon, and Elena Manresa. A distributional framework for matched employer employee data. 2014.

Katarna Borovickova and Robert Shimer. High wage workers work for high wage firms. *Manuscript*, 2017.

Wyatt Brooks, Kevin Donovan, and Terence R Johnson. Mentors or teachers? microenterprise training in kenya. *Manuscript*, 2017.

Martin Browning and Thomas F Crossley. Unemployment insurance benefit levels and consumption changes. *Journal of Public Economics*, 80(1):1–23, 2001.

Francisco J Buera and Ezra Oberfield. The global diffusion of ideas. Technical report, National Bureau of Economic Research, 2016.

Pierre Cahuc, Fabien Postel-Vinay, and Jean-Marc Robin. Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, 74(2):323–364, 2006.

Hector Chade and Jan Eeckhout. Stochastic sorting. In *2013 Meeting Papers*, number 1222. Society for Economic Dynamics, 2013.

Hector Chade, Jan Eeckhout, and Lones Smith. Sorting through search and matching models in economics. *Journal of Economic Literature*, 55(2):493–544, 2017.

Satyajit Chatterjee and Esteban Rossi-Hansberg. Spinoffs and the market for ideas. *International Economic Review*, 53(1):53–93, 2012.

Jonathan Chiu, Cesaire Meh, and Randall Wright. Innovation and growth with financial, and other, frictions. *International Economic Review*, 58(1):95–125, 2017.

Thomas Cornelissen, Christian Dustmann, and Uta Schönberg. Peer effects in the workplace. *American Economic Review*, 2016.

Jan Eeckhout and Boyan Jovanovic. Knowledge spillovers and inequality. *American Economic Review*, 92(5):1290–1307, 2002.

Niklas Engbom and Christian Moser. Earnings inequality and the minimum wage: Evidence from brazil. 2016.

Fatih Guvenen, Fatih Karahan, Serdar Ozkan, and Jae Song. What do data on millions of us workers reveal about life-cycle earnings risk? Technical report, National Bureau of Economic Research, 2015.

Marcus Hagedorn, Tzuo Hann Law, and Iourii Manovskii. Identifying equilibrium models of labor market sorting. Technical report, National Bureau of Economic Research, 2012.

Tom-Reiel Heggedal, Espen R Moen, and Edgar Preugschat. Productivity spillovers through labor mobility in search equilibrium. *Journal of Economic Theory*, 169:551–602, 2017.

Kyle Herkenhoff, Gordon Phillips, and Ethan Cohen-Cole. How credit constraints impact job finding rates, sorting & aggregate output. Technical report, National Bureau of Economic Research, 2016.

Boyan Jovanovic. Misallocation and growth. *American Economic Review*, 104(4):1149–71, 2014.

Boyan Jovanovic and Rafael Rob. The growth and diffusion of knowledge. *The Review of Economic Studies*, 56(4):569–582, 1989.

Rasmus Lentz and Nicolas Roys. Training and search on the job. Technical report, National Bureau of Economic Research, 2015.

Jeremy Lise and Fabien Postel-Vinay. Multidimensional skills, sorting, and human capital accumulation. *(Working Paper)*, 2015.

Robert E Lucas and Benjamin Moll. Knowledge growth and the allocation of time. *Journal of Political Economy*, 122(1):1–51, 2014.

Erzo G.J. Luttmer. An assignment model of knowledge diffusion and income inequality. *Federal Reserve Bank of Minneapolis Research Department Staff Report*, 715, 2014.

Alexandre Mas and Enrico Moretti. Peers at work. *The American economic review*, 99(1): 112–145, 2009.

Alexander Monge-Naranjo. Foreign firms and the diffusion of knowledge. 2012.

Marc-Andreas Muendler, Jennifer Poole, Garey Ramey, and Tamara Wajnberg. Job concordances for brazil: Mapping the classificaçao brasileira de ocupaçoes (cbo) to the international standard classification of occupations (isco-88). *University of California, San Diego, unpublished manuscript*, 2004.

Emily Nix. Learning spillovers in the firm. Technical report, mimeo, 2015.

Jesse Perla and Christopher Tonetti. Equilibrium imitation and growth. *Journal of Political Economy*, 122(1):52–76, 2014.

Fabien Postel-Vinay and Jean-Marc Robin. Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, 70(6):2295–2350, 2002.

Michel Serafinelli. Good firms, worker flows and local productivity. Technical report, University of Toronto, Department of Economics, 2015.

Robert Shimer and Lones Smith. Assortative matching and search. *Econometrica*, 68(2): 343–369, 2000.

James R Spletzer. Inequality statistics from the lehd. In *June 2014 FESAC meetings, http://www. census. gov/fesac/pdf/2014-06-13/Spletzer_Background. pdf*, 2014.

Andrey Stoyanov and Nikolay Zubanov. Productivity spillovers across firms through worker mobility. *American Economic Journal: Applied Economics*, 4(2):168–98, 2012.