Worker Mobility and the Diffusion of Knowledge*

Kyle Herkenhoff Jeremy Lise Guido Menzio Gordon Phillips

October 4, 2017

Preliminary and Incomplete

Abstract

We develop a theory of teams to measure the way knowledge diffuses across workers. We extend the frictional sorting framework of Postel-Vinay and Robin [2002] to allow for workers 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. Using a new methodolgy, we estimate the knowledge diffusion process and the degree of worker complementarities in production with micro wage data and job mobility patterns from the LEHD, as well as startup patterns from the Integrated LBD. Our estimated parameters imply both positive peer effects (lower types learning from higher types) and negative peer effects (higher types dislearning from lower types). We find that eliminating positive peer effects would lower output by 7%, eliminating negative peer effects would increase output by 11%, and eliminating on-the-job-search would reduce output by 17%. Our estimates also imply strong production complementarities. Setting production complementaries to zero, but still allowing for learning and worker flows, would reduce output by 52%. Remarkably, under the estimated parameters, welfare would only be .02% higher if we reshuffled workers to achieve the planner's allocation.

*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 FGV, SAET, SaM-Barcelona GSE, and the SED. 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? What is the role of the workplace? And, what is the role of worker mobility? In this paper, we develop a model which allows us to estimate the role of peer effects and worker mobility on knowledge diffusion. By doing so, we contribute to a fast and growing literature on knowledge diffusion (*inter alia* Jovanovic and Rob [1989], Luttmer et al. [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 theory more generally allows for workers to learn or dislearn).

Our quantitative contribution is to solve and estimate the model using 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.

To jointly estimate peer effects and sorting, our innovation is to view the data through the lens of a search and matching model. 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. Our assumption is that the unemployment spell resets both wage stickiness and built-up rents, which allows us to measure learning and sorting without resorting to fixed effects, e.g. Nix [2015] and Cornelissen et al. [2016]. Even though most workers will experience a wage loss following job loss, it is the *relative* wage loss between workers who had better or worse past colleagues that allows us to identify learning. The relationship between an individual's past wage and future coworker ability, in conjunction with the model's structure, allows us to estimate the degree of complementarity in production.

Using this methodology, we establish several new facts: (1) coworker wages are positively related to future coworker wages and employer size, (2) coworker wages have a positive impact on individual wages, (3) the number of coworkers has a negative effect on individual wages, (4) coworkers wages disproportionately impact workers earning below the mean wage of the employer, (5) workers with higher earning coworkers are much less likely to spin off a business. We estimate the model to match these facts. In particular, we run the same regressions on our model as we do in the data, and we use simulated method of moments to determine the parameter values.

Our first finding is that peer effects are an important determinant of US output. Our estimated parameters imply both positive peer effects (lower types learning from higher types) and negative peer effects (higher types dislearning from lower types). Setting all peer effect parameters (positive peer effects and negative peer effects) to zero, but still allowing for worker mobility, would reduce steady state output by 3%. While the positive and negative peer effects are roughly offsetting, their impact on output, separately, is large. We find that eliminating positive peer effects would lower output by 7% and eliminating negative peer effects would increase output by 11%.

Our second finding is that the interaction between worker mobility and learning is large and positive. Restricting workers' ability to conduct on-the-job search to zero, but still allowing for peer effects, would reduce steady state output by 17%. This large loss primarily reflects changes in sorting patterns. However, mobility and learning interact non-linearly. Shutting off both learning and worker mobility reduces output by 25%, which is significantly larger than the simple sum of losses from shutting off each channel alone.

Our third finding is that worker complementarity is a larger determinant of US output than learning and/or mobility. If we set our estimated worker complementary to zero, but still allow for unrestricted learning and worker flows, output would drop by 52%.

In our model, the net present value of wages to be delivered to a worker is determined by Bertrand competition. This results in several inefficiencies. Teams that hire low-type workers and train those workers are only compensated for the value of the worker in their existing match, not the value of the worker to poaching firms (or whomever benefits the most from having that trained worker). Thus there are not as many teams which pair low-type workers with high-type workers (we call these 'schools') in the decentralized equilibrium relative to the social planner's problem. However, despite generating more schools, we find that the social planner's allocation would only modestly increase output by .02%.

The paper proceeds as follows. Section 2 discusses the related literature in more detail. Section 3 describes the LEHD, our empirical approach, and our main regression results. Section 4 includes the model. Section 5 describes the calibration. Section 6 illustrates the main knowledge diffusion decomposition, and Section 7 compares the decentralized economy the to planner's solution. Section 8 concludes.

2 Related Literature

Our discussion focuses on sorting models with dynamic types. We refer readers to Chade, Eeckhout, and Smith [2017] for a comprehensive survey of sorting models, both with and without dynamics types. Anderson and Smith [2010] generalize the assignment model of Becker [1973] to a dynamic setting in which partner's types may change over time. They show that the conditions for positive assortative matching laid out by Becker [1973] unravel in their framework. In followup work, Anderson [2015] derives sufficient conditions for PAM in assignment models with dynamically evolving types. Chade and Eeckhout [2013] consider a class of search and matching models in which types are realized after matching. 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.

Relevant quantitative work includes Lentz and Roys [2015] who 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. They estimate their model on the NLSY, and they show that training decreases as search frictions rise, in contrast to earlier partial equilibrium findings. Likewise, Lise and Postel-Vinay [2015] allow the vector of workers' types to evolve as a function of the type of job the worker obtains; they estimate their model on the NLSY as well. Herkenhoff et al. [2016] allow workers' human capital to evolve over the course of a match; firms dynamically respond by increasing or decreasing investment in capital, and they estimate their model on the LEHD-TransUnion dataset. Using German data, Gulyas [2016] uses firm firing patterns in response to a negative firm-type-shock to identify sorting. Our contribution to this literature is to develop a search and matching framework in which workers influence each other's types, and then to estimate the speed of learning and degree or production complementarity using the LEHD.

There is also a growing trade literature that models knowledge diffusion through firm and entrepreneur interactions, (e.g. Monge-Naranjo [2012] and Buera and Oberfield [2016]). In the development literature, Brooks, Donovan, and Johnson [2017] measure knowledge diffusion through a randomized controlled trial in which they matched experienced entrepreneurs with inexperienced entrepreneurs. They find that entrepreneurial mentorship increases the mentee's profits by 20%, largely though local knowledge of suppliers. Our work is also related to the largely theoretic work of Chatterjee and Rossi-Hansberg [2012] on knowledge diffusion through spinoffs.

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. These studies rely on a host of fixed effects to control for selection and sorting.[1] Both Cornelissen et al. [2016] and Nix [2015] find significant, but relatively small, coworker influence on future wages. Our contribution to the empirical literature on peer-effects is to, (i) treat an individual's type as time-varying, (ii) parse-out worker rents using job transitions, and (iii) focus on future employer characteristics to measure sorting.

3 Empirical Approach

The main empirical objective is to measure moments of the data that, in conjunction with the structure of our model, allow us to identify (a) knowledge diffusion and (b) sorting. Unlike the existing literature, we view the data through the lens of a search and matching framework with sorting and human capital accumulation, and we recognize that a worker's

[1] Cornelissen et al. [2016] and Nix [2015] both estimate variants the following log wage equation, with time-invariant own-ability a_i, coworker ability at year t, $a_{-i,t}$, year fixed effects γ_t, and individual characteristics $X_{i,t}$:

$$ln(w_{i,t}) = \gamma_t + \beta_0 a_i + \beta_1 a_{-i,t} + \beta_2 X_{i,t} + \epsilon_{i,t}$$

These ability levels, a_i, are proxied with education levels in Nix [2015], and treated as unobserved variables in Cornelissen et al. [2016], which are then subsequently estimated by inverting and iterating on the least-squares objective function until convergence.

type may change as they accumulate or deccumulate human capital. Our innovation is to focus on workers who transition between employers through a spell of unemployment ('EUE' transitions). By focusing on individuals who temporarily lose their jobs, we are able to disentangle wage stickiness and any bargained wage increases coming from outside offers (e.g. Postel-Vinay and Robin [2002]) from human capital. The relative wage loss between workers with differing quality of prior colleagues lets us identify peer-effects without resorting to fixed effects. The correlation between an individual's prior wage and future quality of colleagues, in conjunction with the model's structure, allows us to estimate the degree of complementarity in production.

In particular, we estimate the following regressions. Let i index people who make an EUE transition and let t index years. Let $Y_{i,t+1}$ be the dependent variable of interest, e.g. the individual i's log annual earnings in year t+2 or the log average earnings of coworkers at individual i's primary employer in year t+2.[2] Let $\tilde{w}_{i,t}$ denote an individual's annual earnings (which we will refer to as an individual's wage in the LEHD) at his/her primary employer in year t. Our first outcome of interest is the log of this wage, $w_{i,t} = ln(\tilde{w}_{i,t})$. Our second outcome of interest is the average annual wage of the coworkers of an individual at his/her primary employer in year t, $\tilde{w}_{-i,t}$. Let $w_{-i,t} = ln(\tilde{w}_{-i,t})$ denote the log of this variable. We also show results for the startup rate and future employer size. We estimate the following specifications:

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

Our regression controls, $X_{i,t}$, include firm size, state dummies, 1-digit sic dummies, race dummies, gender dummies, education dummies, quadratics in age and tenure, as well as year fixed effects. Appendix 3.1 includes additional details regarding the regressions.

3.1 Data Description

Our empirics are based on the Longitudinal Employer-Household Dynamics (LEHD) database. 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. We combine the LEHD, by a unique scrambled identifier, with the Integrated Longitudinal Business Database (ILBD) that covers the universe of pass-through entities; however, we are only able to link information on sole proprietorships

[2]The primary employer is defined as the employer that paid the worker the most in a given year.

to the LEHD. Therefore, our set of potential observations includes the universe of privately-employed individuals and sole-proprietors between 2001 and 2008 from the 11 states for which we have LEHD data: California, Illinois, Indiana, Maryland, Nevada, New Jersey, Oregon, Rhode Island, Texas, Virginia, and Washington. Even though we use the universe of coworkers and firm characteristics in our estimation, our main regressions will focus on a smaller random sample of workers within those states.

Our main sample consists of prime age (24 to 65) males with positive earnings in year t who complete an EUE transition. 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, and then obtain a job at a different primary employer in year t+1. Since human capital transmits through workplace interaction, we further restrict our sample to workers who are at single-unit plants within a state that have between 2 and 250 employees in year t. These plants are identified by state-level firm identifiers, called State Employment Identification Numbers (SEINs), and our classification between single and multi-unit firms comes from ES-202 files supplied by the states. Since EUE spells for women may capture other forces than job loss, we restrict ourselves to males.

We count individuals as employed if they earn $1k or more in a given quarter. We count individuals as non-employed if they earn less than $1k in a given quarter. All variables which are not bounded above are winsorized at the 1% level, and nominal variables are deflated using the CPI. All reported standard errors are clustered at the SEIN level.

3.2 Summary Statistics

Table 1 provides summary statistics in our main sample of EUE job transitioners.[3] 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 earned approximately $30k per annum. 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. On average, they earned $1.6k in self-employed earnings.

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

Table 1: Summary Statistics of Job-switchers (Source: LEHD/ILBD, 2001-2008)

Variable	**Mean**	**SD**
Age	38.2	9.5
Imputed Years of Education	13.0	2.9
Tenure (Years)	2.9	2.3
Annual Earnings of Coworkers (Primary Employer)	30,436	21,743
Annual Earnings (Primary Employer)	39,365	28,267
Self-Employment Income	1,667	9,062
Number of observations	55000	

3.3 LEHD Regression Results

Table 2 shows the relationship between past coworker wages, measured in year t, and future individual wages, measured in year t+2. These correlations, in combination with the structure of the model, are used to inform the learning process. Columns (1) and (2) illustrate that across all workers, a 10% greater coworker wage is associated with a .4% greater individual wage. Columns (3) and (4) restrict the sample to those who are below the mean wage at their previous firm in year t. For that set of workers, a 10% greater coworker wage in year t is associated with a 1.3% greater individual wage in year t+2, roughly 3 times as large as the average affect. Columns (5) and (6) restrict the sample to those above the mean wage, and we see a much more muted relationship.

Table 3 shows the relationship between past individual wages, measured in year t, and future coworker wages, measured in year t+2. These correlations, in combination with the structure of the model, are used to inform the degree of worker complementarity, and thus largely determine sorting patterns. Columns (1) and (2) illustrate that across all workers, a 10% greater individual wage in year t is associated with a .99% greater coworker wage in year t+2. Columns (3) and (4) restrict the sample to those who are below the mean wage at their firm. For that set of workers, a 10% greater coworker wage is associated with a 1.29% greater future coworker wage. In columns (5) and (6) which restrict the sample to those above the mean wage, a 10% greater coworker wage is associated with a .998% greater future coworker wage. These results suggest that 'better' workers ultimately find 'better' colleagues. Viewed through the lens of our model, these point estimates imply strong productive complementarity between types.

Table 4 measures the rate at which workers spin-off their own self-employed ventures as a function of their coworker wages and their own wages. The dependent variable is an indicator if the individual transitioned into self-employment. The dummy takes a value of 1 if the individual earned more than 5k in self-employed earnings in year t+2, and earned less than 5k in self-employed earnings in year t+1. Columns (1) and (2) show that a 10% greater coworker wage is associated with a .02% lower probability of the individual spinning off a new self-employed venture. Columns (3) and (4) show that coworker wages have little impact on the spinoff rate of workers who earned below the mean wage at their firm in year t. Columns (5) and (6) show that the effect of coworker wages on spinoff rates is largely governed by workers whose earnings are above the mean.

Table 5 illustrates the relationship between coworker wages in year t and employer size in year t+2. The coefficients are therefore semi-elasticities. Column (1) shows that a 10% greater coworker wage is associated, on average, with the individual finding a job at a firm with 34 more employees in year t+2. Columns (3) and (4) show that for individuals who were below the mean wage at their firm in year t, the average point estimate implies a strong relationship between coworker wages and future employer size; however, the coefficients are insignificant at the standard thresholds. For those above the mean of their firm's wage in year t, there is a significant impact of coworker wages on future employer size, but when controls are added, that point estimate decreases and becomes insignificant.

Table 2: Future Individual wages in year t+2 as a function of coworker wages in year t (Source: LEHD/ILBD 2001-2008)

Dependent Variable	(1) Log Individual Wage, Year t+2	(2) Log Individual Wage, Year t+2	(3) Log Individual Wage, Year t+2	(4) Log Individual Wage, Year t+2	(5) Log Individual Wage, Year t+2	(6) Log Individual Wage, Year t+2
Sample	All	All	Below Mean Wage	Below Mean Wage	Above Mean Wage	Above Mean Wage
Log Average Firm Wage, Year t	0.0428***	0.0484***	0.134***	0.125***	0.00959*	0.0159***
	(0.00356)	(0.00362)	(0.0127)	(0.0127)	(0.00514)	(0.00519)
Log Average Individual Wage, Year t	0.498***	0.469***	0.299***	0.289***	0.587***	0.557***
	(0.00443)	(0.00467)	(0.0118)	(0.0118)	(0.00664)	(0.00691)
Number of Firm Employees, Year t		-0.000212***		-0.000145**		-0.000282***
		(3.71e-05)		(6.34e-05)		(4.53e-05)
Controls	N	Y	N	Y	N	Y
R-Squared	0.285	0.306	0.159	0.194	0.324	0.341
Observations	55000	55000	18000	18000	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, gender dummies, education dummies, quadratics in age and tenure, as well as year fixed effects.*

Table 3: Future coworker wages in year t+2 as a function of coworker wages in year t (Source: LEHD/ILBD 2001-2008)

Dependent Variable	(1) Log Average Firm Wage, Year t+2	(2) Log Average Firm Wage, Year t+2	(3) Log Average Firm Wage, Year t+2	(4) Log Average Firm Wage, Year t+2	(5) Log Average Firm Wage, Year t+2	(6) Log Average Firm Wage, Year t+2
Sample	All	All	Below Mean Wage	Below Mean Wage	Above Mean Wage	Above Mean Wage
Log Average Firm Wage, Year t	0.271***	0.247***	0.264***	0.235***	0.261***	0.237***
	(0.00908)	(0.00945)	(0.0249)	(0.0251)	(0.0151)	(0.0156)
Log Average Individual Wage, Year t	0.0849***	0.0990***	0.106***	0.129***	0.0902***	0.0998***
	(0.0110)	(0.0116)	(0.0232)	(0.0233)	(0.0190)	(0.0197)
Number of Firm Employees, Year t		0.000729***		0.000671***		0.000772***
		(8.75e-05)		(0.000133)		(0.000114)
Controls	N	Y	N	Y	N	Y
R-Squared	0.032	0.039	0.028	0.037	0.030	0.038
Observations	55000	55000	18000	18000	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, gender dummies, education dummies, quadratics in age and tenure, as well as year fixed effects.*

Table 4: Future Self-employment in year t+2 as a function of coworker wages in year t (Source: LEHD/ILBD 2001-2008)

Dependent Variable	(1) Transition into Self-Empl, Year t+2 (d)	(2) Transition into Self-Empl, Year t+2 (d)	(3) Transition into Self-Empl, Year t+2 (d)	(4) Transition into Self-Empl, Year t+2 (d)	(5) Transition into Self-Empl, Year t+2 (d)	(6) Transition into Self-Empl, Year t+2 (d)
Sample	All	All	Below Mean Wage	Below Mean Wage	Above Mean Wage	Above Mean Wage
Log Average Firm Wage, Year t	-0.00201**	-0.00156	0.00311	0.00212	-0.00451***	-0.00382**
	(0.000953)	(0.000988)	(0.00300)	(0.00305)	(0.00152)	(0.00156)
Log Average Individual Wage, Year t	0.0130***	0.0123***	0.00524*	0.00456	0.0175***	0.0173***
	(0.00121)	(0.00130)	(0.00271)	(0.00278)	(0.00203)	(0.00213)
Number of Firm Employees, Year t		-4.32e-05***		-4.45e-05***		-4.39e-05***
		(9.57e-06)		(1.48e-05)		(1.24e-05)
Controls	N	Y	N	Y	N	Y
R-Squared	0.003	0.005	0.001	0.005	0.003	0.006
Observations	55000	55000	18000	18000	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, gender dummies, education dummies, quadratics in age and tenure, as well as year fixed effects.*

Table 5: Future employer size in year t+2 as a function of coworker wages in year t (Source: LEHD/ILBD 2001-2008)

Dependent Variable	(1) Firm Size, Year t+2	(2) Firm Size, Year t+2	(3) Firm Size, Year t+2	(4) Firm Size, Year t+2	(5) Firm Size, Year t+2	(6) Firm Size, Year t+2
Sample	All	All	Below Mean Wage	Below Mean Wage	Above Mean Wage	Above Mean Wage
Log Average Firm Wage, Year t	340.3***	250.7***	412.7	105.6	217.4*	163.4
	(88.06)	(89.62)	(340.1)	(340.9)	(121.6)	(125.2)
Log Average Individual Wage, Year t	-1,272***	-1,212***	-1,105***	-990.7***	-1,213***	-1,174***
	(107.7)	(113.3)	(330.3)	(337.0)	(154.1)	(158.0)
Number of Firm Employees, Year t		-0.0877		0.198		-0.101
		(0.863)		(1.810)		(0.924)
Controls	N	Y	N	Y	N	Y
R-Squared	0.003	0.018	0.001	0.017	0.003	0.017
Observations	55000	55000	18000	18000	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, gender dummies, education dummies, quadratics in age and tenure, as well as year fixed effects.*

3.4 Robustness

Appendix F includes additional robustness checks and reruns our specifications on Brazilian data.[4] 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 to measure the role of general versus industry specific human capital, and we find point estimates that are similar for industry switchers, (iv) we compare our results to standard methods, including Nix [2015], 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.

4 Model

In this section, we develop a model of knowledge diffusion through team learning and job flows. The three most important features of our model include (i) workers have heterogeneous human capital, (ii) coworkers influence each other's evolution of human capital, (iii) workers search on the job, potentially starting their own team or displacing less productive members on existing teams. As workers transition between teams, they accumulate human capital, or lose human capital, as a function of the time spent with their coworkers as well as the human capital of their coworkers. The distribution of worker human capital is therefore endogenous and a function of worker mobility and the worker-coworker learning process. Our model therefore drops the assumption of fixed, exogenous, 'firm types'; rather, the distribution of 'firm types' is the set of potential teammates and the human capital of those potential teammates.

Time is discrete and infinite. Our focus is on the steady state, and so we suppress time subscripts. There is a unit measure of finitely lived individuals and a fixed exogenous measure F of entrepreneurs (we will also refer to these entrepreneurs simply as *firms*). Entrepreneurs are homogeneous, but each worker has a time-varying type indexed by $i \in \{1, \ldots, N\} \equiv \mathbb{N}$. Newborn workers draw their type from the cumulative distribution function $\Gamma : \mathbb{N} \to [0, 1]$. Each period, agents die with probability χ and are replaced by a newborn. Let μ_i denote the endogenous measure of workers of type i in the economy, $\sum_{i=1}^{N} \mu_i = 1$.

There are up to three members of a team. Every team must include an entrepreneur

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

(which we will also call a *firm*). Entrepreneurs and workers cannot produce on their own. Therefore, the only productive relationships include matches between one entrepreneur and one worker (we refer to these matches as *single worker firms*), or matches between one entrepreneur and two workers (we refer to these matches as *two worker firms*). We also assume that each entrepreneur has a symmetric, homogeneous blueprint which governs how much they can produce as a function of the number as well as the types of workers they hire; thus, unmatched entrepreneurs, which we will also call *idle firms*, are uniform in terms of productive capability. Let $f : \{0\} \cup \mathbb{N} \times \{0\} \cup \mathbb{N} \to \mathbb{R}$ summarize the production function. Single worker firms produce $f(i,0) = f(0,i)$ per period. Two worker firms produce $f(i,j) = f(j,i)$ per period.

There are three distributions of workers we must follow: u_i is the measure of type i agents who are unemployed, $e_{i,0}$ denotes the measure of type i agents at *single worker firms*, and $e_{i,j}$ is the measure of type i agents who have a coworker of type j.

Workers search for firms in a frictional market. Firms do not meet firms, and workers do not meet workers. This assumption has important consequences for the way $e_{i,j}$ is interpreted and used in the value functions. For production, the ordering of the subscripts is irrelevant, $e_{i,j} = e_{j,i}$, i.e. the number of type i agents with a j coworker is the same as the number of type j agents with a type i coworker. However, from a firm's perspective, the order of subscripts matters: they care if they meet a type i agent that has a type j coworker (an $e_{i,j}$ agent) as opposed to a type j agent that has a type i coworker (an $e_{j,i}$ agent), since those two agents have differing amounts of human capital and will contribute differently to team production and learning. On the other hand, from a worker's perspective, since they meet firms and not workers, meeting any firm with an (i,j) team or (j,i) team results in the same set of options. Table 6 summarizes the notation.

Table 6: Worker Types

Employment status (Col)	Unemployed	Work at single worker firm	Work at two worker firm				
			Coworker Type				
	u		1	2	...	N	Measure
Worker type 1	u_1	$e_{1,0}$	$e_{1,1}$	$e_{1,2}$		$e_{1,N}$	μ_1
Worker type 2							
⋮							
Worker type N	u_N	$e_{N,0}$	$e_{N,1}$	$e_{N,2}$		$e_{N,N}$	μ_N

Turning to the firms, the measure of single worker firms is $F_1 = \sum_{k=1}^{N} e_{k,0}$, the measure of two worker firms is $F_2 = \frac{1}{2}\sum_{k=1}^{N}\sum_{l=1}^{N} e_{k,l}$ (correcting for the symmetry of $e_{i,j}$), and the measure of vacant firms is given by the residual, $F_0 = F - F_1 - F_2$.[5]

When workers search, they randomly draw from the measure of firms. With probability λ_0 unemployed workers randomly sample a firm. We allow workers to search-on-the-job, so with probability λ_1, matched workers randomly sample a firm as well. Thus the total number of effective searchers is $\lambda_0 u + \lambda_1(1-u)$ (which we ensure is less than F). Since there is random search, $\frac{F_0}{F}$ is the probability an effective unit of search effort results in a contact with a vacant firm, $\frac{e_{k,0}}{F}$ is the probability an effective unit of search effort results in a contact with a single-worker firm with a type k worker, and $\frac{e_{l,k}+e_{k,l}}{2F}$ (correcting for the symmetry of $e_{i,j}$) is the probability an effective unit of search effort results in a contact with a two worker firm in which one member of the team is type k and one member of the team is type l.

What makes this entire set of contacts economically meaningful is that we allow for replacement hiring. So if a worker meets a two worker firm (a 'full' team) this may still generate two employment transitions (a new hire and new separation). With replacement hiring, two worker firms who contact a new worker may replace one of their team members with the newly contacted worker.

Let $S \in \{u, 0, \{j\}\}$ denote the workers' status (unemployed u, employed at single worker firm 0, or employed at two worker firm with coworker type j). We assume that workers' types change over time according to the conditional pdf, $g : \mathbb{N} \cup S \to \mathbb{N}$, $g_S(i_+ \mid i)$ where the subscripted S captures the dependence of a workers' learning process on their coworkers' type.

An existing productive relationship can end for a number of reasons: (i) the worker and firm can endogenously dissolve a match, (ii) the firm can replace a worker with a different worker, (iii) the worker can leave to another firm, or (iv) the match exogenously ends with probability δ.

To pin down the division of match value between workers and firms we assume that when a firm meets an unemployed worker the firm and worker Nash bargain over the marginal surplus. When a firm meets a worker who is already employed, it engages in Bertrand competition with the worker's current firm (see Postel-Vinay and Robin [2002], Dey and Flinn [2005], Cahuc et al. [2006]). Let σ denote the worker's bargaining power. Once a worker and firm are matched, they operate to maximize joint surplus.

[5]The division by two is to avoid double counting firms, e.g. an i type with a j coworker, $e_{i,j}$, is at a single firm A. By definition, from the coworker's perspective, the j type has an i coworker, $e_{j,i}$, who is also at the same firm A. There is only one firm, so when we want to recover the mass of firms by summing worker counts, when we deal with the teams, we must divide by 2.

4.1 Timing

The timing of events is as follows:

1. Exogenous separation and search (exogenous separations with probability δ, endogenous meetings between firms and workers, and replacements are made).
2. Dismissals (firms may fire any of their workers).
3. Production occurs according to $f : \{0\} \cup \mathbb{N} \times \{0\} \cup \mathbb{N} \to \mathbb{R}$.
4. Type changes are realized according to $g : \mathbb{N} \cup S \to \mathbb{N}$.
5. Worker deaths and births.

4.2 Value Functions

Let Π_0 be the value of an unmatched firm. Let $\Pi_1(i)$ be the value to an unmatched firm of hiring a type-i worker from unemployment, $\Pi_1(i,0)$ be the value to an unmatched firm of hiring a type-i worker from a firm who has no co-workers. Finally, let $\Pi_1(i,j)$ be the value to an unmatched firm of hiring a type-i worker from a firm where she is working with a type-j co-worker. Let $U(i)$ be the value of unemployment for a worker of type-i. Let $V_1(i)$ be the joint value of a match to a firm and worker of type-i. Finally, let $V_2(i,j)$ be the joint value of a match to a firm and workers of type i and j.

The value, then, of an unmatched firm is

$$\begin{aligned}\Pi_0 =& f(0,0) + \beta\Big[\sum_{i=1}^{N} \frac{\lambda_0 u_i}{F} \max\{\Pi_1(i) - \Pi_0, 0\} \\ &+ \sum_{i=1}^{N} \frac{\lambda_1 e_{i,0}}{F} \max\{\Pi_1(i,0) - \Pi_0, 0\} \\ &+ \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\lambda_1 e_{i,j}}{F} \max\{\Pi_1(i,j) - \Pi_0, 0\} + \Pi_0\Big].\end{aligned}$$

There are three events that change the value of a vacant firm. The vacant firm can meet an unemployed worker with probability $\frac{\lambda_0 u_i}{F}$. The firm Nash bargains with the worker over surplus, and the worker's bargaining weight is σ. The vacant firm can meet a worker at a single worker firm with probability $\frac{\lambda_1 e_{i,0}}{F}$. In this case, Bertrand competition implies the poaching firm must promise to pay the worker their marginal product in the old match, $\hat{V}_1(i) - \Pi_0$. Or the firm can meet a type i worker who has a coworker of type j (here, the ordering of subscripts matters). This occurs with probability $\frac{\lambda_1 e_{i,j}}{F}$. Again, Bertrand competition implies the poaching firm must pay the worker their marginal product in the old match, $\hat{V}_1(i,j) - \hat{V}_1(j)$, which is a function of whom the contacted worker (i) has as a coworker (j). The change in value to the vacant firm from hiring a type-i worker then depends on the worker's outside options:

$$\begin{aligned}
\Pi_1(i) - \Pi_0 &= (1-\sigma)[\widehat{V}_1(i) - U(i) - \Pi_0] \\
\Pi_1(i,0) - \Pi_0 &= \widehat{V}_1(i) - [\widehat{V}_1(i) - \Pi_0] - \Pi_0 = 0 \\
\Pi_1(i,j) - \Pi_0 &= \widehat{V}_1(i) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] - \Pi_0,
\end{aligned}$$

where to make the value functions more tractable, we separate them into pre-dismissal continuation values, $\{\Pi_0, V_1(i), V_2(i,j)\}$, and dismissal-stage continuation values, $\{\widehat{V}_1(i), \widehat{V}_2(i,j)\}$. The fact that there is no change in value to hiring a worker from a single worker firm is a consequence of homogeneous firms.

The value of an unmatched firm can then be written as:

$$\begin{aligned}
\Pi_0 = & f(0,0) + \beta\Big[\sum_{i=1}^{N} \frac{\lambda_0 u_i}{F}(1-\sigma)\max\{\widehat{V}_1(i) - U(i) - \Pi_0, 0\} \\
& + \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\lambda_1 e_{i,j}}{F}\max\{\widehat{V}_1(i) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] - \Pi_0, 0\} + \Pi_0\Big]
\end{aligned}$$

If a currently employed worker contacts another firm and leaves their existing match, under Bertrand competition, the value transfered to the worker by the poaching firm is equal to the worker's marginal product in their original match, and thus the ex-ante joint continuation value of the original match remains unchanged. If a worker contacts another firm and remains in their existing match, depending on the quality of the offer, the worker will simply receive a larger share of the joint surplus, but the joint continuation value remains unchanged. Thus the contacts of the worker do not impact joint continuation values. On the other hand, if a firm contacts an employed or unemployed agent and opts to hire that agent, the firm's productive capabilities will change and the joint continuation value will adjust upwards.

At the dismissal stage for single worker firms, the worker-firm pair jointly decide whether to keep operating, or separate and obtain their respective continuation values. In the case of a separation, the firm becomes vacant and thus receives a continuation value Π_0, and the worker becomes unemployed and thus receives a continuation value $U(k)$ (since $\widehat{V}_1(k)$ is a joint value).

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

Through contacts made by the firm, an existing single worker firm may form a two worker firm (a 'team'). The origin of whom the firm contacts – whether the contacted worker came from unemployment, a single worker firm, or a two worker firm – determines the cost of poaching that worker and forming the team. Let k_+ denote the realization of the worker's new type. The joint value of a single worker firm is given by,

$$\begin{aligned} V_1(k) =& f(k,0) + \beta \mathbb{E}_{k_+}\Big[\delta[U(k_+) + \Pi_0 - \widehat{V}_1(k_+)] + \chi[\Pi_0 - \widehat{V}_1(k_+)] \\ &+ \sum_{i=1}^{N} \frac{\lambda_0 u_i}{F}(1-\sigma)\max\{\widehat{V}_2(k_+, i) - U(i) - \widehat{V}_1(k_+), 0\} \\ &+ \sum_{i=1}^{N} \frac{\lambda_1 e_{i,0}}{F}\max\{\widehat{V}_2(k_+, i) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_1(k_+), 0\} \\ &+ \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\lambda_1 e_{i,j}}{F}\max\{\widehat{V}_2(k_+, i) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] - \widehat{V}_1(k_+), 0\} + \widehat{V}_1(k_+)\Big] \end{aligned}$$

At the dismissal stage for two worker firms, the worker-worker-firm trio jointly decide whether or not to continue operating or separate. The joint continuation value for two worker firms is therefore:

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

Two worker firms cannot grow any larger by assumption; they must therefore dismiss one of their workers in order to add a new member to the team, leading to replacement hiring. The joint continuation values below reflect that before the dismissal stage, the two worker firm will now be forced to make a decision about which worker to swap out of the team if they decide to hire a new worker. The joint value of a two-worker firm with a type k and type l worker is given below.

$$\begin{aligned}V_2(k,l) =& f(k,l) + \beta\mathbb{E}_{k_+,l_+}\Big[\delta(U(k_+) + \widehat{V}_1(l_+) - \widehat{V}_2(k_+,l_+)) + \delta(U(l_+) + \widehat{V}_1(k_+) - \widehat{V}_2(k_+,l_+)) \\ &+ \chi(\widehat{V}_1(l_+) - \widehat{V}_2(k_+,l_+)) + \chi(\widehat{V}_1(k_+) - \widehat{V}_2(k_+,l_+)) \\ &+ \sum_{i=1}^{N} \frac{\lambda_0 u_i}{F}(1-\sigma)\max\{0, \widehat{V}_2(k_+,i) + U(l_+) - U(i) - \widehat{V}_2(k_+,l_+), \\ &\qquad \widehat{V}_2(i,l_+) + U(k_+) - U(i) - \widehat{V}_2(k_+,l_+)\} \\ &+ \sum_{i=1}^{N} \frac{\lambda_1 e_{i,0}}{F}\max\{0, \widehat{V}_2(k_+,i) + U(l_+) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k_+,l_+), \\ &\qquad \widehat{V}_2(i,l_+) + U(k_+) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k_+,l_+)\} \\ &+ \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\lambda_1 e_{i,j}}{F}\max\{0, \widehat{V}_2(k_+,i) + U(l_+) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] - \widehat{V}_2(k_+,l_+), \\ &\qquad \widehat{V}_2(i,l_+) + U(k_+) - [\widehat{V}_2(i,j) - \widehat{V}_1(j)] - \widehat{V}_2(k_+,l_+), 0\} + \widehat{V}_2(k_+,l_+)\Big]\end{aligned}$$

Let σ denote an unemployed worker's bargaining power. The value of unemployment to a type k individual is given by,

$$\begin{aligned}U(k) =& b(k) + \beta\mathbb{E}_{k+}[(1-\chi)U(k_+) + \frac{\lambda_0 F_0}{F}\sigma\max\{\widehat{V}_1(k_+) - \Pi_0 - U(k_+), 0\} \\ &+ \sum_i \frac{\lambda_0 e_{i,0}}{F}\sigma\max\{\widehat{V}_2(k_+,i) - \widehat{V}_1(i) - U(k_+), 0\} \\ &+ \sum_i\sum_j \frac{\lambda_0 e_{i,j}}{2F}\sigma\max\{\widehat{V}_2(k_+,i) + U(j) - \widehat{V}_2(i,j) - U(k_+), \\ &\qquad \widehat{V}_2(k_+,j) + U(i) - \widehat{V}_2(i,j) - U(k_+), 0\}]\end{aligned}$$

The bargaining protocol determines promised values at the time of worker transitions, but it does not determine the wage paths that deliver those promised values. In order to map the model to the data, we must take a stance on wage setting (see Section 4.5). With linear utility, however, the way the firm delivers the value to the worker does not impact mobility decisions.

4.3 Distributions of Unemployed and Employed Workers

While the value functions for joint surplus are extremely tractable, the potential set of worker flows is significantly more complicated, since contacts that do not impact the value of match may actually result in a worker transition. 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 into four components: the distribution before dismissals occur, $u_k^{pre\ dis}$, the distribution before learning occurs, $u_k^{pre\ lean}$, the distribution before births/deaths occur, $u_k^{pre\ birth}$, and the end of period distribution u_k. There are several events that result in the flow of a type k worker into unemployment prior to the dismissal stage. The first three terms in equation (2) account for workers flowing out of unemployment by meeting vacant, single worker firms, or two worker firms. The next three terms in equation (2) account for workers flowing into unemployment by being kicked off of teams 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 (2) accounts for exogenous layoffs.

$$
\begin{aligned}
&u_k^{pre\ dis} = u_k - u_k \lambda_0 \frac{F_0}{F} \mathbb{I}(\widehat{V}_1(k) - \Pi_0 - U(k) > 0) \\
&- u_k \sum_i \frac{\lambda_0 e_{i,0}}{F} \mathbb{I}(\widehat{V}_2(k,i) - U(k) - \widehat{V}_1(i) > 0) \\
&- u_k \sum_i \sum_j (\lambda_0 \frac{e_{i,j}}{2F} \mathbb{I}(\max\{\widehat{V}_2(k,j) + U(i) - U(k) - \widehat{V}_2(i,j), \widehat{V}_2(i,k) + U(j) - U(k) - \widehat{V}_2(i,j)\} > 0)) \\
&+ \sum_i e_{k,l} \frac{\lambda_0 u_i}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l)\} > 0) \\
&\qquad \times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l)) \\
&+ \sum_i e_{k,l} \frac{\lambda_1 e_{i,0}}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l)\} > 0) \\
&\qquad \times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l)) \\
&+ \sum_i \sum_j e_{k,l} \frac{\lambda_1 e_{i,j}}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l)\} > 0) \\
&\qquad \times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l)) \\
&+ \delta \Big(e_{k,0} + \sum_i e_{k,i} \Big)
\end{aligned} \tag{2}
$$

Equation (3) describes the distribution of unemployment in the dismissal stage. The first line of equation (3) captures single worker firm dismissals, and the second and third lines capture two worker firm dismissals. For two worker firms, either both workers can be dismissed or one worker can be dismissed.

$$
\begin{aligned}
&u_k^{pre\ type} = u_k^{pre\ dis} + e_{k,0}^{pre\ dis}\mathbb{I}(\widehat{V}_1(k) < \Pi_0 + U(k)) \\
&+ e_{k,l}^{pre\ dis}\mathbb{I}[max\{U(k)+U(l)+\Pi_0 - V_2(k,l), V_1(k)+U(l)-V_2(k,l), V_1(l)+U(k)-V_2(k,l), 0\} = U(k)+U(l)+\Pi_0 - V_2(k,l)] \\
&+ e_{k,l}^{pre\ dis}\mathbb{I}[max\{U(k)+U(l)+\Pi_0 - V_2(k,l), V_1(k)+U(l)-V_2(k,l), V_1(l)+U(k)-V_2(k,l), 0\} = V_1(l)+U(k)-V_2(k,l)]
\end{aligned} \tag{3}
$$

After the dismissal stage, production occurs, and then types evolve 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'). The evolution of types for the unemployed is therefore given by,

$$u_k^{pre\ birth} = u_k^{pre\ type} + \sum_{j \neq k} u_j^{pre\ type} g_u(k \mid j) - \sum_{j \neq k} u_k^{pre\ type} g_u(j \mid k) \tag{4}$$

Agents then die at a rate of χ, and newborns enter the economy initially unemployed. The newborns draw their type from CDF Γ (let γ denote the PDF). The resulting end-of-period mass of unemployed agents is therefore given by,

$$u_k = (1-\chi)u_k^{pre\ birth} + \chi\Big(\sum_k u_k^{pre\ birth} + \sum_k e_{k,0}^{pre\ birth} + \sum_k \sum_l e_{k,l}^{pre\ birth}\Big)\gamma(k) \tag{5}$$

For a given mass of employed agents, equations (2) through (5) define a fixed point for the unemployment rate of type k agents, u_k. Appendix A includes the other two fixed point equations for the mass of employment at single worker firms, $e_{k,0}$, and two worker firms, $e_{k,l}$, which, when solved together, yield the equilibrium.

4.4 Equilibrium Definition

A steady state recursive competitive equilibrium consists of a set of policy functions describing the optimal transitions of agents, and a distribution of unemployed agents $\{u_k\}$, agents at single worker firms $\{e_{k,0}\}$, and agents at two worker firms $\{e_{i,j}\}$, such that the distributions are consistent with the policy functions.

4.5 Wages

To map the model to the data, we must take a stance on wages. Our main assumption is that firms can commit to pay wages, i.e. firms have access to perfect capital markets. The wage payment is then assumed to remain constant unless a worker changes employers or receives an outside offer. Because of the assumption of firm commitment, an individual's wage is not a function of the offers their coworker receives. The value of an employed worker at a single-worker firm with wage w is therefore given by,

$$
\begin{aligned}
W_1(k,w) =& w + \beta \mathbb{E}_{k+}[\delta U(k_+) + \lambda_1 \frac{F_0}{F} \max\{W_1(k_+,w), U(k_+), \min\{\widehat{V}_1(k_+) - \Pi_0, \widehat{V}_1(k_+) - \Pi_0\}\} \\
&+ \sum_i \lambda_1 \frac{e_{i,0}}{F} \max\{W_1(k_+,w), U(k_+), \min\{\widehat{V}_1(k_+) - \Pi_0, \widehat{V}_2(k_+,i) - \widehat{V}_1(i)\}\} \\
&+ \sum_i \sum_j \lambda_1 \frac{e_{i,j}}{2F} \max\Big\{W_1(k_+,w), U(k_+), \min\big\{\widehat{V}_1(k_+) - \Pi_0, \\
&\qquad \max\{\widehat{V}_2(k_+,i) + U(j) - \widehat{V}_2(i,j), \widehat{V}_2(k_+,j) + U(i) - \widehat{V}_2(i,j)\}\big\}\Big\} \\
&+ (1 - \delta - \lambda_1 - \chi) \max\{W_1(k_+,w), U(k_+)\}]
\end{aligned}
$$

In each of the possible continuation cases (meeting a vacant firm, a single worker firm, or team), the wage is adjusted upwards or downwards to deliver the promised continuation value. For example, if a k type worker at a single worker firm chooses to form a team with another type l worker, the worker captures their entire marginal product from the prior match ($\widehat{V}_1(k_+) - \Pi_0$), and the new wage, $\tilde{w}$, is adjusted to deliver this value:

$$W_2(k,l,\tilde{w}) = \widehat{V}_1(k) - \Pi_0$$

Appendix D includes the continuation value of two-worker firms under our wage setting assumptions.

5 Calibration

The model is calibrated so that one period is one month. The discount factor is set to $\beta = .99$, which corresponds to a discount rate of 12% per annum, or $\approx 1\%$ per month. We assume there are $N = 7$ types, so that $i \in \{1, \ldots, 7\}$. Let $\mu : \mathbb{N} \to \mathbb{R}$ map type indeces into productive abilities. We assume μ is a linear mapping and we normalize the lowest type's productive ability to 1, $\mu(1) = 1$. We estimate the highest type's productive ability, $\mu(N) = \bar{\mu}$, to match the within-occupation p90/p10 residual wage ratio.[6] In particular, we estimate the following regression on the pooled 2000-2016 CPS Merged Outgoing Rotation Groups,[7]

$$ln(w_{i,t}^{CPS}) = \sum_j \beta_j \mathbb{I}(Occ_{i,t} = j) + \Gamma X_{i,t} + \epsilon_{i,t}$$

Let the residual wage be given by the estimated residuals, $w_{i,t}^R = e^{\hat{\epsilon}_{i,t}}$. We estimate $\bar{\mu}$ to match the observed p90 to p10 wage ratio, $\frac{p_{90}(w_{i,t}^R)}{p_{10}(w_{i,t}^R)}$. 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 in the economy (non-residualized), $\frac{\bar{w}_{i,t}(UE)}{\bar{w}_{i,t}}$. Home production is given by $b(k) = .5\mu(k)$.

The death rate, χ, is .003% per month, corresponding to a 30 year working life. For newborns, we assume that the type-distribution of newborns is a truncated normal distribution over the type support with mean m_{new} and variance v_{new}. We parameterize the mean and variance parameters to target the mean wage of 24-30 year old individuals to the average wage in the economy (non-residualized), $\frac{\bar{w}_{i,t}(24-30)}{\bar{w}_{i,t}}$ and the variance to match the p90 to p10 wage ratio for this subgroup, $\frac{p_{90}(w_{i,t}^R(24-30))}{p_{10}(w_{i,t}^R(24-30))}$.

The production function is given by,

$$f(x,y) = \begin{cases} \mu(y) & if\ x = 0 \ \ (solo\ production) \\ \mu(x) + \mu(y) + \rho\mu(x)\mu(y) & if\ x > 0,\ y > 0 \ \ (team\ production) \\ \mu(x) & if\ y = 0 \ \ (solo\ production) \end{cases}$$

ρ controls the degree of complementarity between types. Using the LEHD, we estimate ρ to match the relationship between individual wages prior to layoff and future coworker

[6] The linear mapping from types to productive abilities is therefore given by, $\mu(i) = \frac{N-\bar{\mu}}{N-1} + \frac{\bar{\mu}-1}{N-1} i$.

[7] In all CPS regressions, unless otherwise stated $X_{i,t}$ includes 2-digit occupation, gender, marital status, education, and race controls.

wages. We estimate the following regression in the LEHD:

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

We set ρ so that the relationship between prior individual wages and future coworkers wages, β_2, can be reproduced by the model.

Let S= $\{u, 0, \{j\}\}$ denote the worker's status: unemployed ('u'), single worker firm ('0'), or two worker firm with coworker j. We impose a learning technology that allows workers to move up and down by one type, at most, between periods:

$$g_S(k_+ \mid k) = \begin{cases} \max\{k-1, 1\} & \text{with prob } P_L(S,k) \\ k & \text{with prob } 1 - P_L(S,k) - P_H(S,k) \\ \min\{k+1, N\} & \text{with prob } P_H(S,k) \end{cases}$$

We place several additional restrictions on the learning technology. Unemployed individuals may not gain human capital, and so $P_H(u,k) = 0$. Unemployed individuals lose human capital a rate $P_L(u,k) = \alpha_{U,L}$. We estimate $\alpha_{U,L}$ to match the elasticity of replacement wages with respect to unemployment duration in the pooled 2000-2016 CPS Merged Outgoing Rotation groups. We restrict our sample to those who have valid wage observations in the 4th and 8th waves of the survey. There is exactly one year in between those survey waves. Let $y^w_{i,8}$ denote weekly earnings in the 8th wave and $y^w_{i,4}$ denote weekly earnings in the 4th wave. Let X_i denote a static set of controls measured in the 8th wave of the survey. We isolate all completed EUE spells with corresponding duration dur_i in weeks, and we estimate the following regression:

$$\ln \frac{y^w_{i,8}}{y^w_{i,4}} = \gamma_0 + \gamma_1 \ln dur_i + \gamma X_i + \epsilon_i$$

We set $\alpha_{U,L}$ to match the coefficient, γ_1, the elasticity of the replacement wage with respect to the length of the unemployment spell.

Individuals working on their own learn at a rate $P_H(0,k) = \alpha_{0,H}$ and dislearn at a rate $P_L(0,k) = \alpha_{0,L}$. We estimate $\alpha_{0,H}$ to match average real wage growth among employed workers, and $\alpha_{0,L}$ to match the probability of a wage decline among job-to-job transitioners.

Individuals working on a team learn at the rate they would learn on their own plus a premium for having coworkers that are better, $P_H(j,k) = P_H(0,k) + \alpha_{T,H} \max\{j-k, 0\}$. They also dislearn at the rate they would dislearn on their own plus a premium for having coworkers that are worse, $P_L(j,k) = P_L(0,k) + \alpha_{T,L} \max\{k-j, 0\}$.

To estimate $\alpha_{T,H}$ and $\alpha_{T,L}$, we use data on the responsiveness of re-employment wages to prior coworker wages. In particular, we use the estimated regressions in Section 3 for individuals who were above the mean wage at their prior firm and individuals who were below the mean wage at their prior firm:

$$\bar{w}_{-i,t+2} = \beta_0^{below} + \gamma_t^{below} + \beta_1^{below} w_{i,t} + \beta_2^{below} \bar{w}_{-i,t} + \Gamma^{below} X_{i,t} + \epsilon_{i,t} \quad \forall i \ s.t. \ w_{i,t} \leq \bar{w}_{-i,t}$$

$$\bar{w}_{-i,t+2} = \beta_0^{above} + \gamma_t^{above} + \beta_1^{above} w_{i,t} + \beta_2^{above} \bar{w}_{-i,t} + \Gamma^{above} X_{i,t} + \epsilon_{i,t} \quad \forall i \ s.t. \ w_{i,t} > \bar{w}_{-i,t}$$

We estimate the same regressions in our model, and we set $\alpha_{T,L}$ to match β_2^{above} and $\alpha_{T,H}$ to match β_2^{below}.

We normalize the measure of firms, F, to a unit mass, and the remaining parameters that govern the unemployed job contact rate, the employed job contact rate, and the job destruction rate, $\{\lambda_0, \lambda_1, \delta\}$, are set to target the job finding rate, the job-to-job transition rate, and the unemployment rate. We measure these flow rates using the pooled 2000-2016 CPS Merged Outgoing Rotation Groups.

5.1 Model Fit

Table 7 summarizes the model's fit relative to the targeted moments. In order to match the observed correlation between prior individual wages and future coworker wages in the LEHD regressions, the model requires a strong degree of complementarity, $\rho = 1.2$.

The team learning parameters $\alpha_{T,L}$ and $\alpha_{T,H}$ are large, implying that a worker of type i paired with a coworker of type $i+1$, will advance to their coworker's level, ceteris paribus, in 13 months (=1/.075). The rate of dislearning implies that a worker of type i paired with a coworker of type $i-1$, will dislearn to their coworker's level, ceteris paribus, in 20 months (=1/.05). Individuals at single-worker firms learn relatively slowly, advancing one skill level every 4.16 years, ceteris paribus. They lose one skill level every 8.3 years, ceteris paribus.

The rate of dislearning among the unemployed is much stronger, with the unemployed losing one skill level every 3.3 months (=1/.3). The large rate of dislearning is identified from the elasticity of replacement rates with respect to duration of unemployment, and our estimates in the CPS imply that this elasticity is large and negative. Even with such a large dislearning parameter, the model struggles to generate the observed replacement rate elasticity.

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 and job finders. To do so, the model requires a large bargaining parameter for the workers, of .97; however, this bargaining weight is only applicable for the initial transition out of unemployment into a job. The model also requires a very low skill level among newborn types. This skill level should be interpreted as skills independent of occupation and official degrees since the wage dispersion measures we target parse these covariates out of the wage. In the current calibration, the model produces too little wage dispersion among the young, but population wage dispersion is in line with the data.

Table 7: Calibration: Model Moments vs. Data Moments (Source: CPS and LEHD)

Parameter	Value	Description	Model	Data	Moment Description	Source
$\bar{\mu}$	1.65	Top productive ability	2.875	2.912	p90/p10 wage ratio	CPS (2000-2016)
σ	0.97	Bargaining weight worker	0.787	0.783	UE wage to avg wage ratio	CPS (2000-2016)
m_{new}	1	Mean newborn type	0.824	0.820	24yo-30yo mean wage to avg wage ratio	CPS (2000-2016)
v_{new}	0.06	Variance of newborn types	2.303	2.622	24yo-30yo p90/p10 wage ratio	CPS (2000-2016)
ρ	1.2	Productive complementarity of types	0.108	0.099	Elasticity of future coworker wage to prior individual wage	LEHD (2001-2008)
$\alpha_{U,L}$	0.3	Unemployed Pr. Trans. Down	-0.001	-0.029	Duration/Replacement Rate Elasticity (Controlling for exp and age)	CPS (2000-2016)
$\alpha_{U,H}$	0	Unemployed Pr. Trans. Up	NA	NA	NA	Restriction
$\alpha_{0,L}$	0.01	Solo Employed Pr. Trans. Down	0.527	0.475	Pr. real wage declines, EE (Nominal 11.1%)	CPS (2000-2016)
$\alpha_{0,H}$	0.02	Solo Employed Pr. Trans. Up	0.074	0.064	Elasticity of wage with respect to age	CPS (2000-2016)
$\alpha_{T,L}$	0.05	Coworker Learning Pr. Trans. Down	0.068	0.016	Elasticity of future individual wage to prior coworker wage, above mean	LEHD (2002-2007)
$\alpha_{T,H}$	0.075	Coworker Learning Pr. Trans. Up	0.133	0.125	Elasticity of future individual wage to prior coworker wage, below mean	LEHD (2002-2007)
λ_0	0.33	Unemployed Job Contact Rate	0.210	0.219	UE Flow Rate, Monthly	CPS (2000-2016)
λ_1	0.45	Employed Job Contact Rate	0.017	0.017	EE Flow Rate, Monthly	CPS (2000-2016)
δ	0.0076	Layoff Rate	0.010	0.009	EU Flow Rate, Monthly	CPS (2000-2016)

5.2 Steady State Distribution of Workers

Table 8 includes the estimated type distribution. Roughly 24% of workers are of the lowest type and only about 6% are the highest type. This should be interpreted as a measure of human capital over and above formal education measures. Figure 1 is a contour map of the team distribution (Figure 2 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 6.8% of all teams are composed solely of the lowest type, i.e. (1,1) teams. Figure 3 plots the distribution of workers across all occupations and teams. Figure 3 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. An unemployed type N agent that runs into a team is much more likely to replace one of the existing team workers than an unemployed type 1 agent. The result is that type 1 agents have fewer contacts that result in a job transition, and type N agents are much more likely to be employed and on teams.

To formalize these graphical measures of sorting, Table 9 shows that the Spearman rank correlation coefficient among worker types is .49, which is consistent with several recent estimates of sorting 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 31% of the population is in a school. As we will discuss in Section 7, the wage setting mechanism disincentivizes firms from educating workers. This is because firms are unable to capture the social surplus of training workers; they are compensated for the value of the worker at their own firm, not the value of the worker to the firm that wants the worker most.

Table 8: Type Distribution

Worker Type						
1	2	3	4	5	6	7
0.24	0.18	0.16	0.14	0.12	0.10	0.06

Table 9: Sorting Moments

Sorting Moments	**Value**
Spearman rank correlation coefficient	0.49
Spearman rank correlation coefficient (UE transitioners)	0.56
Fraction in school (Coworker type higher than own type)	0.31

Figure 1: Contour map of team distribution

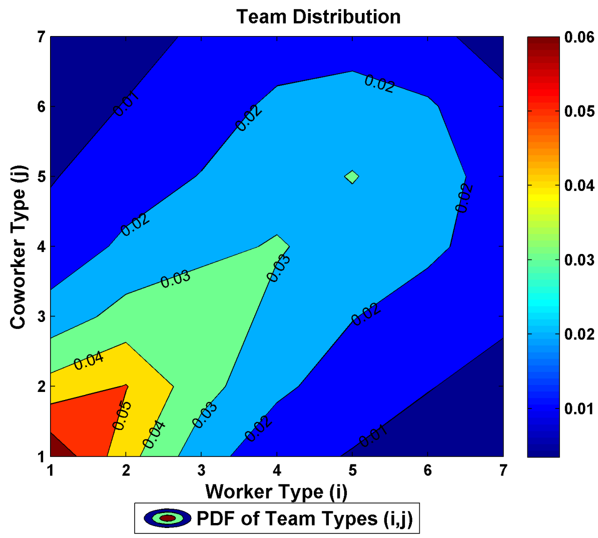


Figure 2: PDF of team distribution

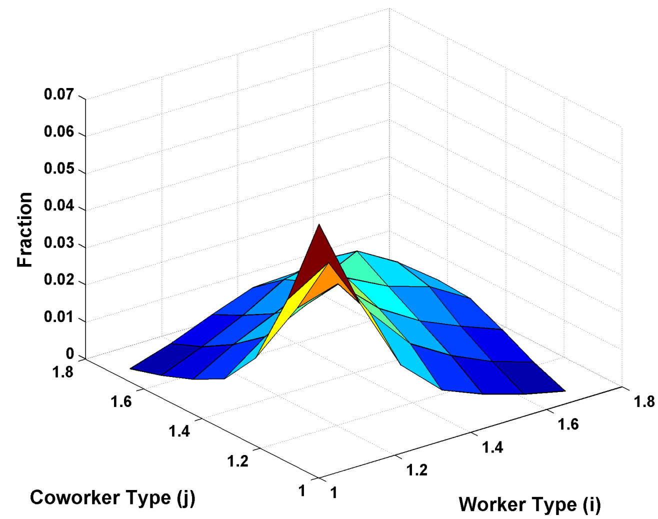
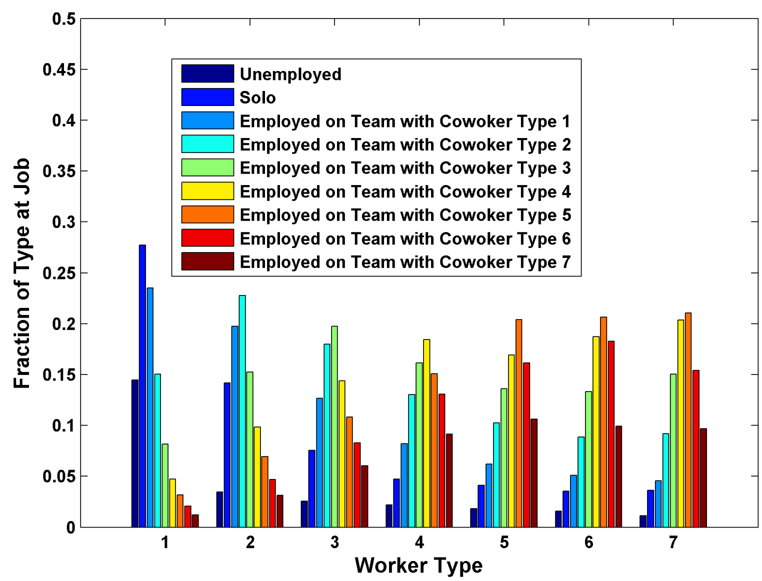


Figure 3: Distribution of Workers Across Occupations and Teams



6 Sources of Knowledge Diffusion

In this section, we use the calibrated model to understand the determinants of knowledge diffusion. We conduct three counterfactuals, (i) inhibiting all forms of coworker learning and dislearning ($\alpha_{T,L} = 0,\ \alpha_{T,H} = 0$), but still allowing for worker mobility ($\lambda_1 > 0$), (ii) inhibiting worker mobility ($\lambda_1 = 0$), but still allowing for coworker learning ($\alpha_{T,L} > 0,\ \alpha_{T,H} > 0$), and (iii) shutting down both worker mobility and coworker learning ($\lambda_1 = 0,\ \alpha_{T,L} > 0,\ \alpha_{T,H} > 0$).

Table 10 quantifies the impact of coworker peer effects and worker mobility on output. Column (2) of Table 10 shows that if we allow workers to move between jobs but not be influenced by their coworkers (either positively or negatively), then output drops by 3 percentage points. Inequality, measured by the p90 to p10 wage ratio, increases by 2.5%. The reason is that roughly the same measure of high type workers are matched with lower type workers (i.e. the fraction of workers in a school remains stable at 31%), but high types are no longer passing along their knowledge and training the lower types. This is also captured by our measure of learning (see Section 3), which focuses on the relationship between prior coworker wages and future individual wages for EUE transitioners. If we run the same regressions on simulated model output, the elasticity of a displaced individual's wage with respect to prior coworker wages is barely discernible from zero, implying a much less important role of coworker peer-effects for the equilibrium distribution of human capital.

The relatively small output loss from shutting down learning ($\alpha_{T,H} = \alpha_{T,L} = 0$) includes two offestting effects: the removal of positive peer effects (which tends to lower output), and the removal of negative peer effects (which tends to raise output). In Section 6.1, we separately isolate those two effects, and we show that the output losses of only setting the positive peer effects to zero are much larger.

Column (3) shows that relative to the baseline economy, disallowing on-the-job mobility lowers output by 17 percentage points. With a very strong degree of worker complementarity, much of the losses from worker mobility are generated by weaker sorting; using the same EUE measures of sorting as our data exercise, the elasticity of future coworker wages with respect to a displaced individual's prior own wage drops from .108 to a precisely estimated 0, implying much less sorting. The role of coworkers in determining the human capital distribution strengthens without worker mobility; again, through the lens of our data exercise, the elasticity of an individual's future own wage with respect to past coworker wages actually strengthens from .126 to .386, implying a much stronger role of coworker peer effects for the equilibrium distribution of human capital.

Lastly, these effects from worker mobility and learning are highly non-linear and interact in meaningful ways. Column (4) shows that output drops by 25% when we rule out both worker mobility and learning; this loss is roughly 25% larger than the simple sum of losses from shutting down each knowledge diffusion channel on its own.[8]

Table 10: Relative Importance of Mobility vs. Learning for Knowledge Diffusion and Output

	(1) Baseline	(2) No Learning ($\alpha_{T,H} = 0, \alpha_{T,L} = 0$)	(3) No Mobility ($\lambda_1 = 0$)	(4) No Mobility and No Learning
Output	3.90	3.78	3.23	2.91
Percent Loss Relative to Baseline		-3.1%	-17.1%	-25.3%
Wage Dispersion, p90 to p10 ratio	2.87	2.94	1.43	1.56
Measure of Population in Schools	0.31	0.31	0.28	0.23
Elasticity of coworker wage with respect to past own wage	0.108	0.140	-0.002	-0.036
Std. Error	(0.008)	(0.009)	(0.008)	(0.007)
Elasticity of own wage with respect to past coworker wage	0.126	0.022	0.386	0.544
Std. Error	(0.009)	(0.011)	(0.028)	(0.018)

6.1 Positive Peer Effects and Negative Peer Effects

In this section, we conduct two additional exercises, (i) remove the positive peer-effects alone, and (ii) remove the dislearning effects alone. Table 11 shows that eliminating positive peer effects alone reduces output in the economy by 7.4%, whereas eliminating negative peer effects would raise output by 11.2%. The Spearman rank correlation coefficient of worker types is unchanged when positive peer effects are eliminated; however, the measure of type 1 individuals (see Table 12 which includes the type distributions) increases enormously, and (1,1) teams become nearly 15% of the model economy's matches. Workers simply learn less, and the skill distribution shifts towards lower types.

On the other hand, eliminating negative peer effects lowers equilibrium sorting by over a factor of two. Type-1 workers learn rapidly, rarely fall back down the type distribution, and as a result, Table 12 shows that the entire distribution of skills shifts toward higher types. The share of type-7 individuals increases by a factor of 3 (from 6% to 18%) when dislearning

[8] The elasticity of one's own wage with respect to prior coworker wages actually strengths as the type distribution becomes near-degenerate with most workers being type 1, and thus having type 1 coworkers. The coworker and worker types become highly correlated, and the worker-wage and coworker-wage regression terms are nearly collinear.

is eliminated. More schools are formed (38% of the population is in a school) and fewer type (1,1) teams exist, since the detrimental learning effects from a type-1 partner are now zero.

Table 11: Relative Importance of Positive and Negative Peer Effects

	Baseline	Eliminate *Positive* Peer Effects ($\alpha_{T,H} = 0$)	Eliminate *Negative* Peer Effects ($\alpha_{T,L} = 0$)
Output	3.90	3.61	4.33
Percent Change Relative to Baseline		-7.4%	11.2%
Spearman rank correlation coefficient	0.49	0.45	0.18
Measure of Population in Schools	0.31	0.25	0.38
Measure of (1,1) teams	0.057	0.148	0.012

Table 12: Type Distribution After Eliminating Positive Peer Effects and Negative Peer Effects

	Worker Type						
	1	2	3	4	5	6	7
Baseline	0.24	0.18	0.16	0.14	0.12	0.10	0.06
Eliminate Positive Peer Effects ($\alpha_{T,H} = 0$)	0.33	0.22	0.16	0.11	0.08	0.06	0.04
Eliminate Negative Peer Effects ($\alpha_{T,L} = 0$)	0.15	0.11	0.11	0.13	0.15	0.16	0.18

6.2 Role of Skill Complementarity for Learning and School Formation

Figure 4 illustrates the team distribution under the assumption of no production complementaries between workers, i.e. $\rho = 0$. The value of a high-type worker is now much higher 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 5 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'). The strong worker complementarity in the baseline calibration incentivizes firms to combine high-type workers, even though no learning occurs.

[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 4: Contour map of team distribution, $\rho = 0$.

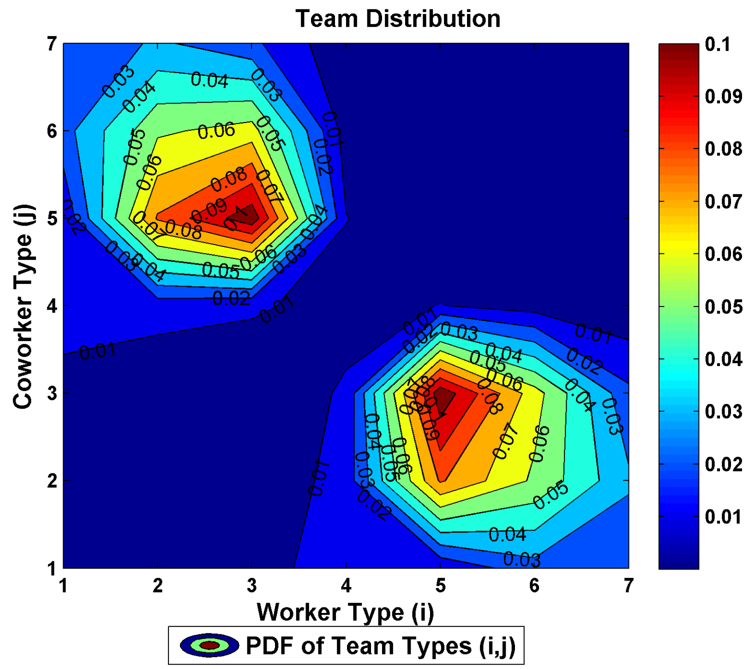


Figure 6 illustrates the percentage of team-workers in schools when there are no production complementarities, i.e. $\rho = 0$. The result is a bimodal distribution of schools. Essentially all type-1, type-2, and type-3 individuals are in schools, and thus being mentored by higher types. As soon as they learn and become a type-4, they spinoff their own school. Relative to the baseline calibration, output is 52% lower with no production complementarities.

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 schools as misallocation.

7 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 Bertrand competition. 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 paid for what

Figure 5: Baseline complementarities ($\rho = 1.2$), fraction of worker types in schools.

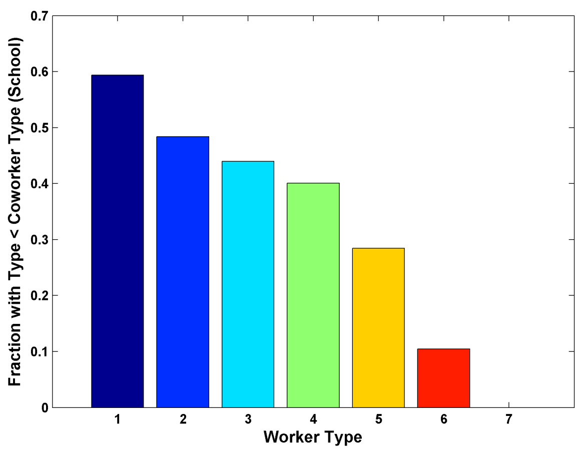
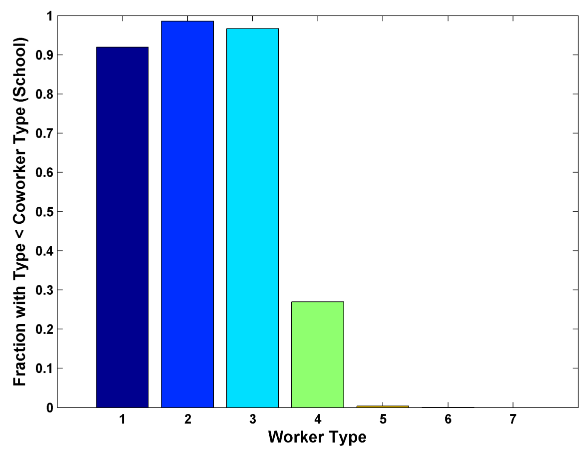


Figure 6: No complementarities in production ($\rho = 0$), fraction of worker types in schools.



the worker could produce in the existing match, rather than the value of the worker to the poaching firm, not enough (or depending on the functional forms, too much) 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 (7) to (18) (evaluated using the planner's value functions and measures of firms), and taking the mass of firms as given

$$n_0 = n - n_1 - n_2 \tag{6}$$

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 (6). 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 E includes the value functions for the planner.

7.1 Planners Allocation vs. Decentralized Allocation

Overall, the planner's allocation and decentralized allocation are quite similar, but there are some subtle and important differences. Table 13 includes the type distribution for both the social planner and decentralized economies. The planner ultimately produces a more skilled workforce. There are .1% more type-7 workers and .1% less type-1 workers. Due to the fact that the planner must make more schools to generate the more skilled workforce, the output gains from the improved type distribution are almost entirely cannibalized by the lost worker production complementaries. Table 14 illustrates that the output gain from moving to the social planner's allocation is .02%.

To understand where this modest output gain originates, Figure 7 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 7 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, breaks up low-skill (1,1) teams, and generates more (3,1) and (1,3) schools. While the gains are small, the patterns are worth discussing. 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.

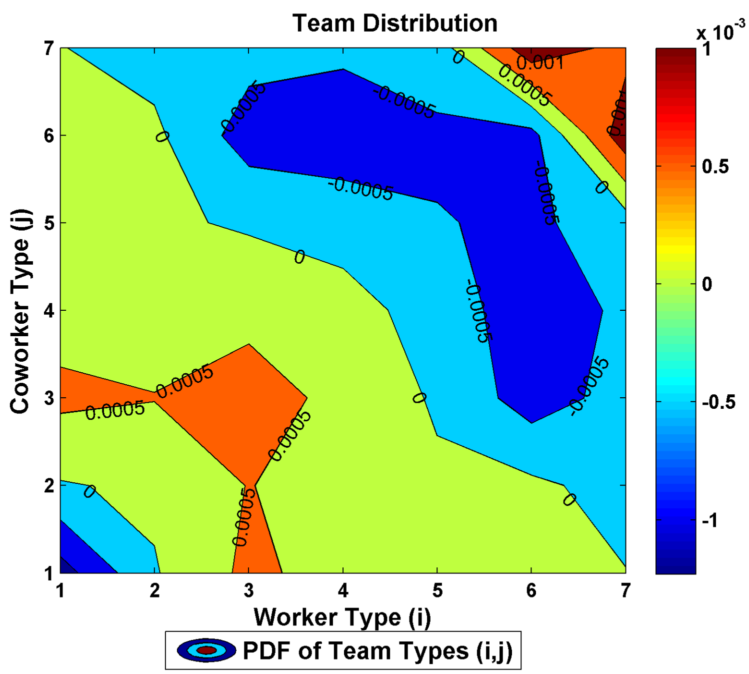
Table 13: Social Planner's Type Distribution

	Worker Type						
	1	2	3	4	5	6	7
Planner	0.240	0.184	0.157	0.138	0.123	0.096	0.063
Decentralized	0.241	0.183	0.155	0.139	0.124	0.097	0.062

Table 14: Social Planner's Output

	Decentralized	Planner
Output	3.8961	3.8967
Percent Gain Relative to Decentralized Output		0.02%

Figure 7: Contour map of *difference* in team distribution across planner and decentralized economies: Social planner's PDF of teams minus decentralized PDF of teams.



8 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. We 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.

We develop a new methodology to measure learning and sorting. Our innovation is to focus on workers who transition between employers through a spell of unemployment ('EUE' transitions). By focusing on individuals who temporarily lose their jobs, we are able to disentangle wage stickiness and any bargained wage increases coming from outside offers (e.g. Postel-Vinay and Robin [2002]) from human capital. The relative wage loss between workers with differing quality of prior colleagues lets us identify peer-effects without resorting to fixed effects. The correlation between an individual's prior wage and future quality of colleagues, in conjunction with the model's structure, allows us to estimate the degree of complementarity in production.

Using this methodolgy, we estimate the knowledge diffusion process and the degree of worker complementarities in production with micro wage data and job mobility patterns from the LEHD. We use our calibrated model to quantify the importance of learning and worker flows for US output. Our estimated parameters imply strong positive peer effects and negative peer effects. We find that eliminating positive peer effects would lower output by 7% and eliminating negative peer effects would increase output by 11%. Restricting the on-the-job search rate to zero, but still allowing for peer effects, would reduce steady state output per capita by 17%. Although the economy features inefficiencies, under the estimated benchmark parameters, welfare would only be .02% higher if we reshuffled workers to achieve the planner's allocation. Overall, our results imply that peer effects are an important determinant of US output.

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.

Gary S Becker. A theory of marriage: Part i. *The Journal of Political Economy*, pages 813–846, 1973.

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.

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, 03 2006. URL http://ideas.repec.org/a/ecm/emetrp/v74y2006i2p323-364.html.

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.

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

Steven J Davis, John Haltiwanger, Ron S Jarmin, Cornell J Krizan, Javier Miranda, Alfred Nucci, and Kristin Sandusky. Measuring the dynamics of young and small businesses: Integrating the employer and nonemployer universes. Technical report, National Bureau of Economic Research, 2007.

Matthew S. Dey and Christopher J. Flinn. An equilibrium model of health insurance provision and wage determination. *Econometrica*, 73(2):571–627, 03 2005. URL http://ideas.repec.org/a/ecm/emetrp/v73y2005i2p571-627.html.

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

Andreas Gulyas. Identifying labor market sorting with firm dynamics. Technical report, Mimeo, UCLA, 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 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 GJ Luttmer et al. 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.

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

A Expressions for Distributions of Unemployed and Employed Workers

A.1 Unemployed

Similar to the Bellman equations, we split the distribution of type k unemployed workers into four components: the distribution before dismissals occur, $u_k^{pre\ dis}$, the distribution before learning occurs, $u_k^{pre\ learn}$, the distribution before births and deaths occur, $u_k^{pre\ birth}$, and the end of period distribution u_k. 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^{pre\ dis} =& u_k - u_k \lambda_0 \frac{F_0}{F} \mathbb{I}(\widehat{V}_1(k) - \Pi_0 - U(k) > 0) \\
& - u_k \sum_i \frac{\lambda_0 e_{i,0}}{F} \mathbb{I}(\widehat{V}_2(k,i) - U(k) - \widehat{V}_1(i) > 0) \\
& \underbrace{- u_k \sum_i \sum_j (\lambda_0 \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))}_{worker\ meets\ firm\ (i,j)\ and\ is\ hired} \\
& + \sum_i e_{k,l} \frac{\lambda_0 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)}_{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))}_{firm\ replaces\ k} \\
& + \sum_i e_{k,l} \frac{\lambda_1 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)}_{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))}_{firm\ replaces\ k} \\
& + \sum_i \sum_j e_{k,l} \frac{\lambda_1 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)}_{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))}_{firm\ replaces\ k} \\
& + \delta\Big(e_{k,0} + \sum_i e_{k,i}\Big)
\end{aligned} \tag{7}$$

A.2 Solo Employed

The flows 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}^{pre\ dis} =& e_{k,0} + \delta \sum_i e_{k,i} - \delta e_{k,0} + \underbrace{u_k \lambda_0 \frac{F_0}{F} \mathbb{I}(\widehat{V}_1(k) - \Pi_0 - U(k) > 0)}_{unempl\ hired} \\
& - \underbrace{e_{k,0} \frac{\lambda_1 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_1 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_1 \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_0 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_1 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_1 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_1 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_1 \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_1 \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_1 \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{8}
$$

A.3 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}^{pre\ dis} = & e_{k,l} - \underbrace{\delta e_{k,l}}_{lose\ k} - \underbrace{\delta e_{k,l}}_{lose\ l} + \underbrace{u_k \frac{\lambda_0 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_0 \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_0 \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_1 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_1 \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_1 \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} \\
& + \underbrace{e_{k,0} \frac{\lambda_0 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_1 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_1 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_1 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_1 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_1 \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} \tag{9}
$$

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

(from previous page) ...

$$
\begin{aligned}
&- \underbrace{e_{k,l}\lambda_1 \frac{F_0}{F}\mathbb{I}(\widehat{V}_1(l) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \Pi_0 > 0)}_{coworker\ poached\ by\ idle} \\
&- \underbrace{\sum_i e_{k,l}\lambda_1 \frac{e_{i,0}}{F}\mathbb{I}(\widehat{V}_2(i,l) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \hat{V}_1(i) > 0)}_{coworker\ poached\ by\ solo\ firm} \\
&- \underbrace{\sum_i \sum_j e_{k,l}\lambda_1 \frac{e_{i,j}}{2F}\mathbb{I}(\max\{\widehat{V}_2(i,l) + U(j) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \hat{V}_2(i,j), \widehat{V}_2(l,j) + U(i) - (\widehat{V}_2(k,l) - \hat{V}_1(k)) - \widehat{V}_2(i,j)\} > 0)}_{coworker\ poached\ by\ firm\ team} \\
&- \underbrace{\sum_i e_{k,l}\frac{\lambda_0 u_i}{F}\mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l)\} > 0)}_{firm\ meets\ unempl} \\
&- \underbrace{\sum_i e_{k,l}\frac{\lambda_1 e_{i,0}}{F}\mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l), 0\} > 0)}_{my\ firm\ meets\ solo} \\
&- \sum_i \sum_j \underbrace{e_{k,l}\frac{\lambda_1 e_{i,j}}{F}\mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l)\} > 0)}_{my\ firm\ meets\ i\ from\ team\ (i,j)}
\end{aligned}
$$

A.4 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.

$$
\begin{aligned}
&u_k^{pre\ type} = u_k^{pre\ dis} + e_{k,0}^{pre\ dis}\mathbb{I}(\widehat{V}_1(k) < \Pi_0 + U(k)) \\
&+ \underbrace{e_{k,l}^{pre\ dis}\mathbb{I}[max\{U(k) + U(l) + \Pi_0 - V_2(k,l), V_1(k) + U(l) - V_2(k,l), V_1(l) + U(k) - V_2(k,l), 0\} = U(k) + U(l) + \Pi_0 - V_2(k,l)]}_{dimiss\ both} \\
&+ \underbrace{e_{k,l}^{pre\ dis}\mathbb{I}[max\{U(k) + U(l) + \Pi_0 - V_2(k,l), V_1(k) + U(l) - V_2(k,l), V_1(l) + U(k) - V_2(k,l), 0\} = V_1(l) + U(k) - V_2(k,l)]}_{dismiss\ k}
\end{aligned} \tag{10}
$$

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

$$
e_{k,0}^{pre\ type} = e_{k,0}^{pre\ dis} + \underbrace{e_{k,l}^{pre\ dis}\mathbb{I}[max\{U(k) + U(l) + \Pi_0 - V_2(k,l), V_1(k) + U(l) - V_2(k,l), V_1(l) + U(k) - V_2(k,l), 0\} = V_1(k) + U(l) - V_2(k,l)]}_{dismiss\ l} \tag{11}
$$

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

$$e_{k,l}^{pre\ type} = e_{k,l}^{pre\ dis} + \underbrace{e_{k,l}^{pre\ dis}\mathbb{I}[max\{U(k)+U(l)+\Pi_0 - V_2(k,l), V_1(k)+U(l)-V_2(k,l), V_1(l)+U(k)-V_2(k,l), 0\} = 0]}_{no\ dismissal} \tag{12}$$

A.5 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'). The evolution of types for the unemployed is therefore given by,

$$u_k^{pre\ birth} = u_k^{pre\ type} + \sum_{j\neq k} u_j^{pre\ type} g_u(k \mid j) - \sum_{j \neq k} u_k^{pre\ type} g_u(j \mid k) \tag{13}$$

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

$$e_{k,0}^{pre\ birth} = e_{k,0}^{pre\ type} + \sum_{j\neq k} e_{j,0}^{pre\ type} g_0(k \mid j) - \sum_{j \neq k} e_{k,0}^{pre\ type} g_0(j \mid k) \tag{14}$$

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

$$e_{k,l}^{pre\ birth} = e_{k,l}^{pre\ type} + \sum_{(i,j)\neq(k,l)} e_{i,j}^{pre\ type} g_j(k \mid i) g_i(l \mid j) - \sum_{(i,j)\neq(k,l)} e_{k,l}^{pre\ type} g_l(i \mid k) g_k(j \mid l) \tag{15}$$

B Birth and Deaths

$$u_k = (1-\chi)u_k^{pre\ birth} + \chi\Big(\sum_k u_k^{pre\ birth} + \sum_k e_{k,0}^{pre\ birth} + \sum_k \sum_l e_{k,l}^{pre\ birth}\Big)\Gamma(k) \tag{16}$$

$$e_{k,0} = (1-\chi)e_{k,0}^{pre\ birth} + \chi \sum_{l} e_{k,l}^{pre\ birth} \tag{17}$$

$$e_{k,l} = (1-2\chi)e_{k,l}^{pre\ birth} \tag{18}$$

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, part-time) 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/startup 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. For measures of self-employment, we use the ILBD. The ILBD is described in greater detail in Davis et al. [2007],

D Wage Setting

This section includes the continuation value for an individual on a team given our wage setting assumptions. Let $R(k,l)$ denote the probability that the type k worker in a (k,l) team is replaced. The value of an employed type k worker who has a type l coworker and a promised wage w is given by,

$$\begin{aligned}
W_2(k,l,w) =& w + \beta \mathbb{E}_{k+,l+}[(\delta + R(k_+,l_+))U(k_+) \\
&+ \lambda_1 \frac{F_0}{F} \max\{W_2(k_+,l_+,w), U(k_+), \min\{\widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+), \widehat{V}_1(k_+) - \Pi_0\}\} \\
&+ \sum_i \lambda_1 \frac{e_{i,0}}{F} \max\{W_2(k_+,l_+,w), U(k_+), \min\{\widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+), \widehat{V}_2(k_+,i) - \widehat{V}_1(i)\}\} \\
&+ \sum_i \sum_j \lambda_1 \frac{e_{i,j}}{2F} \max\Big\{W_2(k_+,l_+,w), U(k_+), \min\big\{\widehat{V}_2(k_+,l_+) - \widehat{V}_1(l_+), \\
&\qquad \max\{\widehat{V}_2(k_+,i) + U(j) - \widehat{V}_2(i,j), \widehat{V}_2(k_+,j) + U(i) - \widehat{V}_2(i,j)\}\big\}\Big\} \\
&+ (1 - \delta - \lambda_1 - \chi - R(k_+,l_+)) \max\{W_2(k_+,l_+,w), U(k_+)\}]
\end{aligned}$$

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

$$\begin{aligned}
R(k,l) =& \sum_i \frac{\lambda_0 u_i}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l)\} > 0) \\
&\times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - U(i) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - U(i) - \widehat{V}_2(k,l)) \\
&+ \sum_i \frac{\lambda_1 e_{i,0}}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l)\} > 0) \\
&\times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - (\widehat{V}_1(i) - \Pi_0) - \widehat{V}_2(k,l)) \\
&+ \sum_i \sum_j \frac{\lambda_1 e_{i,j}}{F} \mathbb{I}(\max\{\widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l), \widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l)\} > 0) \\
&\times \mathbb{I}(\widehat{V}_2(i,l) + U(k) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l) > \widehat{V}_2(k,i) + U(l) - (\widehat{V}_2(i,j) - \widehat{V}_1(j)) - \widehat{V}_2(k,l))
\end{aligned}$$

E 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:

$$Planners\ Opportunity\ Cost\ of\ Match = \sum_i \lambda_0 \frac{u_i}{n} \max\{\widehat{V}_1^P(i) - U^P(i), 0\}$$
$$+ \sum_i \lambda_1 \frac{e_{i,0}}{n} \max\{\widehat{V}_1^P(i) - \widehat{V}_1^P(i), 0\} + \sum_i \sum_j \lambda_1 \frac{e_{i,j}}{n} \max\{\widehat{V}_1^P(i) + \widehat{V}_1^P(j) - \widehat{V}_2^P(i,j), 0\}$$

The value of an unemployed worker to the planner is,

$$\begin{aligned}
&U^P(k) = b(k) + \beta \mathbb{E}_{k_+} \big[U^P(k_+) \\
&+ \lambda_0 \frac{n_0}{n} \max\{\widehat{V}_1^P(k_+) - U^P(k_+), 0\} \\
&+ \lambda_0 \sum_i \frac{e_{i,0}}{n} \max\{\widehat{V}_2^P(k_+, i) - \widehat{V}_1^P(i) - U^P(k_+), 0\} \\
&+ \lambda_0 \sum_i \sum_j \frac{e_{i,j}}{2n} \max\{\widehat{V}_2^P(k_+, i) + U(j) - \widehat{V}_2^P(i,j) - U^P(k_+), \\
&\qquad \widehat{V}_2^P(k_+, j) + U^P(i) - \widehat{V}_2^P(i,j) - U^P(k_+), 0\} \big]
\end{aligned}$$

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

$$\begin{aligned}
&V_1^P(k) = f(k,0) + \beta \mathbb{E}_{k+} \{ \widehat{V}_1^P(k+) \\
&+ \lambda_1 \frac{n_0}{n} \max\{\widehat{V}_1^P(k_+) - \widehat{V}_1^P(k_+), 0\} \\
&+ \lambda_1 \sum_i \frac{e_{i,0}}{n} \max\{\widehat{V}_2^P(k_+, i) - \widehat{V}_1^P(i) - \widehat{V}_1^P(k_+), 0\} \\
&+ \lambda_1 \sum_i \sum_j \frac{e_{i,j}}{2n} \max\{\widehat{V}_2^P(k_+, i) + U^P(j) - \widehat{V}_2^P(i,j) - \widehat{V}_1^P(k_+), \widehat{V}_2^P(k_+, j) + U^P(i) - \widehat{V}_2^P(i,j) - \widehat{V}_1^P(k_+), 0\} \\
&+ \sum_i \lambda_0 \frac{u_i}{n} \max\{\widehat{V}_2^P(i, k_+) - U^P(i) - \widehat{V}_1^P(k_+), 0\} \\
&+ \sum_i \lambda_1 \frac{e_{i,0}}{n} \max\{\widehat{V}_2^P(i, k_+) - \widehat{V}_1^P(i) - \widehat{V}_1^P(k_+), 0\} \\
&+ \sum_i \sum_j \lambda_1 \frac{e_{i,j}}{n} \max\{\widehat{V}_2^P(i, k_+) + \widehat{V}_1^P(j) - \widehat{V}_2^P(i,j) - \widehat{V}_1^P(k_+), 0\} \\
&- \underbrace{\sum_i \lambda_0 \frac{u_i}{n} \max\{\widehat{V}_1^P(i) - U^P(i), 0\} - \sum_i \lambda_1 \frac{e_{i,0}}{n} \max\{\widehat{V}_1^P(i) - \widehat{V}_1^P(i), 0\} - \sum_i \sum_j \lambda_1 \frac{e_{i,j}}{n} \max\{\widehat{V}_1^P(i) + \widehat{V}_1^P(j) - \widehat{V}_2^P(i,j), 0\}}_{opportunity\ cost\ to\ planner} \\
&+ \delta(U^P(k) - \widehat{V}_1^P(k)) \}
\end{aligned}$$

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_1 \frac{n_0}{n} \max\{\widehat{V}_1^P(k_+) + \widehat{V}_1^P(l_+) - \widehat{V}_2^P(k_+,l_+), 0\} \\
&+ \lambda_1 \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_1 \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_1 \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_1 \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_1 \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_0 \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_1 \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_1 \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_0 \frac{u_i}{n} \max\{\widehat{V}_1^P(i) - U^P(i), 0\} \\
&- \sum_i \lambda_1 \frac{e_{i,0}}{n} \max\{\widehat{V}_1^P(i) - \widehat{V}_1^P(i), 0\} \\
&- \sum_i \sum_j \lambda_1 \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}
$$

F 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

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

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 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.

F.1 Empirical Approach: Brazilian Data

Our empirical approach follows Section 3. 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 $w_{-i,t} =$

[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] Brazils 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."'

$ln(\tilde{w}_{-i,t})$ denote the log of this variable.

F.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:

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

F.3 Baseline Specifications: Brazilian Data

Table 15 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

[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

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 16 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 16 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 the initial job loss.

Lastly, Table 17 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.

F.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 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

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"

Table 15: Baseline Job-Transition Specification for Learning. 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.*

Table 16: Time Series Impact of Coworker Wages on Future Individual Wages. 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.*

Table 17: 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.*

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 18 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 workers colleagues by 10 percentage points increases that workers wages in the following year by 0.3%, which is significant at the 1% level." The discrepancy between the Brazilian and Swedish data may 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 18: 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.*