Firms as Learning Environments: Implications for Earnings Dynamics and Job Search[*]

Victoria Gregory[†]

Federal Reserve Bank of St. Louis

(click here for latest version)

January 28, 2021

Abstract

This paper demonstrates that heterogeneity in firms' promotion of human capital accumulation is an important determinant of life-cycle earnings inequality. I use administrative micro data from Germany to show that different establishments offer systematically different earnings growth rates for their workers. This observation suggests that that the increase in inequality over the life cycle reflects not only inherent worker variation, but also differences in the firms that workers happen to match with over their lifetimes. To quantify this channel, I develop a life-cycle search model with heterogeneous workers and firms. In the model, a worker's earnings can grow through both human capital accumulation and labor market competition channels. Human capital growth depends on both the worker's ability and the firm's learning environment. I find that heterogeneity in firm learning environments accounts for 40% of the increase in the cross-sectional earnings variance over the life cycle, and that this mechanism is especially important for young workers. I then show that this variation in labor market histories partially shapes the worker-specific income profiles estimated by reduced-form statistical earnings processes. Finally, because young workers do not fully internalize the benefits of matching to high-growth firms, changes to the structure of unemployment insurance policies can incentivize these workers to search for better matches.

[*]I am immensely grateful for the support and advisement from Guido Menzio, Tom Sargent, and Gianluca Violante. I also thank Serdar Birinci, Katka Borovičková, Sydnee Caldwell, Chase Coleman, Niklas Engbom, Miguel Faria-e-Castro, James Graham, Ben Griffy, Fatih Karahan, Julian Kozlowski, Jeremy Lise, Lars Ljungqvist, Elena Manresa, Abdou Ndiaye, Jonathan Payne, Laura Pilossoph, B. Ravikumar, Francisco Roldán, Yongseok Shin, Chris Tonetti, David Wiczer, and seminar participants at NYU, UC San Diego, the Federal Reserve Board, Carnegie Mellon, VMACS Junior, Insper, Yale, and UConn, as well as the Federal Reserve Banks of New York, St. Louis, Minneapolis, San Francisco, and Dallas. I am grateful to Joel Steinberg for outstanding research assistance. This study uses the Linked-Employer-Employee Data (LIAB) longitudinal model 1993-2014 (LIAB LM 9314) from the IAB. Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and subsequently remote data access under project number fdz1726. I gratefully acknowledge financial support from NYU's MacCracken Fellowship and Dean's Dissertation Fellowship. Part of this work was completed during Ph.D Dissertation Fellowships at the Federal Reserve Bank of New York and the Federal Reserve Bank of St. Louis. The views expressed herein are solely those of the author and do not necessarily reflect those of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

[†]Contact: victoria.gregory@nyu.edu

1 Introduction

Earnings dispersion across workers rises over the life cycle: there is more inequality among older workers than among younger workers. Studying the life-cycle patterns of inequality provides clues about the sources of overall earnings dispersion. This paper argues that nearly half of the rise in inequality over the life cycle is caused by differences in the firms by which workers are employed. At some firms, earnings grow systematically faster, even controlling for the growth that is specific to their employees. As different workers spend different amounts of their lives in high wage-growth firms, earnings inequality rises over the life cycle. This finding shows that persistent earnings inequality is not purely a matter of intrinsic heterogeneity among workers, but also a matter of luck.

A long literature has studied the sources of earnings inequality. An important contributor is human capital disparities across workers. These differences between individuals may be present at labor market entry and develop further as workers gain job experience.[1] Another source of earnings inequality comes from search frictions. Similar workers looking for jobs differ in the types of offers they receive. This determines whether they are able to match with high-paying firms and how much their earnings grow on the job. As a result, inequality in earnings arises due to luck in the search process.[2]

In this paper, I offer a new insight into the interactions between these two sources of inequality, and quantify how it contributes to the rise in earnings inequality over the life cycle. To do so, I delve into the sources of earnings growth. Motivated by the empirical finding that the growth rate of earnings differs across employers, I argue that luck of the draw in employer, due to search frictions, matters for a worker's growth rate of human capital. I build a search model of the labor market in which earnings can grow due to: differences in ability across workers, labor market competition, and differences in human capital promotion, or "learning environments," across firms. I use the model along with micro data to disentangle these channels and find that the firm component of human capital is a core contributor to the increase in cross-sectional earnings variance over the life cycle. I then show that these results matter for understanding the determinants of the labor income process, and for the role of policy in alleviating the inefficiencies induced by search frictions.

Using an administrative matched employer-employee data set from Germany, I show that establishments offer systematically different earnings growth rates to their workers. My data set allows me to observe the complete workforce of a subset of establishments and track workers through other jobs and through unemployment. I employ a two-way fixed effects specification to attribute growth in earnings to both worker and establishment effects. I document significant variation in

[1] See Huggett, Ventura and Yaron (2011) for an exploration of how initial human capital levels and differences in human capital growth rates across workers impact lifetime inequality.

[2] Hornstein, Krusell and Violante (2011) and Bagger et al. (2014) quantify the effect of search frictions on wage dispersion and wage growth, respectively.

earnings profiles between establishments. This finding suggests that similar workers, even workers who may have inherently similar earnings growth rates, will experience different earnings trajectories depending on the establishment they match with.

To understand the economic mechanisms that lead to this finding, I build a life-cycle search model of the labor market. The model features workers who search for jobs at firms that differ along two dimensions, productivity and learning environment.[3] These firm attributes correspond to two reasons that can explain why earnings growth rates differ between firms. The first, productivity, affects a labor market competition channel. More productive firms are better able to raise wages to prevent workers from moving to competitor firms. The second, learning environment, governs the extent to which firms promote human capital accumulation. Some firms offer faster speeds of on-the-job learning, which increases productivity, and therefore wages in both the current job and subsequent jobs.

The key features of the model generate heterogeneity in earnings profiles across workers, even for similar workers employed at different firms. Workers in the model search on and off the job, accumulating human capital via learning-by-doing as they gain job experience. The speed of human capital growth for a given worker depends temporarily on the learning environment of the firm that the worker is matched with and permanently on the worker's level of learning ability. Apart from human capital growth, a worker's earnings growth is also impacted by labor market competition. Because workers can receive outside job offers while employed, they can also obtain earnings increases by moving to better paying firms or by using competing job offers to bargain for raises at their current firm.

The model implies that workers face trade-offs between a firm's productivity and learning environment. Because their ability to accumulate human capital declines over the life cycle, workers change how they value these two components between different ages. Learning environment is highly valued early in life, when human capital accumulation is highest. Workers who match to firms with better learning environments early in life receive permanently higher earnings throughout their lifetime. As human capital accumulation declines later in life, learning environment becomes irrelevant and workers only make decisions based on the firm's productivity. These changes in trade-offs drive the job search dynamics in the model and have quantitative impacts on the major sources of earnings dispersion across workers.

Identifying the parameters of this model is challenging because there are many distinct components to earnings growth: worker ability, firm productivity, and firm learning environment. In order to discipline the parameters, I construct new moments from the data that are separately informative about each of these growth components and use an indirect inference technique to match them in the model. The first set of moments disentangles firm productivity from learning environment and worker ability by comparing the earnings growth patterns of different-aged

[3] I focus on these two because in the past literature, both have been identified as major contributors to an *individual's* life-cycle earnings growth. See the survey by Rubinstein and Weiss (2006), or for models, Bagger et al. (2014) and Bowlus and Liu (2013).

workers employed at the same firm. Assuming human capital accumulation is low for older workers, I construct an informative measure of human capital accumulation across firms by exploiting the differences in within-job earnings growth of older versus younger workers. The second set of moments disentangles the worker component from the firm components of growth. I use two-way (worker and firm) fixed effects models on earnings growth, while taking into account the biases associated with estimating these statistical models in both the data and structural model.

My parameterization method also enables me to assign a measure of learning environment to a subset of establishments in the data. I then show how this measure relates to other observable characteristics of the establishment. I find that it is not a purely industry or establishment size story: within these categories, there is still considerable variation in learning environment. I also link my measurements with survey data completed by the managers of these establishments. Here, I find that my learning environments are correlated with various aspects of the establishment's on-the-job training and apprenticeship programs. These results not only provide some context on what learning environment may be driven by, but also confirm that my measurement is actually related to characteristics that are designed to promote human capital accumulation on the job.

Next, I use the model to decompose life-cycle earnings profiles. I first examine the mean earnings profile and find that human capital drives about two-thirds of the life-cycle increase in earnings. Note also that in my setting, the human capital component implicitly contains a job search element because as workers move between firms, their speed of human capital acquisition is altered. I then shut down the worker component of human capital growth to quantify how much of a worker's human capital stock is acquired through firms. This turns out to be around 58%, despite the fact that my estimates imply higher average human capital growth on the worker rather than firm side. Here, workers are even more driven towards higher learning environments, meaning that more human capital is accumulated.

Next, I decompose the life-cycle profile of the log earnings variance. I find that the increase in earnings variance is almost entirely driven by dispersion in human capital. This result comes from both the heterogeneity in worker learning ability and firm learning environment. These two features mean that human capital grows at heterogeneous rates across workers. As a result, the dispersion in human capital increases as workers age. On the other hand, the dispersion in the components of earnings coming from labor market competition decreases. This is because workers settle into a more homogeneous set of higher paying firms and extract a larger share of the match surplus. These are the standard forces present in a textbook job ladder model.

I next assess the contribution of differences in firm learning environments and find that they account for 41% of the increase in the life-cycle earnings variance. This result comes from an experiment in which I turn off all heterogeneity in worker learning ability. In this setting, all human capital disparities arise solely due to luck in which firms workers meet. In addition, the impact of firms is is concentrated early on in workers' careers. After the first 15 years in the labor market, about 85% of earnings dispersion is due to human capital differences. Of this, half of the additional variance relative to labor market entry comes from the long-term impacts of workers' previous matches. As workers are able to catch up to each other and move to better firms, the role of firms declines.

My findings imply that firms play an important role in the formation of workers' human capital. This result sheds light on the properties of reduced-form labor income processes. Statistical models of earnings estimated from panel data on workers find that individuals appear to face different earnings profiles. These tend to be attributed to permanent worker heterogeneity, like learning ability.[4]

Using the earnings realizations generated by the model, I estimate some of the commonly-used labor income processes from the literature. The model is able to generate the relevant features of the income process. In particular, I find that the income process picks up profile heterogeneity, even in the version of the model without permanent differences in worker ability. This signifies that some of the heterogeneity in income profiles commonly attributed to worker effects come from the series of firms a worker matches with over their lifetime, which is not detectable in the panel data sets that are typically used in this context. My result also points to the existence of a new stochastic component of income growth that could alter consumption and savings behavior in incomplete markets models.

The model also has implications for worker welfare and the design of unemployment insurance (UI) policies. My findings also suggest that some of the variation in earnings growth comes about due to search and matching frictions (or differences in luck), and not due to permanent, individual heterogeneity in skill. The jobs workers accept, particularly early on in life, have permanent impacts on human capital and hence lifetime inequality. When workers have limited bargaining power, they do not fully internalize the long-term impacts of human capital accumulation. As a result, the decentralized allocation of workers to firms is inefficient. The structure of UI in the model impacts workers' ranking of firms, which means it can be used to affect the allocation.

I find that age-dependent UI schedules can improve welfare and reduce lifetime inequality relative to the benchmark model. The best UI schedules offer the highest benefit levels to young workers and reduce them with age. This UI benefit pattern induces young workers to be selective in which jobs to accept early on, particularly along the learning environment dimension. Welfare improves since the matches formed result in persistently higher lifetime earnings. Inequality is reduced by giving all workers a chance to find jobs that will boost their earnings throughout their lives. This experiment offers an example in which UI policies impact long-term outcomes, in contrast to most other settings where they are used as insurance for short-term episodes like job loss.

[4] Huggett, Ventura and Yaron (2011)'s model generates this type of profile heterogeneity through differences in worker learning ability and idiosyncratic shocks to human capital. The process of job mobility in my model offers a microfoundation to their idiosyncratic shocks.

1.1 Related literature

This paper is related to several strands of literature. Understanding the formation of human capital has been a longstanding research goal, going back to Becker (1962), Ben-Porath (1967), and Heckman (1976). A more recent complementary set of work, most notably, Herkenhoff et al. (2018) and Jarosch, Oberfield and Rossi-Hansberg (2019), explores how the quality of one's coworkers impacts human capital. This study, in contrast, views firm differences in earnings growth as coming from intrinsic firm characteristics. I also emphasize the ability of this channel to account for life-cycle features of earnings, and identify the model via establishment fixed effects. Luttmer (2014) also looks at a setting where people learn from others, but there is randomness in individual discovery. The resulting variation is likely similar to what I explore, but does not rely on search.

This work also relates to the long literature on the determinants of life-cycle earnings profiles: for a survey, see Rubinstein and Weiss (2006). There has been more recent work, such as Bagger et al. (2014) and Bowlus and Liu (2013), that decomposes the contributions of human capital growth, labor market competition, and bargaining power to life cycle earnings growth. This work performs a similar decomposition, but emphasizes how heterogeneous firm learning environments shape the earnings *variance* profile. Another recent paper by Karahan, Ozkan and Song (2019) features worker-level heterogeneity in human capital and job ladder risk and assesses the contribution of each to lifetime earnings inequality. Here, I focus more on the earnings growth components and allow for firm as well as worker effects on those.

Another paper that has explored the forces behind the earnings variance profile is Huggett, Ventura and Yaron (2011). They use exogenous human capital shocks and worker learning ability heterogeneity in a consumption/savings model to generate the increase in life-cycle variance. More broadly, the focus of the paper is to study the roles of initial conditions (level of human capital, learning ability, wealth) versus luck (shocks to human capital) in determining heterogeneity in lifetime income. In contrast, this work explores another "luck" channel that contributes to the rise in life-cycle earnings variance: the types of firms workers meet in a frictional labor market. Because my focus is only on forces that could explain the rise in variance, I only concentrate on a single initial condition, differences in learning ability.[5]

This paper also draws features from several prominent labor search models. The wage bargaining protocol adopts the sequential auction framework of Cahuc, Postel-Vinay and Robin (2006). Some of its features are also reminiscent of of Bagger et al. (2014) and Jarosch (2015). Like Bagger et al. (2014), I allow for deterministic human capital growth and adopt piece-rate wage contracts. As in Jarosch (2015), firms differ according to two dimensions: there, productivity and separation rate; here, productivity and learning environment. My model can also be cast as a special case of Lise and Postel-Vinay (2015). They allow workers and jobs to have multi-dimensional attributes,

[5] There is also a literature that relates long-term worker outcomes to observable features like graduating in a recession (Kahn (2010)) and the size of the first employer (Arellano-Bover (2019)).

and workers can acquire skills at different rates that depend on the job they are matched with. I interpret my dimensions of worker and firm heterogeneity in different ways, which restricts how they enter output and human capital accumulation, compared with Lise and Postel-Vinay (2015)'s more general setup. In addition, Engbom (2020) features a model in which workers in some jobs endogenously choose more training than in others, in line with my empirical findings.

The results of this study also connect to the vast literature that estimates statistical models of the labor income process. Some classic examples are MaCurdy (1982), Abowd and Card (1989), and Meghir and Pistaferri (2004).[6] Other studies have explored the possibility of endogenizing this labor income risk. Two potential sources are human capital (Huggett, Ventura and Yaron (2011)) and job-to-job mobility (Low, Meghir and Pistaferri (2010), Lise, Meghir and Robin (2016)). These are both present in my model and enable it to generate the main characteristics of the stochastic labor income process.

This study also closely relates to the work of Hause (1980), Baker (1997), Guvenen (2009), and Guvenen (2007) on income profile heterogeneity. Using panel data on workers' income, this research finds evidence that individuals face heterogeneous income growth rates. Here, I propose a potential source of this variation, in which the earnings profiles of different firms partially piece together a given individual's life-cycle earnings path.

Finally, my work also represents an extension to the existing body of work relating firms and labor market outcomes (Abowd, Kramarz and Margolis (1999); Card, Heining and Kline (2013)). This strand of research documents dispersion in firm-specific wage premia that impact the level of wages for all employees within the firm. In some countries, the firm component accounts for a non-trivial share of wage inequality.[7] Here, I document a similar fact, but for wage *growth*. In addition, this literature has focused on the impacts of contemporaneous firm/worker relationships. This paper introduces one mechanism in which a worker's previous employers impacts his or her earnings in the future.

There have also been studies that link firms to earnings dynamics such as Friedrich et al. (2019) and Engbom and Moser (2020). Their goal is to quantify the transmission of firm-level shocks to workers' stochastic wage processes, finding a large contribution of firms to the variance of wages over the life cycle and throughout time. In contrast, I study the persistent impacts of firm-specific wage growth trends, yet also find a substantial role for firms in accounting for the cross-sectional life-cycle variance.

The remainder of this paper proceeds as follows. Section 2 presents some motivating evidence from the data that demonstrates the extent of the establishment heterogeneity in earnings profiles. Section 3 describes the search model that allows for sources of earnings growth to differ between firms. In Section 4, I discuss how I use the data to identify the new features that my model in-

[6] For a detailed survey, see Meghir and Pistaferri (2011).

[7] Past findings that support this include Song et al. (2015) for the United States, Card, Heining and Kline (2013) for Germany, Håkanson, Lindqvist and Vlachos (2015) for Sweden, Borovičková and Shimer (2017) for Austria, Schaefer and Singleton (2016) for the United Kingdom, and Alvarez, Engbom and Moser (2015) for Brazil.

troduces. Section 5 discusses the parameter values, model fit, and provides additional context on what the learning environment measures in the data capture. Section 6 presents the model's predictions and counterfactuals for the life-cycle earnings profiles. Section 7 estimates reduced-form earnings processes from the model's earnings outcomes. Section 8 shows how changes in unemployment benefits schedules affect worker outcomes in the model. Section 9 concludes.

2 Motivating Evidence

2.1 Data Description

The main data source is an administrative matched employer-employee dataset from Germany, provided by the Research Data Center of the German Federal Employment Agency at the Institute for Employment Research (IAB). The Linked-Employer-Employee Data (LIAB) longitudinal model combines administrative employment records with unemployment benefit receipts from the German social security system. The structure of this dataset enables me to observe the complete workforce of a random sample of establishments, as well as the employment biographies of the workers employed at these sample establishments.[8] For a detailed description of this data set, see Klosterhuber et al. (2013), Fischer et al. (2009), and Heining et al. (2014).

All establishments in Germany are required to submit an annual record for each employee that worked there at any time in that year. The annual employment records in the data come in spell format and indicate the exact dates in each year during which the worker was employed at the establishment. Each record contains an establishment identifier and average daily earnings during the spell, as well as other observables like age, gender, education level, occupation, industry, and a full-/part-time indicator. The LIAB dataset contains all employment records for every worker employed at a subset of establishments between the years 2002 and 2010. Therefore, in these years I observe the complete workforce of these sample establishments. Beyond that, I get the employment biographies for each of these workers from 1993 to 2014. This means that I can track the worker through establishments not in the main sample, and through unemployment spells.

My baseline sample only uses the employment records of full-time workers, aged 20 to 60. I reorganize the data by first converting it from spell format to a monthly panel.[9] Much of the analysis involves constructing a wage for each year of job tenure. To do this, I re-aggregate all the employment spells to the annual level using the average of the wages over each 12 month interval. All wages are in real terms, deflated by the German CPI with base year 2010. In the end, the results

[8] I observe data on establishments, rather than firms, meaning two Starbucks would be considered separate entities. It is not possible in these data to aggregate the establishments into their parent firms.

[9] The level of observation in the original data set is a spell, which is at the longest 1 year for a worker who is employed at a specific establishment for the entire year. There are shorter spells that cover the partial calendar years of employment. For example, if somebody works at an establishment from August 15, 2009 to March 2, 2011, there would be 3 records for the worker: one for August 15, 2009 to December 31, 2009; another for January 1, 2010 to December 31, 2010; another for January 1, 2011 to March 2, 2011.

described in this section are derived from approximately 13.6 million worker-year observations, with approximately 1.1 million unique workers and 381,000 unique establishments.[10] For further details on the construction of the main sample and summary statistics, see Appendix A; the creation of the annual panel is discussed in Appendix B.1.

2.2 Heterogeneity in establishment-level earnings profiles

The goal of this section is to provide descriptive evidence on the heterogeneity in earnings growth rates across establishments in the data, while also controlling for differences in worker growth rates. I carry out a simple empirical exercise which shows that establishments offer systematically different earnings growth profiles.

I run regressions that are variations on the two-way fixed effects specification of Abowd, Kramarz and Margolis (1999): instead of the log wage *level* on the left-hand side, I use the *growth* in log wages. For worker i, employed at establishment j in year t, wage growth is defined as $\Delta \log w_{ijt} = \log w_{ijt} - \log w_{ij,t-1}$. I run regressions of the following form, with log wage growth as the dependent variable:

$$\Delta \log w_{ijt} = \alpha_i + \psi_j + \gamma_t + \boldsymbol{\beta} \mathbf{X_{ijt}} + \varepsilon_{ijt} \tag{1}$$

The covariates include a worker fixed-effect, α_i, an establishment fixed-effect, ψ_j, a set of year dummies, γ_t, and a set of time-varying worker and establishment characteristics $\mathbf{X_{ijt}}$. Note that all wage growth observations use only the observations of job-stayers, meaning that they do not include any wage growth that occurs during job-to-job transitions.

The fixed-effects are identified off workers who switch employers across years. When run in levels, these specifications have been widely used for understanding how innate worker and firm variation contributes to overall wage inequality. The correlation between the fixed effects has also been used to measure assortative matching. In this application, I use this method to separate worker-specific effects on wage growth, which could arise from disparities in learning ability (among others), from establishment-specific wage growth effects, the sources of which will be considered extensively in the model. Worker-specific wage growth effects have been estimated on their own using panel data on workers[11] However, less is known about the extent of the dispersion in the establishment fixed effects.[12]

To get a sense of the dispersion in these fixed effects with a simple interpretation, I first estimate a version of (1) without any of the time-varying worker or establishment observables (imposing $\boldsymbol{\beta} = \mathbf{0}$). This yields a distribution of worker and establishment fixed effects. Histograms of each

[10] The establishment count includes establishments that are not in the core sample from 2002 to 2010.

[11] For example, Guvenen (2009)'s HIP (Heterogeneous Income Profiles) process allows workers to experience different permanent growth rates in income, along with some stochastic components.

[12] To my knowledge, the only other study that has analyzed a similar specification is Sørensen and Vejlin (2011) who obtain similar results on Danish data.

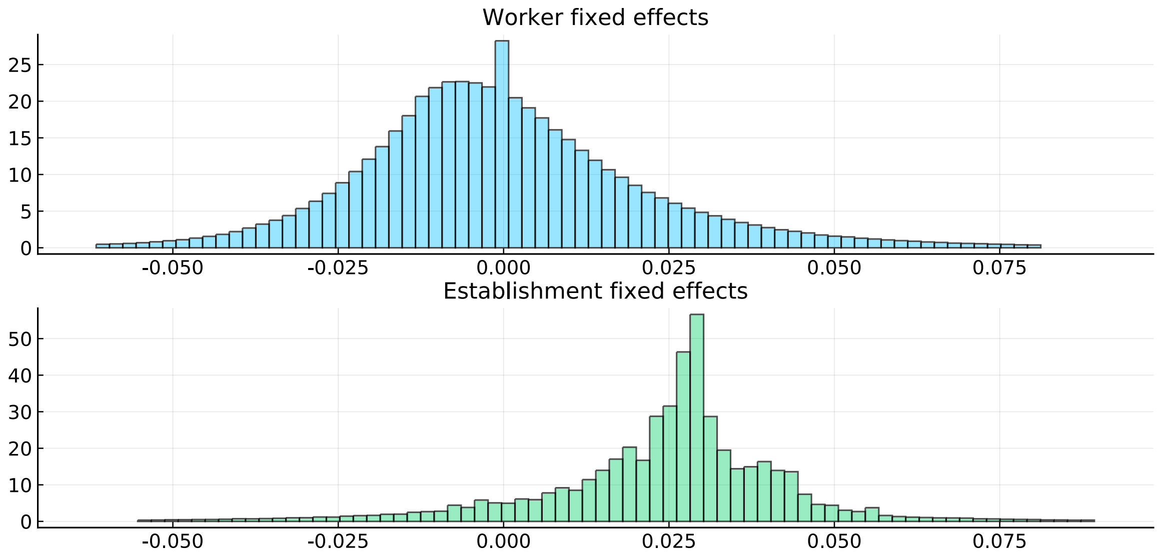

Figure 1: Distributions of worker and establishment wage growth fixed effects.
Histograms of the estimated fixed effects for workers, α_i, and establishments, ψ_j, from equation (1) with $\boldsymbol{\beta} = \mathbf{0}$. The α_i were normalized to have mean 0. For versions of these histograms broken down by education, see Figures B.1 and B.2.

are depicted in Figure 1. The detailed results for this estimation are in Table B.1. This specification with only year dummies allows for a simple interpretation of the fixed effects as the unconditional annual wage growth for a specific person or establishment.[13] I find that the dispersion in establishment effects is almost as large as the dispersion in worker effects:[14] their standard deviations are 0.0262 and 0.0242, respectively.

To better understand the role of this establishment heterogeneity on the wage growth of workers, I put an age and tenure profile in $\mathbf{X_{ijt}}$, common to all workers and establishments. I estimate the following, separately for three different education groups, high school diploma or less, vocational degree holder, and college degree holder:

$$\Delta \log w_{ijt} = \alpha_i + \psi_j + \gamma_t + \beta_1 \text{age}_{it} + \beta_2 \text{age}^2_{it} + \beta_3 \text{tenure}_{it} + \beta_4 \text{tenure}^2_{it} + \varepsilon_{ijt} \tag{2}$$

The detailed results for this estimation are in Table B.2. Figure 2 provides some examples of how the establishment fixed effect impacts wage growth. Each panel constructs cumulative earnings profiles for identical workers who are employed in establishments at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and

[13] In other words, take a worker with a fixed effect of zero employed at a firm with the average fixed effect of 0.024. This would predict an annual change in log wages of 0.024 within the spell.

[14] The well-known limited mobility bias present in AKM biases these variances upward. See Abowd et al. (2004) and Andrews et al. (2008). However, the outliers in these fixed effects distributions massively inflate the variances. For instance, removing the top and bottom 10% reduces the variance of the establishment fixed effects by five times. The difference between the 10^{th} and 90^{th} percentiles is 0.0537 for the worker fixed effects; 0.0443 for the establishment fixed effects. Moreover, the relative dispersion in the two fixed effects does not matter for this motivating exercise, whose main goal is to describe the dispersion in the establishment effects. Separating worker from establishment heterogeneity will be addressed by the structural model.

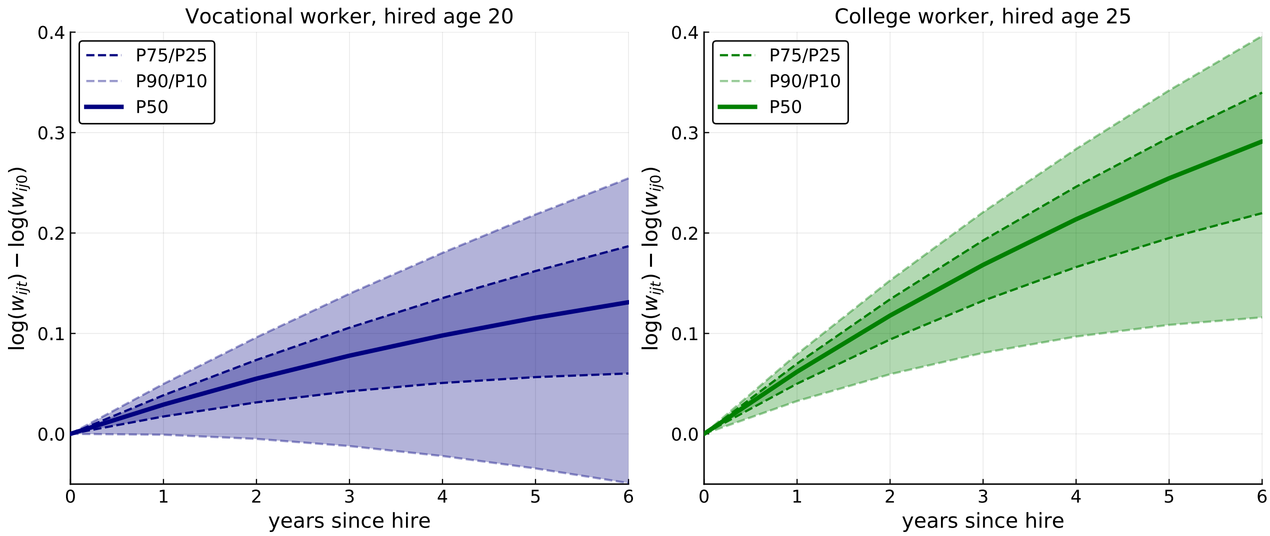

Figure 2: Establishment-specific earnings growth profiles.
Each panel depicts profiles of cumulative earnings growth as a function of tenure for workers with the same education level, age of hire, and fixed effect α_i. Estimates of the age and tenure profiles and fixed effects distributions come from equation (2). Each profile is constructed by computing the predicted values of earnings growth for each implied tenure and age horizon and taking the cumulative sum. Each series from bottom to top corresponds to the earnings growth profile of the establishment at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of the establishment fixed effect, ψ_j, distribution. For more details, see Appendix B.2.

90^{th} percentiles of the fixed effect distribution.[15] For instance, the right panel says that a college-educated worker with a given worker fixed-effect, who is hired by an establishment at age 25, can expect to see between a 0.12 and 0.40 increase in log earnings compared to their starting level after staying 6 years at each establishment. The heterogeneity in the slopes of the establishment wage profiles, captured by the establishment fixed effect, means that similar workers will face very different wage trajectories just depending on their employer. These results suggest that employers themselves, as well as frictional barriers to which establishments workers match to, may play an important role in piecing together an individual's lifetime earnings profile.

The heterogeneity and establishment earnings profiles documented thus far are purely descriptive and have no structural interpretation. Through this empirical exercise, it is not possible to identify their sources and understand how they influence the labor market outcomes of workers. The rest of this paper aims to explore the economic mechanisms that generate them, and properly quantify how much heterogeneity in earnings growth comes from workers and firms. In the next section, I introduce a structural model that formalizes how and why workers and firms exhibit different earnings growth patterns.

[15] Limited mobility bias will also inflate the variance of the distribution that the example trajectories in Figure 2 are based on. I partially address this concern by only taking into account establishments with fixed effects estimates between the 10^{th} and 90^{th}, or 25^{th} and 75^{th} percentiles. Like the results in Figure 1, the large variance is greatly influenced by the outliers. The difference between the 10^{th} and 90^{th} percentiles is 0.0510 for the worker fixed effects; 0.0446 for the establishment fixed effects. See Appendix B.3 for a discussion on how limited mobility bias affects these results.

3 Model

This section develops a search model of the labor market, featuring heterogeneity on both the worker and firm side. There is human capital accumulation, on-the-job search, and wage renegotiation.[16] They key feature is a new source of firm heterogeneity, learning environment, which impacts the speed of its workers' human capital accumulation, and thus earnings. This new dimension[17] introduces a source of persistence in earnings coming from a worker's history of matches. It also encourages workers to change their job search strategies over the life cycle.

3.1 Environment

One side of the economy consists of a unit mass of overlapping generations of workers. Workers face a deterministic life cycle, participating in the labor market from ages $t = 1, 2, \dots, T$. The age distribution is assumed to constant at all times, meaning that a fraction $1/T$ workers of age T leave the labor market each period and are replaced by new entrants. All workers are risk-neutral and consume a single homogeneous good. Their discount factor is β.

Each period, workers can be either employed or unemployed. They also differ in human capital h, and learning ability a. They enter the labor market unemployed and endowed with the same initial level of human capital,[18] but draw learning ability a from a distribution $G(a)$. Learning ability affects an individual's speed of human capital accumulation and is fixed throughout the lifetime.

Search is random and undirected. Unemployed workers receive a job offers each period with probability λ_U and employed workers receive offers with probability λ_E. A job offer is a draw from the exogenous cumulative distribution of firms, $F(\theta)$. The vector θ consists of two components, p and q, where p denotes the firm's productivity and q denotes the firm's learning environment.

Human capital accumulation is modeled as learning-by-doing. Human capital grows whenever a worker is employed, at a rate that depends on the worker's learning ability and age, as well as their employer's learning environment:

$$\log h' - \log h = (a + q)\, d(t) \tag{3}$$

This function says that the amount of human capital accumulated over a period is additive in the

[16] I adopt the sequential auction framework of Cahuc, Postel-Vinay and Robin (2006). Like in Bagger et al. (2014), earnings depend on this endogenous piece-rate as well as human capital.

[17] Like Jarosch (2015), firms differ in two dimensions. In his case, it is productivity and job security; in my case it is productivity and learning environment. Lise and Postel-Vinay (2015) is a more general environment in which workers and firms differ along multiple attributes and, like in this paper, the evolution of workers' skills depends on the firm they are matched with.

[18] This assumption does not affect the increase in the variance profile, the main focus of the paper. Having heterogeneity in initial h would only shift the level of the variance profile. It also simplifies the parameterization because it avoids having to take a stand on the joint distribution of initial (a, h).

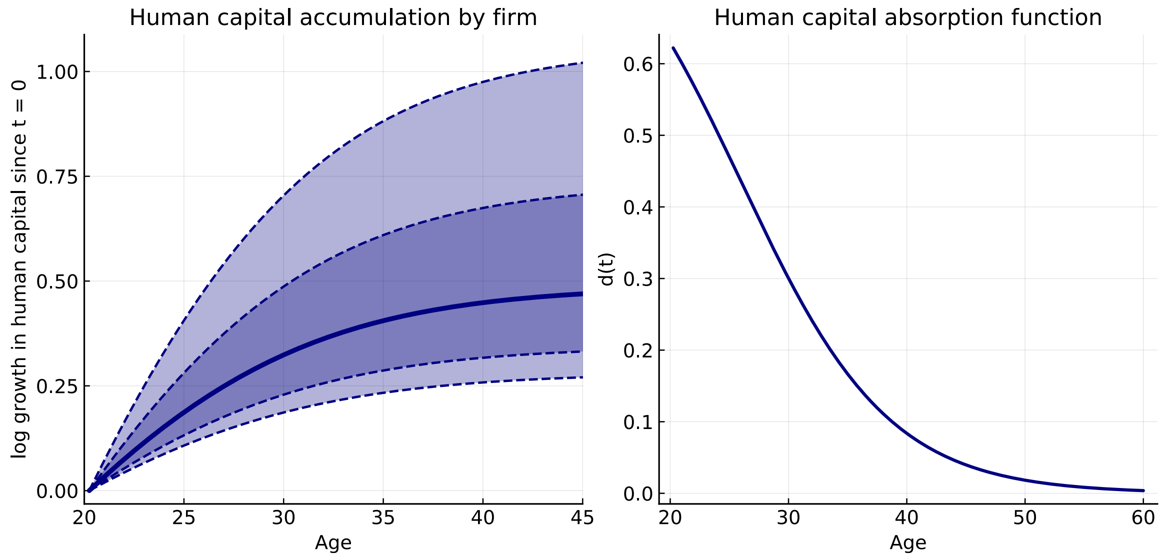

Figure 3: Human capital accumulation and absorption rate functions.
The left panel shows how human capital growth in the calibrated model differs by firm, based on equation (3). It plots the log difference in human capital at age t from the log of its starting value at age 20. Each series from bottom to top corresponds to the human capital profile of the firm at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of the distribution of q, if that worker stays at the firm. Each compares the human capital growth of a worker with the same learning ability a. The right panel shows how the absorption rate function $d(t)$ changes with age.

worker's learning ability and the firm's learning environment.[19] $d(t)$ is a human capital absorption rate function that takes the form:

$$d(t) = \frac{\nu}{1 + \exp(\gamma(t - \alpha))} \tag{4}$$

The functional form in (4) ensures that human capital grows fastest early on in the life cycle. For the same inputs, a young worker accumulates more human capital compared to an old worker. As workers age, growth gradually slows down until at some point, they can no longer accumulate human capital. This captures the effect of forces such as declines in effectiveness of learning or incentives to acquire more human capital that come with approaching retirement.[20] To see how firms and the absorption rate function impact human capital growth, some example profiles are depicted in Figure 3.

The additive portion and the absorption rate function together mean that the human capital production function in (3) will generate an increasing and concave life-cycle pattern of human capital

[19] I have experimented with a more general form of the human capital accumulation function. One can introduce a CES term over (a, q) with an elasticity of substitution parameter that determines how workers sort to firms along these dimensions. It was difficult to identify the elasticity of substitution because the model is unable to generate much sorting – there is no scarcity of jobs because they are just modeled as draws from an exogenous distribution.

[20] This is a reduced-form way to generate behavior that looks like a human capital investment model. Instead of endogenizing the decision to accumulate human capital as in Ben-Porath (1967), this functional form will impose that the earnings profile has the same shape as it would have in such a model.

for a given worker. This will help the match the life-cycle mean earnings profile in the data. The steepness of a worker's earnings profile permanently depends on learning ability and temporarily on the learning environment of the firm that the worker is matched with at a particular time. Human capital transfers perfectly across jobs[21] and does not depreciate in unemployment.[22]

If a worker and a firm form a match, they produce a flow of output ph. While employed, workers earn a flow of income phw, where w is an endogenously determined piece-rate, set according to the rules below. Matches break up with probability δ, and the worker subsequently flows to unemployment, where she earns a flow bh of income.

3.2 Wage Determination

Wages, $w \leq 1$,[23] are piece-rate contracts that determine the share of output paid to the worker. They are fixed and can only be re-bargained when workers move directly from one firm to another (a job-to-job transition) or when the worker receives a sufficiently good offer from another firm. Workers have bargaining power σ.

Let $M_t(a, h, \theta)$ denote the joint (worker + firm) value of a match between firm $\theta = (p, q)$ and a worker of learning ability a, human capital h, and age t. Additionally, let $V_t(w, a, h, \theta)$ be the value of employment to worker (a, h, t) at firm θ and current piece-rate w. Both $M_t(\cdot)$ and $V_t(\cdot)$ are increasing in all arguments.

The rules for updating the wage rate come from Cahuc, Postel-Vinay and Robin (2006) and Dey and Flinn (2005). When a worker employed at incumbent firm θ is contacted by poaching firm θ', the two firms compete for the worker. The outcome is always that the firm who values the worker the most (has the highest joint match value) gets the worker.

Specifically, one of three cases will apply. In the first case, where $M_{t+1}(a, h, \theta') > M_{t+1}(a, h, \theta)$, in which the worker is valued more by the poaching firm, the worker will move from firm θ to firm θ'. The worker's new piece-rate, w'_M will satisfy:

$$V_{t+1}\left(w'_M, a, h', \theta'\right) = M_{t+1}(a, h', \theta) + \sigma\left[M_{t+1}(a, h', \theta') - M_{t+1}(a, h', \theta)\right] \tag{5}$$

[21] I abstract from firm-specific human capital because of past literature that has shown that it is quantitatively unlikely to be as important as general human capital, at least in the long-term. Bagger et al. (2014) do the same, motivated by an argument from Lazear (2009). If all skills are general, but valued differently by each firm as in Lazear (2009), it is not necessary to model different types of human capital for Bagger et al. (2014) to achieve their objective, which was to separate human capital from job search (something I am doing as well). On the quantitative side, Nagypál (2007) finds that the impacts of match-specific human capital are only relevant during the first six months of an employment relationship. In addition, Kambourov and Manovskii (2009) also find a limited role for human capital being firm-specific.

[22] This is for simplicity and does not affect any of the main quantitative results. All that is needed in order to get workers to accept jobs with a large variety of learning environments is that human capital is always growing *less* in unemployment compared to any employment relationship.

[23] Given risk neutrality, in principle, these can be negative: workers may be willing to accept negative starting piece-rates for the opportunity to work at a firm with a particularly high productivity or learning environment.

In other words, the poaching firm delivers a wage that gives the worker the entire joint value at the incumbent firm plus share σ of the additional rents offered by matching with the poaching firm. The previous firm at which the worker was employed, θ, now becomes the worker's relevant outside option.

A second possibility is that the incumbent firm values the worker more than the poacher, but the poacher is able to offer a wage that delivers a value that is greater than the worker's current value. This happens when $M_{t+1}(a,h,\theta') < M_{t+1}(a,h,\theta)$ but there exists a w'_R that satisfies $V_{t+1}\left(w'_R, a, h', \theta\right) > V_{t+1}\left(w, a, h', \theta\right) \geq M_{t+1}(a,h,\theta')$. In this case, the worker stays at the incumbent firm θ, but the wage is re-bargained to make the worker indifferent between staying at θ and moving to θ' while extracting the full output of the match there. w'_R satisfies:

$$V_{t+1}\left(w'_R, a, h', \theta\right) = M_{t+1}(a,h',\theta') + \sigma\left[M_{t+1}(a,h',\theta) - M_{t+1}(a,h',\theta')\right] \tag{6}$$

In this case, the worker is using the outside offer to bargain an increase in the piece-rate. The worker's new relevant outside option is now firm θ', the last job offer received that was used to bargain a piece-rate increase.

The third case is that the outside offer is dominated by a previous one. In that situation, the worker discards the job offer and continues at wage w.

The wage-setting process looks like case one for unemployed workers exiting unemployment and accepting a job at firm θ. Their starting piece-rate w'_u satisfies:

$$V_{t+1}\left(w'_u, a, h', \theta'\right) = U_{t+1}(a,h) + \sigma\left[M_{t+1}(a,h,\theta') - U_{t+1}(a,h)\right] \tag{7}$$

In all cases, the new re-bargained piece-rate implicitly depends on the type of firm that the worker most recently used in a wage negotiation.[24] As workers remain continuously employed, they build up more and better quality outside offers, resulting in higher piece-rates. This process will be referred to as search capital accumulation and I will think of the on-the-job piece-rate increases as the returns to search capital.

3.3 Bellman Equations

All value functions have terminal value 0 when the worker reaches age $T+1$. The value function for an employed worker with age between 0 and T is:

[24] As shown by Cahuc, Postel-Vinay and Robin (2006), these wage setting rules microfound the bargaining game of Rubinstein (1982).

$$V_t(w,a,h,\theta) = phw + \underbrace{\beta\delta U_{t+1}(a,h)}_{\text{separation}} + \underbrace{\beta(1-\lambda_E)(1-\delta)V_{t+1}(w,a,h',\theta)}_{\text{no outside offer arrives}}$$
$$+ \underbrace{\beta\lambda_E(1-\delta)\int \max\{V_{t+1}(w'_M(\theta'),a,h',\theta'), V_{t+1}(w'_R(\theta'),a,h',\theta), V_{t+1}(w,a,h',\theta)\}dF(\theta')}_{\text{worker gets outside offer and decides whether to leave, stay and rebargain, or do nothing}} \tag{8}$$

At age t, the worker's earnings are phw. With probability δ, the worker receives a separation shock and moves to unemployment, without getting to accumulate human capital. If no separation shock and no outside offer arrives, the worker stays at firm $\theta = (p,q)$ on piece-rate w. Human capital increases to h', as governed by (3) and depends on the current firm's learning environment, q. If an outside offer from firm θ' arrives, the worker will either accept it and move to firm θ' on piece-rate w'_M, stay at θ and renegotiate the piece-rate to w'_R, or discard it. The value function in the first two cases corresponds to the promised values from the wage determination rules in (5) and (6). In any of these three cases, human capital is always updated according to the learning environment q of the incumbent firm θ.

The value function of an unemployed worker is the following:

$$U_t(a,h) = bh + \underbrace{\beta\lambda_U \int \max\{V_{t+1}\left(w'_u(\theta'),a,h,\theta'\right), U_{t+1}(a,h)\}dF(\theta')}_{\text{accept or reject job offer}} + \beta(1-\lambda_U)U_{t+1}(a,h) \tag{9}$$

Each period, unemployed workers earn benefits proportional to their human capital, bh. With probability λ_U, they receive a job offer which they can choose to accept or reject. The starting piece-rate is determined by (7). If no offer arrives or it is rejected, the worker continues to age $t+1$ with the same level of human capital h.

Finally, the value function for firm θ paired with worker (a,h,t) is:

$$J_t(w,a,h,\theta) = ph(1-w) + \underbrace{\beta\lambda_E(1-\delta)\int_{\Gamma^t_R(w,a,h,\theta)} J_{t+1}\left(w'_R(\theta'),a,h',\theta\right)dF(\theta')}_{\text{worker stays and renegotiates piece-rate}}$$
$$+ \underbrace{\beta(1-\delta)\left(1-\lambda_E\int_{\Gamma^t_R(w,a,h,\theta)} dF(\theta)\right)J_{t+1}\left(w,a,h',\theta\right)}_{\text{no outside offer arrives, or it is discarded}} \tag{10}$$

The firm's profit is what it produces, ph, minus what it pays its worker, phw, where $w \leq 1$. If the worker leaves, whether to unemployment or to a poaching firm, the firm's continuation value is zero. The continuation value will be updated if the worker receives a job offer which is used to renegotiate the piece-rate. For worker (w, a, h, t) employed at θ, this set of firms is denoted by:

$$\Gamma_R^t(w, a, h, \theta) = \{\theta' | M_{t+1}(a, h', \theta) > M_{t+1}(a, h', \theta'), V_{t+1}\left(w_2', a, h', \theta\right) > V_{t+1}\left(w, a, h', \theta\right) \geq M_{t+1}(a, h', \theta')\}$$

In other words, the worker renegotiates his or her wage at firm θ if firm θ values the worker more than firm θ', but θ can afford to match the maximum value that θ' can offer. If no outside offer arrives, or it is discarded, the match continues with the same piece-rate and human capital is updated according to firm θ's learning environment.

3.4 Joint Match Value

The joint value of the match, $M_t(a, h, \theta)$, is defined as the sum of the worker's value function and the firm's value function: $M_t(a, h, \theta) = V_t(w, a, h, \theta) + J_t(w, a, h, \theta)$. Using equations (8) and (10) and the surplus splitting rules, (5), (6), and (7), we arrive at the following recursive expression for the joint value:

$$\begin{aligned} M_t(a, h, \theta) = ph + \underbrace{\beta\delta U_{t+1}(a, h)}_{\text{unemployment}} + \underbrace{\beta(1-\delta)\left(1 - \lambda_E \int_{\Gamma_M^t(a,h,\theta)} dF(\theta)\right) M_{t+1}\left(a, h', \theta\right))}_{\text{worker stays; wage may be rebargained}} \\ + \underbrace{\beta(1-\delta)\lambda_E \int_{\Gamma_M^t(a,h,\theta)} \left[M_{t+1}\left(a, h', \theta\right) + \sigma\left(M_{t+1}\left(a, h', \theta'\right) - M_{t+1}\left(a, h', \theta\right)\right)\right] dF(\theta')}_{\text{worker moves to firm with higher match value}} \end{aligned} \tag{11}$$

Aside from the impact of human capital accumulation, the joint match value only changes if the worker transitions to unemployment or to another firm, in the set $\Gamma_M^t(a, h, \theta)$, defined as:

$$\Gamma_M^t(a, h, \theta) = \{\theta' | M_{t+1}(a, h', \theta') > M_{t+1}(a, h', \theta)\}$$

This is the set of firms who value the worker more than firm θ. In this case, the updated joint value reflects the value as delivered by the wage setting rule in equation (5). If the worker remains at firm θ, the joint match value is only updated to reflect human capital accumulation, even if the piece-rate changes. This is because changes in the piece-rate are only reflective of a transfer of value from firm to worker. As a consequence, this value function does not depend on the piece-rate.

This function characterizes all job acceptance decisions in the economy and thus is sufficient for determining the steady-state allocation of workers to firms. Once this equation is solved, piece-rates can be backed out from the wage setting equations (5), (6), and (7).

3.5 Equilibrium

Given exogenous distributions $F(\theta)$ and $G(a)$, a **stationary equilibrium** is:

(a) a match value function $M_t(a, h, \theta)$, an employed worker value function $V_t(w, a, h, \theta)$, an unemployed worker value function $U_t(a, h)$, and a firm value function $J_t(w, a, h, \theta)$,

(b) a piece-rate function which depends on (w, a, h, t) and the types of the incumbent and poaching firms, (θ, θ'),

(c) steady state distributions of workers over the state variables (w, a, h, θ, t)

such that:

(i) the value functions are the solutions to the Bellman equations,

(ii) the piece-rates evolve according to the wage setting rules,

(iii) the distributions evolve according to the wage setting rules and the transitions determined by the joint match value function,

(iv) and inflows of worker $(w, a, h, \theta, t) =$ outflows of worker (w, a, h, θ, t)

3.6 Properties of the Model

Next, I discuss a few key implications of this model.

Sources of earnings growth. Earnings in the model are phw. The dynamics of each component play into the growth of overall earnings.

The firm productivity component, p will change whenever the worker makes a job-to-job transition. Thus, the model accounts for the notion of “high" and “low" paying firms, or the job ladder in the traditional sense. In conjunction with each job-to-job transition, as well as on-the-job, the piece-rate w grows as workers obtain outside offers. Increases in the piece-rate reflect increases in search capital as workers accumulate and improve on the outside options they use to renegotiate. This source of growth introduces an indirect effect of firm tenure on earnings growth because workers with longer tenure tend to have received better outside offers throughout the employment spell.

The bargaining setup induces backloaded w contracts. As long as the firm has some bargaining power, it is optimal for it to backload wages and pay the worker well below their marginal product initially. This is because the firm anticipates that the worker will get outside offers in the future and can raise wages to retain them only when they have a credible threat to leave. As a consequence, matches with higher joint values will exhibit steeper earnings profiles, because these firms are better able to compete with others. In a model without learning environment heterogeneity, the slope of a firm's earnings profile would be dictated only by p. But here, much of the future value of the match also depends on the firm's learning environment q through its impact on human capital accumulation. As a result, for a given level of p, workers are willing to accept lower starting

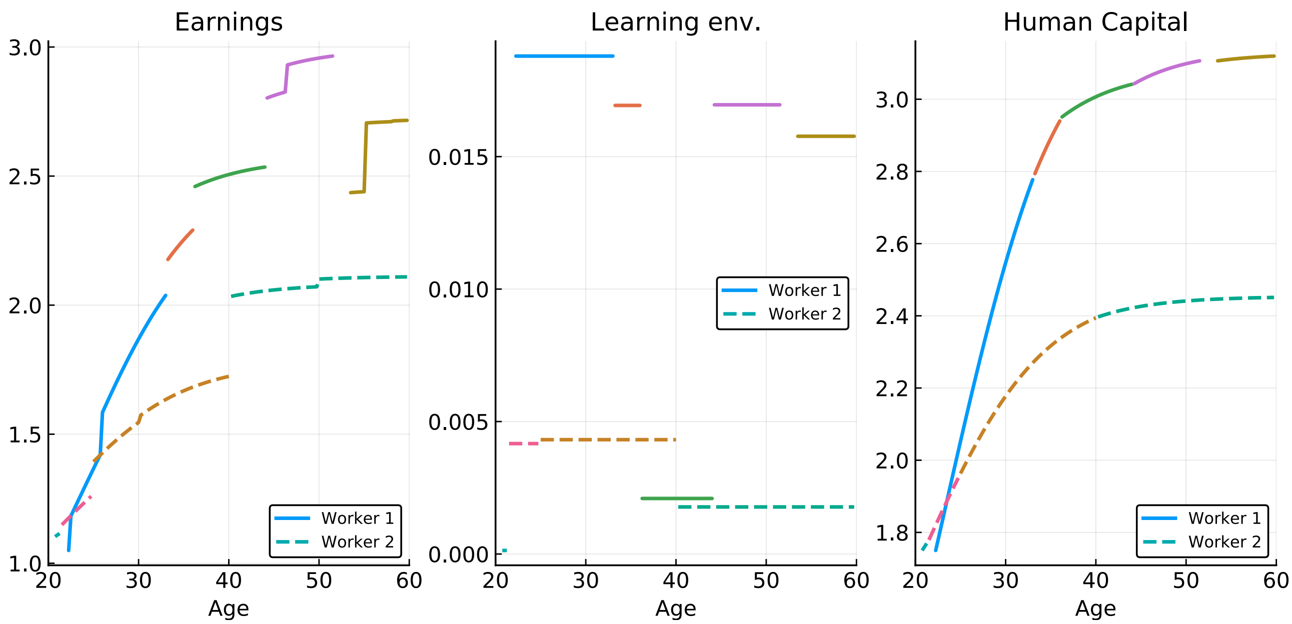

Figure 4: Example paths for workers with same learning ability.
The left panel shows earnings paths for two workers in the solid and dashed lines. Both have the same learning ability, but receive a different series of shocks over their lifetimes. Each separate color represents a spell in a different firm. Gaps (can be seen best in the human capital paths) represent unemployment spells. The middle panel shows the corresponding learning environments of the firms the workers match to. The right panel shows each worker's human capital profile.

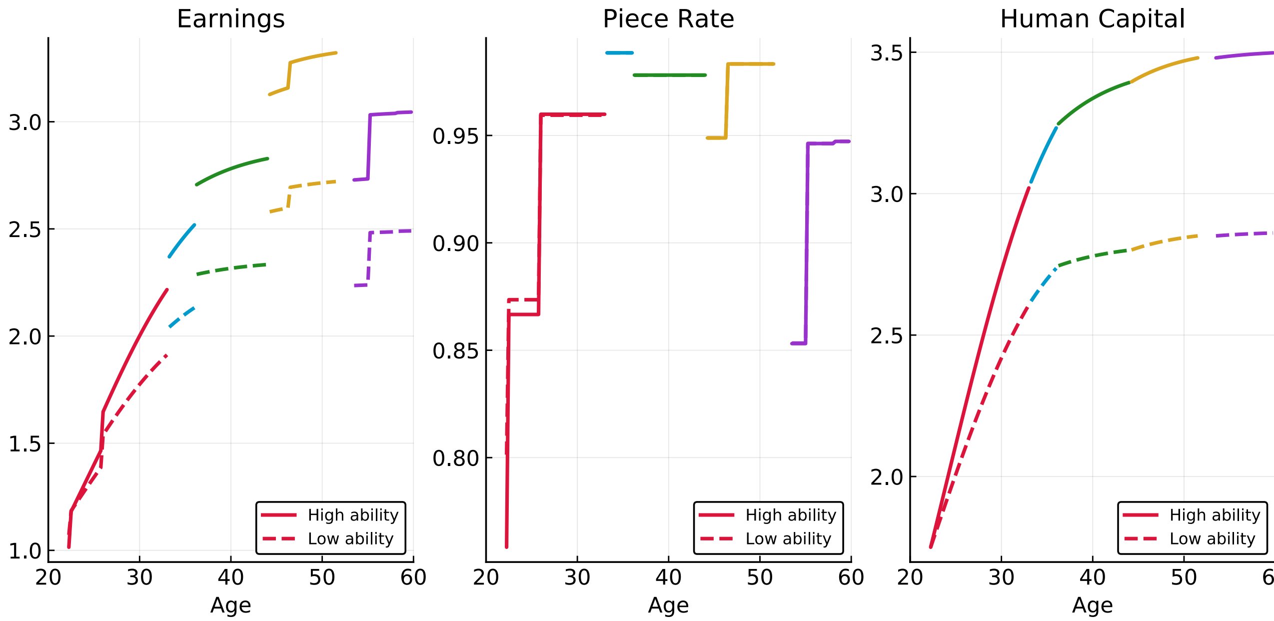

Figure 5: Example paths for workers with different learning abilities, but with the same shocks. The left panel shows earnings paths for two workers. The worker in the solid line has a high a, whereas the worker in the dashed line has a low a. The workers receive the same sets of shocks, and thus meet the same firms. Each separate color represents a spell in a different firm. Gaps (can be seen best in the human capital paths) represent unemployment spells. The middle panel shows the corresponding piece-rates. The right panel shows each worker's human capital profile.

piece-rates in order to work at a firm with a better q.

Finally, increases in human capital, h, directly feed into earnings.[25] Human capital growth depends on the worker's age and learning ability and the firm's learning environment.

To understand the effects of age and learning environment on human capital and earnings, see Figure 4. This figure shows the earnings profiles in the model of two workers with the same learning ability, but who receive different shocks (job offers and separation shocks). Each different-colored line segment represents a spell at a different firm. The middle and right panels also show the learning environment of each match and the corresponding worker's human capital profile. Because Worker 1 consistently meets firms with better learning environments at young ages, his earnings profile is steeper than Worker 2's. In addition, human capital growth flattens for both workers at older ages, regardless of the firms they match with. The outcomes depicted here are an example of the novel mechanism that I explore in this paper: the labor market outcomes of *ex-ante* identical workers differ solely because of luck in which kinds of firms they match with. The main driver is disparities in the firms' learning environments. This is one channel that will impact the life-cycle variance profile of earnings.

[25] An alternative modeling choice would have been to allow for human capital to impact earnings only through increases in the piece-rate. In this setting, earnings would not directly depend on human capital, but increases in the piece-rate would also reflect human capital accumulation since the last outside offer.

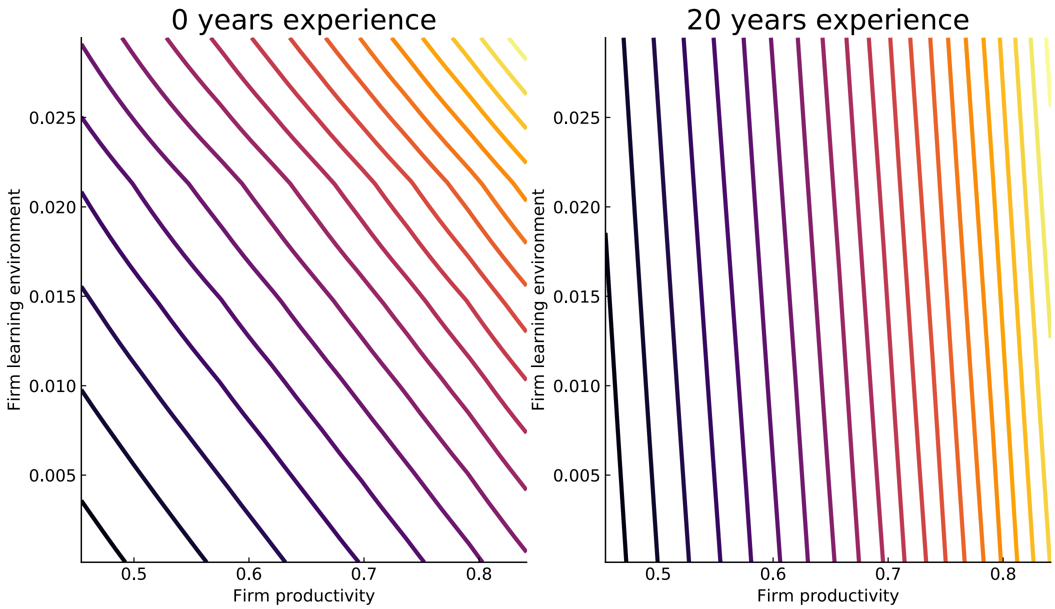

Figure 6: (p, q) indifference curves
Traces out indifference curves in (p, q) space where p corresponds to firm productivity and q to firm learning environment. These are generated for the baseline calibration outlined in Section 4. The contours are defined based on the joint match value as a function of (p, q), which is increasing in both arguments. Worker learning ability and human capital are fixed at the same arbitrary values in the two panels.

Figure 5 highlights the impacts of learning ability and disentangles the sources of earnings growth at different ages. In this figure, there are two agents with different abilities but they meet the same firms over their lifetime. Late in the life cycle, earnings changes during a spell are solely driven by changes in the piece-rate. For example, the changes in the piece-rate that the workers get in the last firm is solely driven by an increase in the piece-rate, but not human capital. In contrast, at younger ages, both the piece-rate and human capital play a role. This insight is going to guide the identification strategy which will aim to separate the contributions of search capital and human capital within firms. Additionally, the earnings of the high ability worker are always growing faster than those of the low ability worker, even though they are always employed by the same firm. This idea will also be used in the identification to quantify the extent of worker versus firm effects on human capital growth.

Job search. The decision to accept a job offer in the model is solely dependent on the comparison between the joint value of the current job (or unemployment) versus the new job. An important determinant of the present value of the match is the growth in human capital that the worker expects to receive over the match. Because human capital growth is highly age dependent, the model creates trade-offs across firms that vary over a worker's life cycle.

Figure 6 illustrates this. Each contour traces out an indifference curve over firm characteristics productivity and learning environment. In each panel, the learning ability and human capital of the worker is held constant; the left panel is for a new labor market entrant and the right panel is for a worker with 20 years of experience in the labor market. The indifference curves earlier in life are flatter than those later on. When young, workers highly value a firm's learning environment because the ability to accumulate human capital diminishes over the life cycle. It is important to match to a high q firm early on in order to receive permanently higher earnings throughout life. When workers are much older, however, the learning environment of the firm becomes irrelevant. Workers only weigh job acceptance decisions by p, generating the nearly vertical indifference curves in the right panel. These changes in workers' job acceptance strategies are crucial for the model's life-cycle dynamics and are the channel through which policies impact the allocation of workers to firms.

4 Identification

Identifying the parameters that determine the outcomes of this model is challenging. An individual's earnings growth contains both worker and firm components. The firm-specific components come from the firm's productivity and learning environment, governed by the joint distribution $F(\theta)$. The worker-specific component comes from the distribution of learning ability, $G(a)$.

Because in the model, the relevance of the sources of earnings growth changes over the life cycle, my identification strategy exploits the differences in earnings patterns over the life cycle. I use an indirect inference method in which I match a set of reduced-form moments in both the model and the data. Using insights from the model, I show why these particular moments are separately informative about the distributions of worker and firm heterogeneity.

I construct two sets of moments. The first aims to separate firm productivity and learning environment. It relies on comparing the earnings growth patterns of different-aged workers within the same establishment (this contains three sub-steps).[26] The second group of moments adds information that helps inform the relative amounts of worker and firm heterogeneity.

I discuss each of these in detail below, and then describe how to identify the more standard features of the model.

4.1 Residual earnings growth of young workers by establishment

First step: establishment-specific returns to search capital. In the first step, I construct a measure of the returns to search capital by establishment. This comes down to estimating establishment-specific earnings profiles with respect to tenure for older workers who are hired out of an unemployment spell. The logic is that this group of workers starts off with the same outside option

[26]One could think of proxying human capital with labor market experience and search capital with job tenure and constructing firm earnings profiles as a function of experience and tenure. However, it is difficult to separate the two effects because whenever tenure increases, experience increases by the same amount.

(unemployment) and can no longer accumulate human capital. As a result, any earnings growth they experience should come only from accumulation of search capital. Through the lens of the model, I am isolating the growth of w in earnings, phw. Because these are estimated on job-stayers p is not growing, and with the assumption on human capital, h is not growing. This idea is also depicted visually in Figure 5. At older ages, human capital is no longer growing, but earnings within-job can still grow if a worker receives a sufficiently good outside offer.

The assumption that little to no human capital is accumulated late in life has been used by Heckman, Lochner and Taber (1998) and later, Huggett, Ventura and Yaron (2011) and Lagakos et al. (2018), among others. The reasoning comes from declines in productivity or the proximity to retirement for older workers. Using the earnings of older workers has enabled these authors to estimate certain parameters of structural models.

The restriction that these workers must be in their first job after an unemployment spell also relies on economic theory. When workers lose their jobs in a sequential auction model like this one, their bargaining position is wiped out. All workers who find new jobs start from the same negotiation benchmark, the value of unemployment, and must get raises by obtaining outside offers. Using workers coming out of an unemployment spell ensures that all of them start from the same benchmark and that workers at the same establishment have in expectation received similar outside offers conditional on tenure. Combined with the older workers restriction, this ensures that the earnings growth of this group of workers is informative about only the establishment-specific returns to search capital.

In order to implement my strategy, I restrict older workers in the data to be ages 50 and up. I locate UE transitions by taking workers who are employed in a given month, but were receiving unemployment benefits in the previous month, or who were not registered in the social security system for between 21 and 365 days. For more information on the construction of this sample and for summary statistics, see Appendix C.1.

To construct the establishment-specific returns to search capital, I run the following random coefficients model:

$$\Delta \log \text{earnings}_{ijt} = \alpha^j + \beta_1^j \text{tenure}_{it} + \beta_2 \text{tenure}_{it}^2 + \varepsilon_{ijt} \tag{12}$$

Importantly, both the intercept and first-order coefficient on tenure differ across establishments, which allows for rich variation in the profiles. Moreover, rather than running OLS separately by establishment, I use a random coefficients model. These statistical models construct earnings profiles for specific establishments by using information about the profiles of other establishments, a concept known as partial pooling. This reduces the noise involved with having small or relatively homogeneous workers employed in some establishments: for establishments like this, the estimates will shift towards the overall mean profile.[27] The statistical model assumes that (α^j, β_1^j)

[27]Nevertheless, I do apply some weak establishment size restrictions on the establishments I include in this regres-

are distributed bivariate normal across the population of establishments and estimates the mean and covariance matrix of that distribution. Using the predicted values of the coefficients, I can construct predicted values for the amount of earnings growth coming from search capital accumulation at each establishment and at each tenure horizon. These will be used in the next step.[28]

Second step: establishment-specific returns to human capital. In the second step, I focus on younger workers in order to construct a set of moments that is informative about the returns to human capital. The main idea is to isolate growth in h in phw. As before, I will be using job stayers, so p is not growing. To separate h from w, I use the establishment's returns from search capital estimated in the first step. The residual is informative about (but necessarily equal to) human capital growth patterns in the establishment.

To ensure that I focus on the part of the life cycle with the fastest human capital growth, the first several years in the labor market, I make restrictions on the ages of the workers and the job spells I include. I want to include a worker's first "real" job in the labor market and use this starting point to construct a measure of experience.[29] I restrict each first job to be the first time the worker appears in the data set, is in a reasonable age range depending on the education of the worker,[30] and lasts at least 90 days. For more information on the construction of this sample and for summary statistics, see Appendix C.2.

Using these job spells, I first compute annual earnings growth at each year of tenure on the job, $\Delta \log \text{earnings}_{ijt}$. Then, using the predicted values, $(\hat{\alpha}_j, \hat{\beta}^1_j)$, obtained in the first step, I can construct a measure informative about how much earnings growth the worker should be getting from search capital accumulation based on the establishment that employs the worker. I construct the residual part of earnings growth as $\Delta \widehat{\log \text{earnings}}_{ijt} = \Delta \log \text{earnings}_{ijt} - \hat{\alpha}_j - \hat{\beta}^1_j \text{tenure}_{it} - \hat{\beta}_2 \text{tenure}^2_{it}$. Finally, like in step 1, I construct establishment-specific human capital returns profiles by estimating another random coefficients model on the residuals:

$$\Delta \widehat{\log \text{earnings}}_{ijt} = \gamma_j + \delta^1_j \text{experience}_{it} + \delta_2 \text{experience}^2_{it} + \varepsilon_{ijt} \tag{13}$$

See the second column of Table C.2 for the full details of the estimates in (13). The moments that I target are based on the cumulative earnings growth profiles constructed from (13). Using the pre-

sion. I include only establishments who have at least 5 worker spells for whom I can compute yearly wage growth, and for which one of these spells lasts at least 5 years. The resulting pattern of earnings profiles looks similar to establishment-by-establishment OLS where I use a stricter sample selection with establishments who have at least 5 workers who stay longer than 5 years. See Table C.2 and Appendix C.4 for more on the comparison between random coefficients and OLS.

[28] Refer to the first column of Table C.2 for the full details of the estimates in (12).

[29] I am careful here about using experience rather than age because in the model, human capital only starts growing upon labor market entry which is interpreted as age 20 for everyone. In the data, this not necessarily the case, so I want to ensure that I am capturing for everyone the right place in the life cycle where human capital (or job experience) starts to grow.

[30] Between ages 17 and 21 for workers with less than a high school degree; 19 and 23 for workers with a high school degree or vocational degree; 21 and 27 for workers with both a high school degree and vocational degree; 24 to 30 for workers with a college degree; 19 to 23 for workers with a missing education level.

dicted values $(\hat{\gamma}_j, \hat{\delta}_j^1, \hat{\delta}_2)$, I compute predicted earnings growth from human capital at experience horizons 1 to 10 for each establishment. Cumulating these gives a predicted cumulative earnings growth at each establishment for each horizon. I target the 10th percentile, 90th percentile, and mean of these distributions at each horizon, obtaining 30 moments. See Appendix C for more details.

Because these moments pick up variation in human capital growth patterns across establishments, they are informative about the distribution of learning environments, q. The shape of these profiles is also informative about γ and α, which control the shape of the absorption rate function. γ determines how steeply human capital declines and α controls the age around which the decline is steepest. Their values are restricted to ones such that human capital growth is zero past age 50, which is necessary to match the assumptions I made with the data.

Third step: correlation between returns to human capital and search capital. I also use the results from step 1 and step 2 to inform how correlated productivity and learning environment should be in the joint distribution $F(\theta)$. I consolidate the results from each step to give me just one measure of each per establishment. To do this, I construct for each establishment the predicted earnings growth that comes from (12) and (13) at tenure and experience levels 1 through 10. I take the average over these 10 to obtain one measure of search capital returns and one measure of human capital returns per establishment. I then target the establishment-level correlation coefficient.[31]

4.2 AKM moments

The moments described in Section 4.1 do not account for variation in worker ability. For instance, if high ability workers sort into high learning environment firms, this will be picked up in these moments. Next, I add additional moments designed to separate the effects of workers versus firms on earnings growth.

For this, I use the AKM two-way fixed effects model from Section 2. I run the following regression in both the data and the model:

$$\Delta \log \text{earnings} = \alpha_i + \psi_j + \gamma_t + \beta_1 \text{age}_{it} + \beta_2 \text{age}_{it}^2 + \beta_3 \text{tenure}_{it} + \beta_4 \text{tenure}_{it}^2 + \varepsilon_{ijt} \tag{14}$$

In order to ensure that the moments from the data and the model are comparable, I need to address the limited mobility bias present in AKM. The AKM fixed effects are identified off of workers who switch firms. When there is a small number of switchers in the data, the fixed effects can only be identified for these workers and for the firms that they visit. Moreover, each of these workers is only employed by a few firms, and each firm may only employ a small number of workers. As a result, the fixed effects estimates become noisy estimates of the true types. This biases the

[31] See Appendix C.3 for further details.

variances of these distributions upward. In addition, the covariance between the fixed effects is biased downward. Intuitively, if a worker fixed effect is overestimated, the firm fixed effect will be underestimated, and vice-versa.

This bias exists in both the data and the model, but to different degrees. The first difference comes from the length of worker histories. The model-simulated data is a balanced panel with exactly 40 years of data per worker. The LIAB data is an unbalanced panel. It only contains on average 14 years of data per worker, with each worker employed in 3 establishments on average. The differences in the lengths of worker histories impacts the precision of the estimates of the worker fixed effects – the more firms I observe a worker in, the better the estimate. To put the model and the data on equal grounding, I randomly truncate the worker histories in the model-simulated data so that I only use on average 14 years of data per worker and 3 establishments per worker when estimating (14).[32]

The establishment sizes also affect the magnitude of the bias. The smaller the establishments, the larger the bias. In the model, workers are matched to firms one-to-one, so to mimic multi-worker firms, I group similar firms together. I bin firms based on their quantiles in the $p \times q$ distribution. I choose the number of quantiles small enough so that I have on average 9 workers per firm like in the real data set.

I target the relative variance of the worker-fixed effect to the establishment fixed-effect, $\mathrm{var}(\alpha_i)/\mathrm{var}(\psi_j)$, and their correlation, $\mathrm{corr}(\alpha_i, \psi_j)$. The variances inform the dispersion in the distributions of worker learning ability and firm learning environment. The correlation informs the degree of sorting on the (a, q) dimension.

4.3 Firm productivity and bargaining power

Unlike the distribution of learning ability, q, heterogeneity in firm productivity, p, is a more standard feature of my model. It informs the dispersion of firm wage premia, and along with the bargaining power σ, how backloaded wages are due to labor market competition forces. Like Jarosch (2015) and Bagger et al. (2014), I will use moments about between- and within-job earnings growth to discipline these. But because early in life these moments are also influenced by human capital accumulation, I will focus on moments from workers above age 50. These moments give me cleaner measures of the forces of the model that are unrelated to human capital.

For between-job growth, I target the mean earnings growth upon a job-to-job transition.[33] For within-job growth, I use the average annual earnings change for job-stayers, the average growth from start to end of a job spell, and the ratio of starting wages to average wages.

[32] There exist econometric methods to correct for the bias. These include Borovičková and Shimer (2017), Bonhomme, Lamadon and Manresa (2019), Andrews et al. (2008), and Kline, Saggio and Sølvsten (2019). They vary in their underlying assumptions and limitations, but they appear to be computationally costly to re-do in the structural model (in keeping with the indirect inference approach) each time a new parameter vector is evaluated.

[33] Because at later ages, the model is not capable of generating job-to-job transitions with wage cuts, I target the mean wage growth of workers aged 50+, conditional on getting a wage increase.

4.4 Transition and replacement rates

I use standard labor market flow moments to identify the arrival rates of job offers on and off the job, λ_E and λ_U, respectively. The job-to-job transition rate identifies λ_E and the job-finding rate identifies λ_U. Because all separations are exogenous, δ can be taken straight from the data. b, the level of unemployment benefits is chosen to match the net replacement rate in Germany as reported by the OECD.[34] In the model, I compare the average earnings in unemployment with the average earnings in employment. The model's period is quarterly, and workers participate in the labor market for 40 years (corresponding to ages 20 through 60 in the data), implying $T = 160$. I follow Herkenhoff et al. (2018) by setting β to a 15% annual discount rate to avoid of negative wages – a high discount factor reduces the desire of workers to take on very steep wage profiles without having to add the complication of concave utility.

4.5 Parameterization

I use Pareto distributions to parameterize a and q.[35] These distributions have shape parameters χ_a and χ_q, respectively. The distribution of p is parameterized as a Beta distribution with parameters χ_p^1 and χ_p^2, with the support shifted by χ_p^3. To further characterize the joint distribution of firms, I introduce a correlation between firm attributes (p, q), called ρ. All together, draws from $F(\theta)$ are correlated draws from the marginal distributions of p and q, the Beta and Pareto distributions defined above.[36] ρ is identified by the correlation of the two firm attributes obtained from each step of the procedure outlined in step 3 of Section 4.1.

Because the moments that identify the parameters are more complicated than just simple functions of the data, the calibration is reminiscent of the indirect inference procedure of Gourieroux, Monfort and Renault (1993). This is a simulated method of moments procedure where the moments can be parameters from reduced form econometric models. These reduced form models, called auxiliary models, can be misspecified, but should be informative about the structural parameters of the model. The structural parameters are chosen to minimize the distance between the auxiliary models estimated on real data and the same ones estimated on simulated data. In this case, the auxiliary models are the cumulative residual earnings growth moments described Section 4.1, the relative variances and the correlation coefficient from the AKM model in growth rates in Section 4.2, as well as the simpler moments described in Sections 4.3 and 4.4. On top of this, I also target the increase in the variance of earnings (from its minimum point to age 60) because I decompose this in the results section as a starting point for my main counterfactual.

[34] See the table here.

[35] Both are shifted so that their support starts at 0 rather than 1.

[36] In practice, I take draws from a bivariate standard normal with correlation ρ, map the draws back to quantiles of the standard normal, and then map these quantiles to the corresponding points in the marginal distributions of p and q. To discretize this for the model solution, I need to assign probabilities to each point on a 2-D grid over these variables. I do this a similar way, making use of approximations of the cdf of the bivariate normal.

Parameter	Value	Target	Data	Model
χ_q	79	$\frac{\text{var}(\alpha_i)}{\text{var}(\psi_j)}$	1.09	1.09
χ_a	77	resid. growth young workers	see Figure 7	see Figure 7
γ	0.04	increase life cycle variance	0.14	0.13
α	6 years	$\text{corr}(\alpha_i, \psi_j)$	-0.49	-0.37
ν	0.87			
χ_p^1	5	between job growth	0.154	0.119
χ_p^2	10	within job growth	0.005	0.003
χ_p^3	0.35	growth over entire spell	0.0	0.013
σ	0.66	starting to avg. wage ratio	0.795	0.848
ρ	-0.02	correlation (p, q) in $F(\theta)$	-0.1202	-0.1276
δ	0.0186	EU rate	0.0186	0.0186
λ_U	0.225	UE rate	0.2025	0.1960
λ_E	0.12	EE rate	0.0288	0.0265
b	0.5	avg U earnings/avg E earnings	0.61	0.58

Table 1: Summary of calibration
The first block of the table corresponds to the parameters that identify the moments informative about human capital accumulation, as described in Sections 4.1 and 4.2. The second block corresponds to the moments that inform the distribution of firm productivity and bargaining power, as described in Section 4.3. Note that the all of the identification within the first two blocks is *joint*, i.e., the parameters in the first two columns do not necessarily map to the moment in the corresponding row. The last block corresponds to the moments that identify the transition rates and replacement rates, described in Section 4.4.

5 Parameter estimates and model fit

5.1 Targeted moments

Table 1 presents a summary of the parameter values and targets. The model fits the data well on most dimensions.

Figure 7 compares the residual earnings growth moments, described in Section 4.1, in the model and the data. The bold lines in the middle show the mean of the cumulative residual earnings growth distribution across firms, and the two dashed lines show the 10th and 90th percentiles. The model fit is excellent, although it implies a little bit too much growth coming from human capital at the mean firm and at the best firms for long experience horizons.

These moments should be interpreted as being informative about disparities in returns to human capital accumulation across firms. The amount of growth and heterogeneity in growth rates is striking and is crucial for the quantitative results. One feature is that the mean shape of this profile looks similar to the overall mean earnings profile which takes into account job-to-job transitions. Thus in general, I find that there is a lot of on-the-job growth to be had early in life, which attributes less overall earnings growth to job-to-job transitions.[37]

[37]In contrast, Bagger et al. (2014) find that most of the earnings growth early on is due to "job shopping." I further

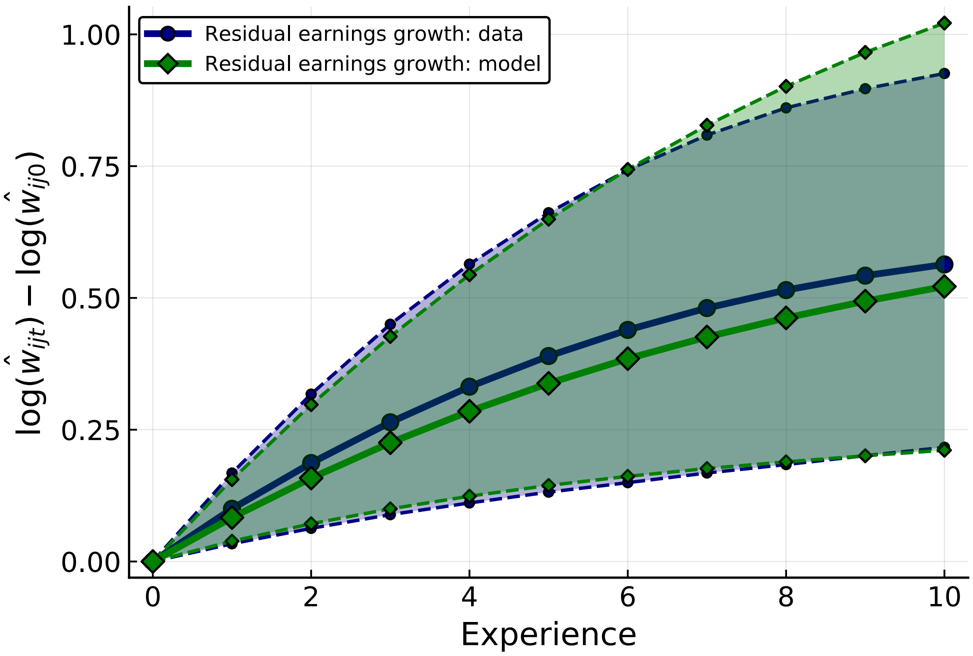

Figure 7: Residual earnings growth moments: model vs. data.
This figure depicts the distribution of earnings growth profiles across firms, when earnings growth due to search capital accumulation is removed, as in the process outlined in Section 4.1. Each marker represents one moment targeted in the calibration procedure.

These moments still pick up differences in worker composition within firms – adding the AKM moments to the estimation separates these and informs how much of the heterogeneity is truly coming from firms. Because the variances in the data are quite close to each other (the ratio of the variance of the worker to the firm effect is 1.09), I will find only slightly more heterogeneity in a, the worker component, than in q, the firm component. This will also be an important driver of the results because there will be a large part of human capital heterogeneity coming from firms.

The values of γ and α in the absorption rate function imply a very gradual decline in human capital accumulation: see Figure 3. The levels of the inputs to the human capital production function, primarily controlled by ν in the numerator of the absorption rate function, impact the measured degree of sorting – the correlation of the AKM fixed effects. However, the large negative value in the model, -0.37, is almost entirely determined by the degree of bias introduced into the model. In contrast, the model's theoretical measure of sorting, the correlation between a and q is approximately zero, for the reasons discussed in footnote 19.

The variance of the distribution of firm productivity is similar to what Jarosch (2015) estimates.[38] The estimate of the worker bargaining power implies that two-thirds of the joint value goes to the

explore this discrepancy in Section 6.1.

[38]It is going to generate less ladder climbing than in Bagger et al. (2014). This is because the model does not take into account permanent differences in the *level* of earnings across workers. The extent to which high-wage (in level) workers climb to high-wage (in level) firms will not be captured here. Jarosch (2015)'s model also does not account for this, so it is reassuring that we both find similar productivity distributions.

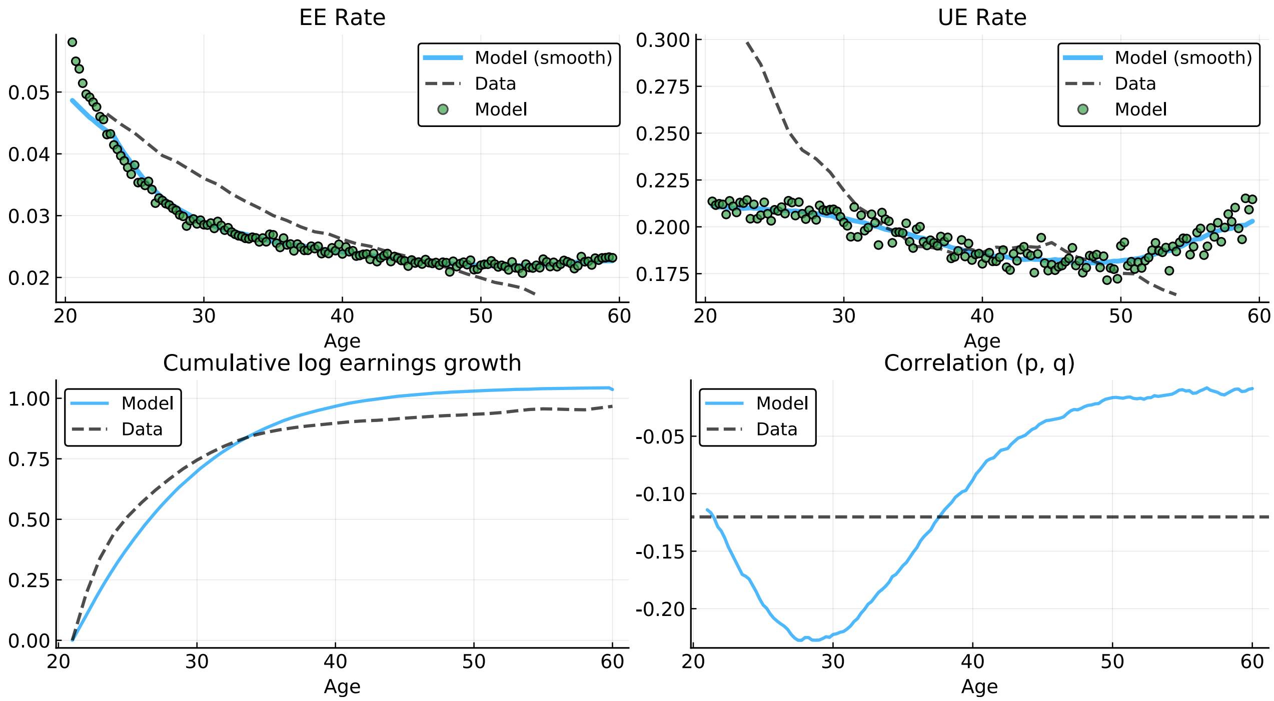

Figure 8: Untargeted moments.
The top left panel compares the job-to-job transition rate by age in the model and the data. In the data, I define a job-to-job transition as two consecutive employment spells with less than 21 days in between them. Because the model is quarterly, I also plot "Model (smooth)", which is a 3-year moving average. The top right panel is the unemployment-to-employment rate. The lower left plots cumulative log earnings growth, which at a given age is defined as the difference in mean log earnings from the log value at age 20. The lower right shows the correlation of productivity and learning environment among the accepted jobs at each age. The "data" line corresponds to the singular correlation measure derived from step 3 in Section 4.1.

worker and generates less earnings growth coming through the search capital channel compared with other studies. I attribute this result to the inclusion of human capital growth. Like in Bagger et al. (2014), the model does not need to attribute so much on-the-job growth to piece-rate increases when human capital growth is allowed.

Finally, the aggregate labor market flow rates match well. As usual, the offer arrival rate is higher in unemployment. This will imply some loss of the option value of search when workers accept employment which will mean that workers sometimes reject job offers.

5.2 Untargeted moments

For further validation, I examine the model's fit to a set of untargeted moments. These are depicted in Figure 8.

Even though I only target the aggregate EE and UE rates, the model can mostly account for their entire life-cycle profiles. In the data, both decline over the life cycle. The model matches the decline in the EE rate well. For the UE rate, I get a decline for the first 30 years and then an increase. The UE rate in the model in the first 10 years is too low: workers in the model are too selective with which jobs they accept early on. The increase at the end comes from workers becoming much less selective at older ages.

I do not target the overall earnings profile, but the model can match this well. This is because I already match the shape of the residual growth moments in Figure 7 from the parameters of the absorption rate function. Finally, I compare the life-cycle profile of the correlation in (p, q). The overall mean of this is targeted (the dashed horizontal line), but the model suggests that the negative correlation found in the data is driven by young workers. These are precisely the young workers who face the relevant trade-off between productivity and learning environment: workers who go to firms with a low learning environment early on must be compensated by a high productivity, generating the negative correlation.

5.3 What is learning environment?

In this section, I offer some insights from the data that help understand which establishment characteristics are tied to learning environment. My identification procedure enables me to assign productivities and learning environments to establishments in the data, based on the estimated earnings profiles outlined in steps 1 and 2 of Section 4.1. Given the predicted values of (α_j, β_j^1) and (γ_j, δ_j^1) from (12) and (13), I construct the predicted values of earnings growth (residual earnings growth) for the first ten years of experience (tenure) for each establishment. Then I average over these ten years to impute a single measure of learning environment (productivity) for each establishment.[39]

5.3.1 Industry and establishment size

In Appendix Tables D.1 and D.2, I illustrate how these measures correlate with the size (number of full-time employees) and industry of the establishments.[40] Note also that in the case of learning environment, these measures are based purely on residuals of earnings growth. Direct evidence to supplement and validate them would be a worthwhile endeavor.

Along the industry margin, it appears that on average the establishments in the manufacturing industries have the highest learning environment measures. This result supports the interpretation of learning environment as a quantification of the scope for *on-the-job* learning. These types of establishments appear to offer more opportunities for learning by doing. This finding could also stem from the widespread presence of apprenticeships in these industries. The more white-collar

[39]Note that I do not use this measure of productivity to inform the distribution of p in the model. Instead I stick to the more standard approach of using job-to-job transitions as discussed in Section 4.3.

[40]The set of establishments that I can impute these for is limited because of the restrictions on the number of workers needed to estimate the earnings profiles. The main constraint is that the establishment needs to have a sufficient amount of older workers hired out of unemployment in order to impute a productivity, and therefore learning environment.

industries have lower learning environment measures. This could mean that pre-existing human capital like the type acquired from schooling may be relatively more important in these industries.

In spite of these observations, the variation in averages between size classes and industries appears to be small. The coefficient of variation is quite stable across categories. A regression of learning environment on a complete set of industry $\times$ size class interactions yields an R^2 of 36%. These observations suggest that heterogeneity in learning environments is a major factor within industries: it reflects a source of uncertainty even for workers who remain in one industry throughout their entire career.

5.3.2 Evidence from the IAB Establishment Panel

I can go further by exploiting the link between the LIAB sample and a survey that the IAB fields each year to managers of establishments, the IAB Establishment Panel. This survey often contains blocks of questions on topics that may indicative of on-the-job human capital accumulation at the establishment. Here, I investigate whether variables related to on-the-job training and apprenticeship programs[41] are correlated with my learning environment measure.

The topics covered in the IAB Establishment Panel vary from year-to-year, but generally include business policy and development, investments, workforce structure, vocational training and apprenticeship programs, personnel recruitment, and working hours, as well as special topics that appear in certain years (for more information on this survey, see Fischer et al. (2009)). The establishments that appear in the IAB Establishment Panel make up the core sample of the LIAB. These are exactly establishments for which I can observe the complete workforce, which means that I can link almost every establishment with a valid learning environment measure back to its responses in the survey.

I focus on the survey blocks related to further training and apprenticeship programs. Specifically, for further training, I look at the types of on-the-job training offered and the topics covered by this further training. For apprenticeship, the most relevant questions ask how many workers participate in, successfully complete, and are hired through the establishment's apprenticeship program. For details on how these responses are aggregated and linked to the LIAB data, as well as more information about the variables and questions, see Appendix E.

In Table 2, I show how each of the further training variables relates to learning environment. Many of the correlation coefficients in the middle column are modest, but are positive and statistically significant. I also regress learning environment on each of the training measures separately, while also including fixed effects for the establishment's industry, size class, and the share of workers within 5-year age bins. The coefficients on each variable are in the right column. For the binary variables (all variables except for the fraction of employees receiving training and the number

[41] In Germany, apprenticeships are a prominent feature of the labor market – nearly three-quarters of workers in my sample have completed a vocational training program. They are required in order for a worker to enter many occupations. Apprenticeships are tied to specific establishments but are also regulated by the government.

Variable	Correlation	Regression Coeff. $\times$ 10
Offers any training	0.045	0.049** (0.021)
Fraction of employees receiving training	0.022	0.011 (0.016)
Types of training offered		
Number of types of training per year	0.091***	0.01*** (0.004)
Number of types offered all years	0.096***	0.006** (0.003)
External courses, seminars, or workshops	0.105***	0.066*** (0.021)
Internal courses, seminars, or workshops	0.085**	0.07*** (0.016)
Further training on-the-job (instruction, initial skill adaptation training)	0.061*	0.02 (0.016)
Participation in lectures, symposia, fairs, etc.	0.086***	0.035** (0.014)
Job rotation	0.078**	0.006 (0.016)
Self-directed study	0.017	0.016 (0.014)
Quality circles, workshop circles, continuous improvement teams	0.014	-0.012 (0.015)
Other	0.032	0.018 (0.016)
1st or 2nd most important training topic		
Business topics	0.099**	0.01 (0.012)
Commercial, scientific, technical, design topics	0.204***	0.007 (0.012)
EDP, information/communication technology	-0.019	0.004 (0.012)
Soft skills (e.g. ability to work in team, conflict management, work organization)	-0.212***	-0.029** (0.011)
Other	-0.06	0.007 (0.012)

Table 2: Learning environment and on-the-job training measures.
The middle column displays the pairwise correlation coefficient between the establishment's learning environment and the on-the-job training variable of interest. The right column is the OLS coefficient on the variable of interest in a regression with learning environment as the dependent variable with fixed effects for size class, industry, and share of employees in 5-year age brackets. Standard errors in parentheses. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Variable	Correlation	Regression Coeff. × 10
Fulfills educational requirements	0.303***	0.093***
		(0.016)
All apprentices retained	0.201***	0.034**
		(0.014)
Fraction of apprentices retained	0.387***	0.093***
		(0.017)
Has successfully completed apprenticeships	0.428***	0.157***
		(0.015)
Fraction of successfully completed apprenticeships	-0.201***	-0.084***
		(0.024)
Apprentice share	-0.095***	-0.125**
		(0.057)
Apprentice share under age 30	-0.052	0.002
		(0.01)

Table 3: Learning environment and apprenticeship variables.
The middle column displays the pairwise correlation coefficient between the establishment's learning environment and the apprenticeship variable of interest. The right column is the OLS coefficient on the variable of interest in a regression with learning environment as the dependent variable with fixed effects for size class, industry, and share of employees in 5-year age brackets. Standard errors in parentheses. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

of types per year and in all years), these numbers can be interpreted as the expected increase in learning environment if each training type (or topic) is offered in every year of the sample.

The types of training that have the strongest and most significant relationship with learning environment are external and internal courses, and to a lesser extent, participation in lectures. For example, offering internal courses would be associated with a 19% increase in the learning environment measure for the average establishment. There is also an association between the variety of types of training offered by the establishment and the learning environment. This appears to be more robust than the link with whether the establishment offers training at all. The topics of training do not exhibit any significant correlation once additional factors are controlled for, suggesting that these may be more related to industry rather than human capital accumulation at the establishment (with the exception of the negative correlation for soft skills). The learning environment is also not related to the share of workers who participate in the training programs.

Table 3 displays the same set of statistics for the apprenticeship variables. Overall, the characteristics of apprenticeship programs have stronger and more robust correlations with learning environment. Learning environment has a negative relationship with the fraction of completed apprenticeships (among all apprentices) and a positive relationship with the fraction of apprentices retained (among completed apprenticeships). Together, this may indicate that a longer apprenticeship program combined with the establishment's willingness to invest in the skills of the workers (as indicated by a high retention rate) become reflected in a higher learning environment. At the same time, the negative (statistically insignificant) coefficient on the share of workers who are apprentices (under age 30) suggest that the learning environment is not mechanically higher based on how many apprentices are at the establishment, and thus, how many workers are directly involved in the apprenticeship program.[42]

Taking all of these results together, I conclude that the learning environments I derived are correlated with features of the establishments that are designed to increase the human capital of employees. This is reassuring because these learning environments were constructed just off of earnings, experience, and tenure data, and were tied to human capital accumulation based only off economic theory. The findings here show that it is not just the presence of training programs that are important for human capital accumulation, but also the variety of types. This suggests that having different forms of training available increases human capital accumulation since a given worker can find a program that best fits his or her learning style. Another takeaway is that these training and apprenticeship programs do not just impact those who participate in them. This result points to a role for intangible or informal features of the establishment that are correlated with the presence and features of these programs. These additional unobservable characteristics may be applicable to all workers and affect human capital accumulation, and therefore get picked up in the learning environment measure.

5.4 Model validation

I also use my learning environment measures to provide direct empirical evidence for one of the key implications of the model. Recall that the indifference curves in Figure 6 imply that valuation of learning environment is highest when workers are young. To check this in the data, I analyze how the age composition of new hires relates to the estimated learning environment.

In the left panel of Figure 9, I divide establishments into 10 learning environment deciles and compute the median and interquartile range of the share of the establishment's hires that are new labor market entrants.[43] New labor market entrants are defined in the same way as discussed in Section 4.1. I find that establishments with better learning environments do tend to hire more new labor market entrants. Moving to the bottom decile to the top decile nearly triples the share of workers for which this establishment represents their first "career job". The correlation coefficient between learning environment and the share of new entrants is 0.31, statistically significant at the 1% level.

The model also implies that workers who are old enough should be indifferent to learning environment. This prediction is also borne out in the data. The right panel of Figure 9 shows that there is no discernible relationship between learning environment and the share of the establishment's hires who are older than 45.[44] The correlation coefficient is 0.022 and is not statistically different from zero.

[42] In this sample, apprentices make up on average 5.4% of each establishment's employment and 19.1% of employment among workers below age 30.

[43] My learning environment measure does not vary from year to year, but the age composition of new hires does. I aggregate up the latter by taking the average over years, in order to end up with one observation per establishment.

[44] Similar results hold within other age groups for workers older than 40.

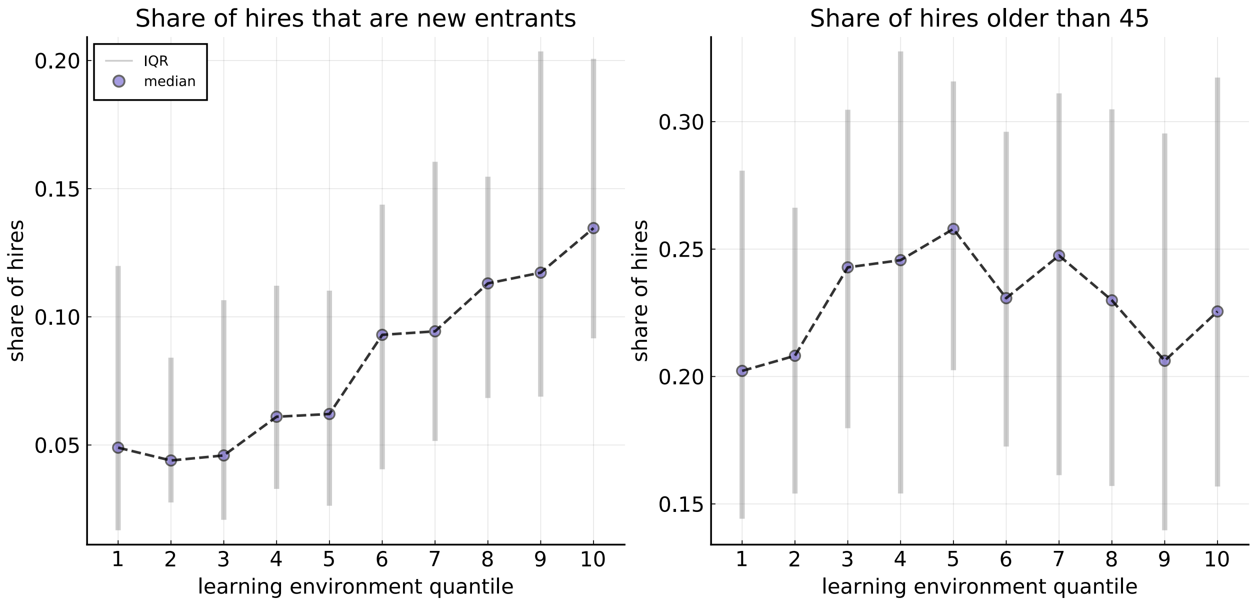

Figure 9: Learning environment vs. composition of hires.
The left panel displays the 25th, 50th, and 75th percentiles of the share of hires that are new entrants by learning environment decile. New labor market entrants are defined as hires with zero experience, computed as described in Section 4.1. The right panel is the same, for share of hires greater than age 45.

6 Quantitative Results: Life-Cycle Earnings Profiles

In this section, I use the model to understand the patterns in the life-cycle earnings profile. To quantify the importance of the firm learning environment channel, I study the model with and without heterogeneity in worker learning ability. By removing *ex-ante* differences in workers, I create a setting in which the only source of heterogeneity in the labor market outcomes of workers comes from the series of firms they happen to match with.[45] In other words, search frictions not only affect how rents are split, but also translate to persistent worker variation. This is the novel interaction put forth by this paper.

6.1 Life-cycle mean profile

Where does the growth in life-cycle earnings come from? In this section, I use the model to explore the sources of life-cycle earnings growth. Since log earnings in the model are the sum of human capital, the productivity of the firm, and the piece-rate, I can decompose the earnings profile into these three components.

[45] I have also done the opposite exercise, in which I turn off differences in firm learning environments but keep the *ex-ante* heterogeneity across workers in Appendix F. However, I argue that this counterfactual is less relevant because it introduces different job search behavior on the part of workers. Workers' job search strategies change because these depend on the distribution of q, but not a. Shutting down a, as done in this section, does not have this effect, and thus truly isolates the effect of one source of heterogeneity.

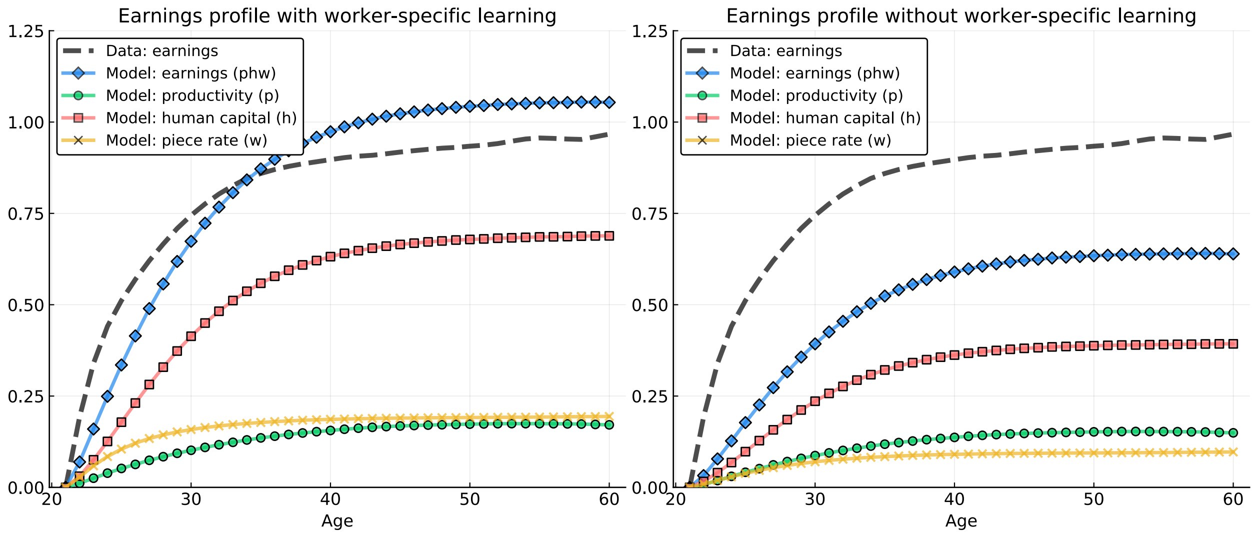

Figure 10: Life-cycle mean of log earnings and decomposition.
The left panel plots the mean of log earnings in the data and in the model where workers are heterogeneous in a. The right panel plots the corresponding means in the version of the model in which there is no worker-specific component of human capital accumulation: $a = 0$ for all workers so that all growth in human capital solely comes from firm learning environments. Each series is derived from the profile of mean log earnings by age. Each is normalized to zero at the start by subtracting the value at age 20.

The left panel of Figure 10 shows the earnings profile in the data, as well as the model counterpart and its three components. Each series is normalized to zero at age 20, so that the interpretation of the y-axis is the difference in average log earnings since age 20. Most of the increase comes from human capital: it drives about 2/3 of growth, whereas the productivity and piece-rate equally drive the remaining 1/3. My decomposition results are quantitatively similar to those of Engbom (2020) who estimates a model which also allows for human capital growth variation across firms.

Bagger et al. (2014) perform almost the same decomposition in a model with heterogeneity in firm productivity, idiosyncratic shocks to match output, and deterministic human capital growth that only depends on age. In contrast to my results, they find a larger role for growth in firm productivity early in life, as workers make a lot of transitions to climb the ladder into a good match, or "job shopping." The differences between our results mainly come from the inclusion of a firm-specific component of human capital growth. I attribute more of the earnings variation between firms to the human capital of its workers, which was partially picked up through the firm's own learning environment. This means that there is less earnings dispersion leftover to come from other sources, captured by the firm's productivity. As a result, the workers in my model have less of a ladder to climb in productivity. Other differences may come from the data used. Bagger et al. (2014) use Danish micro data. I find in that in Germany there is a lot of on-the-job earnings growth, dampening the contributions of job-to-job transitions to earnings growth (see Figure 7). Their findings for Denmark indicate that this may not be the case there. However, I do find that the role of the productivity and piece-rate, the standard job search channels, is highest early on in life, consistent with their results. To see this, note that the share of earnings growth coming from these two sources is highest in the first few years, and then diminishes from then on as human capital keeps growing.

What is the contribution of firms? How much of a worker's stock of human capital comes from the component they accumulate that is firm-specific? To assess this, I simulate a version of the model in which there is no worker-specific component to human capital growth: all workers have learning ability a equal to zero.[46] The earnings profiles generated by this version of the model are depicted in the right panel of Figure 10.

In this economy, there is less human capital acquired, translating to less earnings growth over the life cycle. Here, the growth in average log human capital is 0.397, compared with 0.689 in the full model. This suggests that 57.6% of the human capital stock is acquired through firms.

Thus, I find that a large proportion of the human capital stock is driven by firm learning environments, despite the fact that I estimate a higher average worker learning ability than average firm learning environment.[47] The reason is the endogenous job choices of workers. Workers have the opportunity to visit several firms over their lifetime. Their decisions steer them towards high growth firms, which means they have the opportunity to accumulate more human capital than they would if their ability to learn was completely pre-determined when they enter the labor market.

6.2 Life-cycle variance profile

In this section, I use the model to explore the sources of the patterns of life-cycle inequality. Just as I found for the life-cycle mean earnings profile, I find here that firms and their contribution to human capital accumulation are a core contributor to the increase in life-cycle earnings variance. This result offers a new explanation for rising earnings inequality over the life cycle.

Where does the growth in life-cycle variance come from? The black dashed lines in Figure 11 represent the variance in log earnings at each age from the data.[48] The blue line with the diamonds is the variance profile in the model. It matches by construction because I targeted the increase in life-cycle variance. However, despite the fact that I did not target the general shape, the model accounts for a flattening off after age 40, but not the increase after age 55 or so.

The variance of log earnings in the model can be decomposed into:

[46] The model does not need to be re-solved because the distribution of worker learning ability has no impact on the policy functions.

[47] Doing the opposite exercise, setting all firm learning environments to zero, gives exactly the opposite, 42.4% of acquired human capital coming from workers. This is because of the low degree of sorting in the model.

[48] This series is shifted down to match the lowest point achieved by the corresponding profile in the model. In the data, some fraction of the variance is captured by worker fixed effects in the level of earnings. These are not present in the model. In either case, whether the profile is shifted or not, the increase in variance is the same.

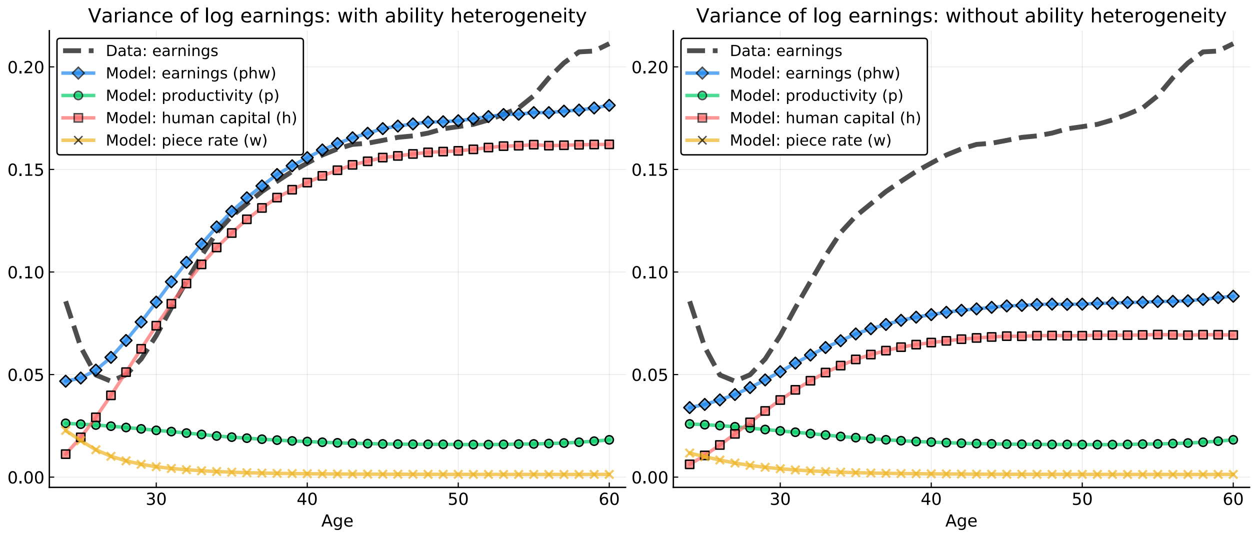

Figure 11: Life-cycle variance of log earnings and decomposition.
The left panel plots the variance of log earnings in the data and in the model where workers are heterogeneous in a. The right panel plots the corresponding variances in the version of the model without heterogeneity in a – all workers have the median ability from the original distribution $G(a)$.

$$\begin{aligned}\text{var}\left(\log \text{earnings}\right) = \text{var}\left(\log p\right) + \text{var}\left(\log h\right) + \text{var}\left(\log w\right) \\ + 2\text{cov}\left(\log p, \log h\right) + 2\text{cov}\left(\log p, \log w\right) + 2\text{cov}\left(\log w, \log h\right)\end{aligned} \tag{15}$$

Each of the variance terms in (15) from the full model are plotted in the left panel of Figure 11 as the green, pink, and yellow lines, respectively.[49] The increase in the variance of human capital clearly drives the overall increase in the variance. The dispersion in human capital increases because workers accumulate human capital at different rates, both because of their different learning abilities and the learning environments of the firms they match with. The flattening out of the variance of human capital roughly coincides with the time at which human capital accumulation is no longer operative, at age 50.

Without human capital accumulation, this model would miss the increase in life-cycle earnings variance. In this scenario, only the firm productivity and piece-rate channels would be operative – the green and yellow lines, respectively. The variance of firm productivity component measures the dispersion in firm wage premia in *levels*. It declines slightly as workers move to higher paying firms over their lives. With this firm productivity distribution, they settle into a smaller set of better-paying firms compared to where they started out. The variance in the piece-rate also declines. As workers build up outside offers and improve their bargaining positions, the distribu-

[49]Most of the covariance terms are small. The only quantitatively large covariance term is the one between human capital and the piece-rate for the first 10 years after labor market entry. This arises because workers with low human capital have even greater incentive to match to firms with better learning environments, and therefore accept very low piece-rates in order to work there.

tion of piece-rates shifts towards 1, its upper bound. Together, these would imply a decrease in the variance of earnings in a model with only these two forces present. Here, however, the increase in human capital dispersion takes over these channels and drives the increase in overall earnings variance.

What is the contribution of firms? The following exercise quantifies the importance of the firm learning environment channel. I shut down the heterogeneity in worker learning ability a, meaning I simulate a version of the model in which everyone has the median learning ability[50] from the original distribution $G(a)$.[51] In this version, all human capital variation arises only from the kinds of firms that workers match with – a new "luck" channel that impacts workers' earnings outcomes. I then recompute the earnings variance profile and decomposition.

The corresponding variances for each component in the version of the model without heterogeneity in learning ability are shown in the right panel of Figure 11. In this counterfactual, the variance of log earnings increases from 0.032 to 0.088. This increase is 41.4% of the increase of the variance of log earnings in the full model, implying that this channel is responsible for about 41% of the increase in life-cycle inequality.

Another interpretation of this result says that the importance of firms is highest early on in worker's lives. This is because early on, workers have limited employment histories and also because they are accumulating human capital very quickly. As a result, a worker's initial match is important. By age 30, 85% of new earnings dispersion comes from human capital. Of the additional variance accumulated since entry, 51% arises due to firm differences. Despite their own abilities, workers who get lucky early on and match to a firm with a better learning environment get a head start over their peers, contributing to inequality among their cohort. But as workers have time to catch up, the influence of firms declines because workers have had time to find better matches. This mechanism also means that there is a component of inequality in lifetime earnings that can be traced back to early labor market experiences; in particular, the identity of a worker's initial match. For further evidence of a similar phenomenon, see Arellano-Bover (2019). He links the size of the firm in which a worker gets their first job to lifetime income, and finds evidence of human capital being a driver of this relationship.

This finding offers a new explanation for rising earnings inequality over the life cycle. Two major insights emerge. First, it is not just a matter of inherent differences across workers. Firms too have an effect on a given worker's earning growth rates and thus contribute to the increased heterogeneity between workers that becomes more pronounced over the life cycle. Second, luck

[50] I set everyone to the median because I only want to shut down the *heterogeneity* in worker human capital growth, but not worker human capital growth itself. As a result, the life-cycle mean profile, but not variance profile, still looks similar to the data.

[51] When I do the opposite exercise in which I turn off differences in firm learning environments but keep the *ex-ante* heterogeneity across workers, I find that worker differences account for about 59% of the increase in variance. This is exactly the remaining share of the 41%. In general, these shares do not have to add to 1. If there were any meaningful sorting in the model, workers could change their job search decisions enough such that the allocations in the two counterfactuals look significantly different.

manifests itself in a novel way. Search frictions impact the amount of human capital workers are able to accumulate. This effect goes beyond the standard role for search in which it only affects how rents are split. As a result, there is an interaction between luck and worker differences because persistent heterogeneity across workers comes about due to variation in labor market histories.

7 Reduced-Form Earnings Process Estimation

The results thus far imply that employers play an important role in the development of the human capital of their workers. Next, I show how this finding matters for the statistical properties of the labor income process. I find that the stochastic properties of workers' earnings in the model are in line with widely-used specifications of the income process. In doing so, I demonstrate how the mechanisms introduced here can provide economic interpretations of some features of the labor income process. The principal result shows that firms are partially responsible for the variation in earnings profiles estimated by these statistical models.

7.1 The Earnings Process

The literature that studies the earnings risk faced by individuals typically models the earnings process as the sum of a persistent and transitory component, and sometimes a life-cycle trend. This flexible specification has been widely used for several decades (for instance, MaCurdy (1982), Abowd and Card (1989), and Meghir and Pistaferri (2004)) has been shown to provide a good fit to the income dynamics observed in the data. When feeding them into consumption/savings models, a simple variation is used in which the random component is an AR(1) plus a transitory shock, as in Heathcote, Storesletten and Violante (2010).

Another modification allows for heterogeneity in the life-cycle trend across individuals. These differences are typically attributed to variation in ability. Guvenen (2007) and Guvenen (2009) estimate this parsimonious specification with and without profile heterogeneity, two cases which he calls RIP (Restricted Income Profiles) and HIP (Heterogeneous Income Profiles). From here on, I adopt his specification and estimate RIP and HIP processes on the model-generated earnings data. I do this in the versions of the model with and without worker learning ability heterogeneity, to quantify how much of the profile heterogeneity is driven by firms.

The log residual earnings of individual i at age h, y_h^i are given by:

$$y_h^i = \alpha^i + \beta^i h + z_h^i + \varepsilon_h^i \tag{16}$$

$$z_h^i = \rho z_{h-1}^i + \eta_h^i \tag{17}$$

where α_i is an individual-specific level of labor income and β_i is an individual-specific growth rate of income.[52] The vector (α_i, β_i) is independently and identically distributed across workers with zero mean, variances σ_α^2 and σ_β^2, and covariance $\sigma_{\alpha\beta}$. Aside from these permanent components of worker heterogeneity, the income process also contains an AR(1) component, z_h^i with persistence parameter ρ, and a purely transitory component, ε_h^i. The shocks to the AR(1) and transitory components are assumed to be independent, with zero mean and variances σ_η^2 and σ_ε^2. Under RIP, the heterogeneity in individual growth rates is shut down: $\sigma_\beta^2 = 0$ and $\sigma_{\alpha\beta} = 0$. Thus the parameters to be estimated are $[\sigma_\alpha, \sigma_\varepsilon, \sigma_\eta, \rho]$ in RIP and $[\sigma_\alpha, \sigma_\varepsilon, \sigma_\eta, \rho, \sigma_\beta, \sigma_{\alpha\beta}]$ in HIP.

With panel data on individuals, the parameters can be identified by using the cross-covariances of labor earnings at different ages. The variances and covariances implied by the income process in (16) and (17) are:

$$\mathrm{var}(y_h^i) = \sigma_\alpha^2 + \sigma_\varepsilon^2 + \left(\frac{1-\rho^{2h+1}}{1-\rho^2}\right)\sigma_\eta^2 + 2\sigma_{\alpha\beta}h + \sigma_\beta^2 h^2 \tag{18}$$

$$\mathrm{cov}(y_h^i, y_{h+n}^i) = \sigma_\alpha^2 + \sigma_{\alpha\beta}(2h+n) + \sigma_\beta^2 h(h+n) + \rho^n \left(\frac{1-\rho^{2h+1}}{1-\rho^2}\right)\sigma_\eta^2 \tag{19}$$

To estimate these income processes in the model, I first need to construct a panel of worker earnings which I will use to compute the analogues of (18) and (19). Importantly, this panel will look more like the PSID, rather than a matched employer-employee data set. I throw out information on firms, and only keep earnings data for each worker by age.

In the model, I impose restrictions that are similar to the ones used on real-life panel data. I aggregate to yearly observations by calculating the total earnings in employment in each year, as long as the worker was employed for at least one quarter. By construction, the model contains 40 years of data for each worker. All cross-covariances are computed on income residuals, obtained by regressing earnings on an age profile. As is standard, I use a GMM procedure to obtain parameter estimates. I search for the parameter set that minimizes the distance between the theoretical moments (18) and (19) and the cross-sectional covariances created from the panel. This amounts to 351 moments and either 4 or 6 parameters. For more details about the implementation, see Appendix G.

7.2 Estimates

U.S. vs. Germany Columns (1) and (2) of Table 4 report Guvenen (2009)'s baseline estimates for the RIP and HIP processes for the U.S. Columns (3) and (4) report the estimates I find for Germany. Comparing the corresponding RIP and HIP estimates across countries, I find that in Germany, the permanent shocks appear to be larger but less persistent, as indicated by the estimates of ρ and σ_η^2. The transitory shocks are smaller. The estimates of σ_β^2 also reveal less slope heterogeneity in Germany. The lower degree of persistence and profile variation in Germany are both consistent

[52] The use of income residuals removes the estimated effects of observable characteristics and common aggregate time trends.

	U.S.: Guvenen (2009)		Germany		Model with a het.		Model w/o a het.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	RIP	HIP	RIP	HIP	RIP	HIP	RIP	HIP
ρ	0.988	0.821	0.9258	0.7829	0.9944	0.9561	0.9688	0.9268
σ_α^2	0.058	0.022	0.0452	0.0601	0	0	0	0
σ_ε^2	0.061	0.047	0.0445	0.0137	0.0758	0.0713	0.0747	0.0695
σ_η^2	0.015	0.029	0.0224	0.0451	0.0093	0.0110	0.0078	0.0101
$\text{corr}_{\alpha\beta}$	-	-23.0%	-	1.29%	-	0	-	0
σ_β^2	-	0.00038	-	0.00013	-	0.00016	-	0.00007

Table 4: Estimating the parameters of the earnings process
Estimates refer to the parameters of the process in equations (16) and (17). In RIP, σ_β^2 and $\sigma_{\alpha\beta}$ are restricted to be zero, thus they are not estimated. The U.S. estimates are taken from Guvenen (2009)'s Table 1 for all workers with the baseline sample. The German estimates are the author's own calculations.

with the fact that the increase in life-cycle earnings variance is lower in Germany compared to the U.S. The results also indicate that the fraction of total cross-sectional inequality attributable to profile heterogeneity is lower in Germany: 30.3% by age 45 compared with about 58% in the U.S.[53]

Model vs. data Columns (5) and (6) show the corresponding estimates from my model with heterogeneity in worker ability.[54] Note that none of the features of the income process were targeted in the parameterization. The model generates a higher degree of persistence and larger transitory shocks compared to the data. However, the profile heterogeneity estimate σ_β^2 is quite close to what I found in the data. The earnings process in the model attributes 37% of the variance of earnings at age 45 to HIP. The differences between RIP and HIP are also consistent with the data. In both the model and data, going from RIP to HIP lowers the variance of the transitory shock and increases the variance of the permanent shock. It also decreases the persistence parameter and instead attributes more differences in individuals to heterogeneity in income profiles. In the version of the model in column (6), the profile heterogeneity is coming from both worker-specific learning ability and firm learning environments.[55]

The characteristics of the earnings process generated by my model show that the model microfounds two features: persistent shocks to earnings and heterogeneity in earnings growth rates. Shocks in the model are persistent because they reflect job-to-job transitions or separations to unemployment, both of which are relatively long-lived. This result has a similar to flavor studies that endogenize earnings risk through job mobility, such as Low, Meghir and Pistaferri (2010), Lise, Meghir and Robin (2016). Individuals also face different income growth rates, because of a combination of their own learning ability and the learning environments and productivities of

[53]To see this, take the terms in (18) that depend on h and compare them to the total variance of income at age h.

[54]Note that in the model, the variances of all terms involving σ_α^2 are zero because there are no permanent differences in the level of earnings across workers.

[55]Examining the autocovariance structure in the model and data provides clues about the source of the differences. See Figure G.1 and the discussion in Appendix G.

their employers. In contrast to my model, Huggett, Ventura and Yaron (2011) are also able to generate estimates similar to the data by explicitly including shocks to human capital and only allowing for heterogeneity in worker learning ability.

What is the contribution of firms? Next, as I did for the counterfactual exercises in Section 6, I turn off all heterogeneity in worker learning ability and give each worker the median value from the original distribution. I re-estimate the labor income processes and present the results in columns (7) and (8) of Table 4. This version still exhibits the bias in ρ and also estimates some dispersion in worker-specific growth rates: about half of what was found in the model with heterogeneity in learning ability. 25% of the earnings variance at age 45 is attributable to profile heterogeneity.[56]

In this version of the model, all income profile heterogeneity is due to the series of firms a worker matches with. An individual's income profile is pieced together by different growth rates offered by various firms. In previous literature, the findings about heterogeneous income profiles across workers were mainly interpreted as fixed worker differences, for instance, coming from learning ability. Models in which individuals vary in their ability to accumulate human capital often serve as a theoretical motivation for the HIP specification. However, I find that even in a version of my model with no *ex-ante* variation across workers, the earnings process still picks up this kind of heterogeneity in growth rates. This suggests that some of these estimated disparities are coming from firms, and are not permanent differences at all. Moreover, the presence of a stochastic firm-specific component of growth adds an additional source of income risk that may have implications for consumption dynamics.

8 Policy Experiments

So far, the findings suggest that many labor market outcomes are not due to permanent variation across workers, but rather come about because of search and matching frictions. This means there is a way for policy to affect the allocation of workers to firms. In this section, I use the model to conduct policy experiments in which the structure of unemployment insurance (UI) impacts the types of jobs that workers are willing to accept.

The trade-offs that workers face between jobs at different points in the life cycle is key to understanding why the types of jobs held by workers affect aggregate outcomes in the model. On one hand, young workers should be very selective about which jobs they accept. The firm's learning environment is important to them because finding good firms along this dimension early in life will boost earnings for their entire lifetime. Moreover, if workers have access to firms with good learning environments, aggregate output is boosted because matches produce more when workers have been able to accumulate more human capital early on in life. Generous unemployment

[56] This is not half of the 37% from the full model because the persistence parameter is lower, which means more of a contribution from HIP.

benefits, especially for young workers, incentivizes them to wait longer for these types of jobs when they enter the labor market. On the other hand, if workers are waiting too long to accept jobs their human capital stagnates, unemployment is high, and output is lower. The right design of unemployment insurance policies can help balance these trade-offs.

8.1 Efficiency

Before discussing the implementation of the policies, I give a brief overview of the efficiency properties of the model. The efficient benchmark described here will be used to evaluate the economic outcomes achieved in the policy experiments. The efficiency properties of this model are in line with Jarosch (2015)'s model.

The equilibrium allocation of workers to firms is inefficient because of an inconsistency in how workers value jobs and how a utilitarian planner would value jobs. In the decentralized equilibrium, workers enter matches that would never be implemented by a social planner. In these cases, the planner would prefer to leave the worker in unemployment or in a previous match. This is because the planner takes into account the potential of that worker to soon form other better matches. These workers therefore exhibit a positive search externality. Note that the inefficiency here is of a partial equilibrium variety. In a model with endogenous vacancy creation, the planner would also care about the congestion externality that an additional unemployed worker would create.

The planner's margin for adjusting the allocation is by choosing the set of acceptable job offers for unemployed workers of each type (a, h, t) and the set of acceptable outside offers for employed workers as a function of (a, h, θ, t). Since all job acceptance decisions are made by comparing joint match values, this comes down to the planner choosing its own joint match value function, $M_t^P(a, h, \theta)$.

The worker's joint match value function will coincide with the planner's when $\sigma = 1$, or all of the bargaining power goes to the worker. In this case, the worker is using the same criteria as the planner when making job acceptance decisions. To see this, notice that in the wage-setting equations, (5), (6), and (7), the worker's value function becomes the same as the match value function when $\sigma = 1$. This means that both the worker and the planner are fully internalizing the entire value of the matches formed.

In this case, as long as $\lambda_E < \lambda_U$, the equilibrium is efficient and the welfare of new labor market entrants is maximized.[57] The offer arrival rate needs to be lower in employment so that there is some option value of search given up when the worker accepts a job – otherwise there is no benefit to the planner of leaving the worker in unemployment.

What are the main differences between the decentralized equilibrium and the planner's allocation?

[57] Because all workers enter unemployed and with the same initial level of human capital h_0, welfare corresponds to $\int U_0(a, h_0) dG(a)$.

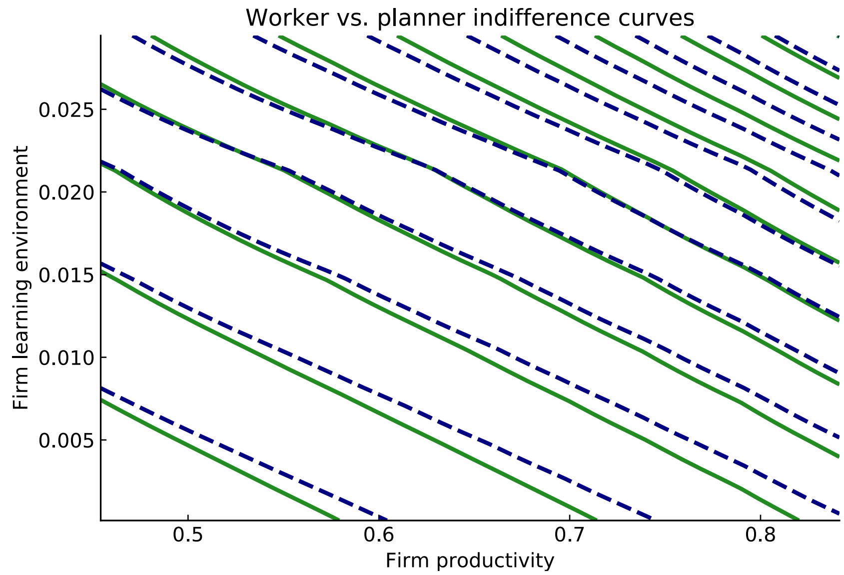

Figure 12: Indifference curves in decentralized (solid line) vs. efficient (dashed line) equilibrium
Traces out indifference curves in (p, q) space. Each curve is a contour of the joint match value function of the workers in the decentralized equilibrium, $M_t(a, h, \theta)$, and of the planner, $M_t^P(a, h, \theta)$. Worker learning ability and human capital are fixed at the same arbitrary values in both economies.

It turns out that workers in the decentralized equilibrium undervalue the learning environment of the firm when making decisions. To see this, compare the worker's and the planner's indifference curves in Figure 12. The planner's indifference curves are flatter, suggesting that the planner wants workers to be more selective on the firm's learning environment. This means that in equilibrium, workers do not fully internalize the long-term benefits of matching to a firm with a good learning environment, creating an inefficiency.

8.2 Policy Environment

Unemployment insurance in the model alters reservation strategies and thus changes the set of jobs that workers accept. The structure of the UI policies, therefore, can help bring the economy closer to the planner's allocation, improve welfare, and affect other outcomes like inequality and output.

In the baseline version of the model explored thus far, unemployment benefits replace some fraction b of a worker's human capital.[58] In this section, I will consider two types of policies that change the setup of UI. The first type is a flat benefits schedule in which I simply vary the value of b. In the second type, the replacement rate depends on age.

[58] This is different than the replacement rate on *earnings*, to which the model was calibrated. Using a replacement rate on human capital means I avoid having to keep track of past earnings throughout each unemployment spell. In the end, this specification still generates a negative relationship between earnings in the past job and subsequent benefits because earnings are also determined by p and w. This relationship is consistent with the data.

In both cases, the replacement rate $b(t)$ will take the following form:

$$b(t) = (b + z(t))h$$

where $z(t) = z_1 t^{-\frac{1}{z_2}}$, and b is the baseline flat unemployment benefit from the calibrated model, 0.5. This form of $b(t)$ says that each unemployed worker receives $z(t)h$ additional insurance above what they receive in the baseline. Therefore an age-dependent schedule will be characterized by a pair (z_1, z_2).

The additional unemployment benefits are funded by a lump sum tax on earnings, B, paid by employed workers.[59] For every policy I consider, I search for the tax on the employed such that the net present value of the additional transfers to the unemployed equals the net present value of the taxes on the employed.[60]

I will study the impact of the policies on four model objects, which are computed in the following ways:

1. **Output**: The amount produced by each match, ph, aggregated across all workers (zero for unemployed workers).
2. **Welfare of new entrants**: The value functions for new labor market entrants. Since all workers enter unemployed and with the same level of human capital, it is the value of unemployment at the initial human capital level integrated over the distribution of worker ability types, $\int U_0(a, h_0) dG(a)$. This is what is maximized by solving the planner's problem, assuming that the planner and the worker discount the future at the same rate.
3. **Variance of lifetime earnings**: Lifetime income is the discounted sum of pre-tax labor income earnings throughout a worker's life. The discount rate is β and inequality is measured as the variance of the log of this object across workers. This is a long-run measure of worker outcomes that takes into account all the events that happen over a worker's career.
4. **Variance of log earnings**: Pre-tax variance of log earnings in the cross-section.

I will compare these outcomes to the efficient benchmark in which $\sigma = 1$ and $b(t) = 0$.[61]

[59] A proportional tax on labor income would be ideal, but it makes the model intractable as the joint match value would depend on the piece-rate.

[60] In the baseline economy, the "additional" transfers to the unemployed and the tax on the employed are zero. This economy in a sense corresponded to a situation where the government held a deficit because it was funding UI but no one was paying for it. Computing the tax in this way keeps things revenue-neutral. Consequently, lowering the replacement rate actually corresponds to giving a transfer to the employed workers.

[61] In the baseline model, like in Jarosch (2015), unemployment benefits are interpreted as and mapped to the net replacement rate in the data, and the flow value of unemployment is zero. Therefore, the planner's benchmark corresponds to an economy where there are no benefits and no flow value of unemployment: $b(t) = 0$.

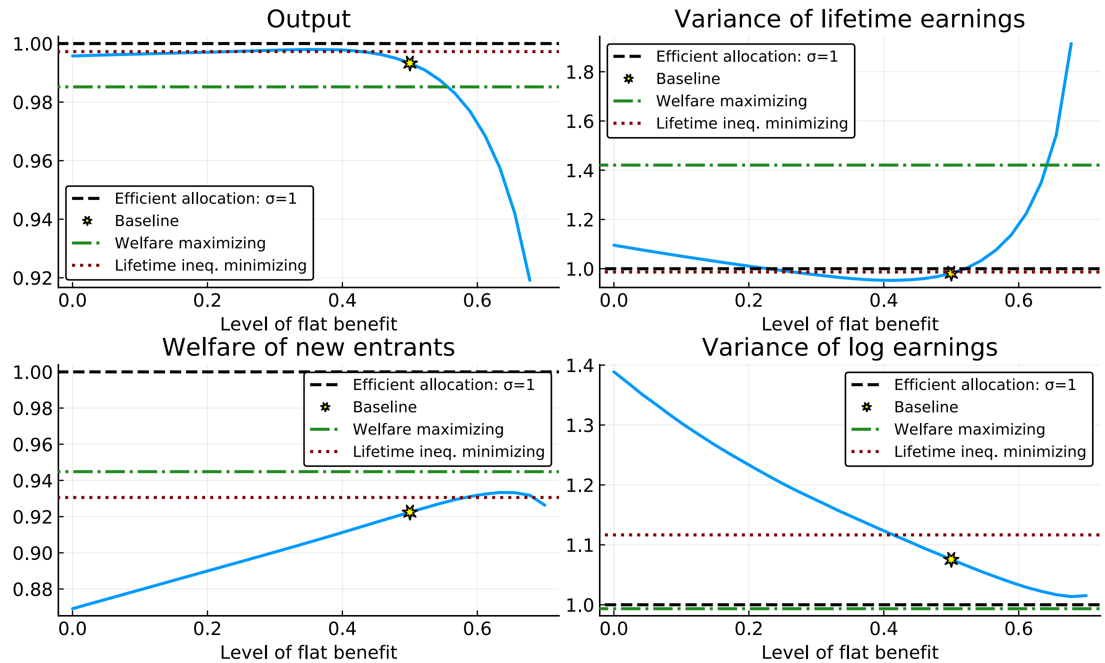

Figure 13: Output, welfare, lifetime earnings inequality, and cross-sectional earnings inequality across different UI schedules. The horizontal lines correspond to the levels achieved by either the planner, the age-dependent UI schedule that maximizes welfare, or the age-dependent UI schedule that minimizes lifetime earnings inequality, as discussed in Section 8.4. All are normalized so that the values are 1 for the planner's allocation.

8.3 Flat UI schedules

I start off by studying the effects of varying the level of the replacement rate of human capital, still keeping it constant across age. I consider z_1 between -0.5 and 0.15 with $z_2 = \infty$, corresponding to replacement rates between 0 and 0.65.[62]

Figure 13 shows the impacts of these different policies on output, welfare of new entrants, lifetime income inequality, and the variance of log earnings. The dashed black line indicates the level achieved by the planner's allocation, which is normalized to 1 in all sub-figures. To understand the effects of changing the level of flat UI benefits, consider the paths drawn by the solid blue lines. The starred points indicate the outcomes from the baseline model with $z_1 = 0$, or a replacement rate of 0.5. As benefits are raised, output rises slightly as workers prefer to accept better jobs that enable them to produce more and accumulate more human capital. It drops off steeply if benefits get too high because unemployment goes up, directly impacting output. It also indirectly impacts output as less human capital is accumulated because workers spend more time in unemployment. These opposing forces are what generate the relative flatness of output for modest levels of UI. The welfare of new entrants maxes out at some point beyond which lifetime utility starts

[62]Above this point, unemployment becomes too high, which means there are less employed workers to fund the UI benefits, which means the tax becomes too high, which further discourages working, and the economy disintegrates.

to decrease as workers expect to be unemployed for longer, pay more taxes, and accumulate less human capital. The U-shaped pattern of lifetime income inequality arises because higher benefits initially induce all workers to take up jobs that boost their lifetime earnings. Eventually, however, inequality rises because some workers luck out and find good initial jobs quickly, whereas others are induced to wait a long time to find good jobs, which means their human capital stagnates in the meantime. The variance of log earnings declines with the benefit level. This is because in the cross-section workers accept a smaller set of jobs, reducing inequality among employed workers. Many of the outcomes seen here are reminiscent of results like Acemoglu and Shimer (1999) in which unemployment insurance increases output and Acemoglu (2001) in which unemployment insurance shifts the employment distribution towards “good" jobs and improves welfare. In both settings, UI benefits allow workers to “find the right match" but I emphasize that this is especially important early in life.

The efficient allocation maximizes welfare and achieves relatively low levels of both lifetime and cross-sectional inequality. There are flat benefits schemes that can come close to achieving both levels of inequality, but they correspond to two different levels. To improve lifetime inequality, the level of benefits cannot be too high in order to prevent stagnation of human capital in unemployment. However, neither can achieve the welfare levels associated with the efficient allocation. Going from the baseline to the best flat UI schedule only brings welfare about 13% closer to the welfare achieved by the planner.

8.4 Age-dependent UI schedules

Next, I ask whether an age-dependent UI schedule can generate an allocation closer to the planner's. Targeting unemployment benefits towards the workers whose policy functions are least aligned with the planner's should improve welfare while at the same time not hurting other workers (through the tax) too much. In this economy, young workers' decisions are most misaligned because of the undervaluation of learning environment. Thus, they are the ones who should be steered the most into high learning environment firms.[63] So in this section, I will focus on benefits schedules that are high when workers are young, and drop off as they age.

The form of the parameterized $b(t)$ function means that each age-dependent schedule will be characterized by an intercept, controlled by z_1, that determines the overall level and a “slope", controlled by z_2, that determines how steeply they drop off.[64] For more details on how these parameters together change outcomes, see Appendix H. To search for a UI schedule that would maximize welfare, I fixed the intercept at the productivity level of the best firm.[65] This rules out schedules that may pay an unemployed worker more than he or she would earn in any employ-

[63] In principle, I could have made the replacement rate depend on all of the state variables including ability and human capital. However, in the real world these are hard to observe.

[64] In this part, I re-cast $b(t)$ to $z(t)h$. The (z_1, z_2) parameters should be interpreted in this context. To compute the “additional" transfer needed to find the corresponding tax, I just subtracted off the baseline benefit level, $0.5h$.

[65] This corresponds to $p = z_1 = 1.067$.

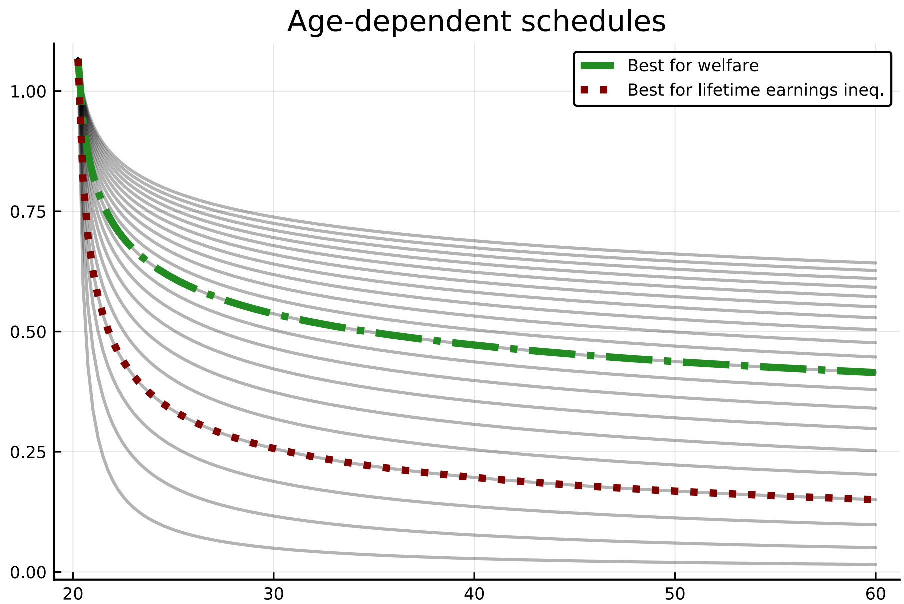

Figure 14: Age-dependent policies that achieve different outcomes. The greyed out lines vary the slope of the replacement rate function.

ment relationship. This restriction constrains me to looking at schedules of the form displayed in Figure 14. I focus on the outcomes of two of these schedules, one that maximizes welfare and one that minimizes lifetime earnings inequality.

In both cases, younger workers receive higher UI. These schedules encourage workers to be more selective across jobs early on in life compared to what they would do in the baseline. However, they drop off quickly in order to incentivize workers to become less selective and accept a job before the window to accumulate human capital runs out. Michelacci and Ruffo (2015) also find that higher UI is optimal for younger workers. Their result is driven by the fact that the young tend to be unable to smooth consumption during unemployment, and want jobs anyway to in order to accumulate human capital. In this paper, the human capital accumulation incentive is also there, but there is the additional uncertainty of whether the jobs a worker encounters will have good opportunities for human capital accumulations. Here, higher benefits while young compensates workers for the risk of not finding a good match right away. My result also has a similar flavor to Farhi and Werning (2013), who find that labor taxes should rise with age for optimal risk-sharing. In my setting, the *transfer decreases* with age, while the main tax burden is on older workers who face less uncertainty over their lifetime earnings.

The red and green dashed lines in Figure 13 indicate the levels of output, welfare, lifetime inequality, and cross-sectional inequality achieved by each of the age-dependent policies highlighted in Figure 14. The lifetime inequality-minimizing policy improves welfare, but at the cost of worsening cross-sectional inequality. This is result is again related to the overall level of these two benefits schedules: a lower level is needed in order to generate low levels of lifetime earnings inequality. Moving to the welfare-maximizing schedule from the baseline brings welfare 29% closer to the planner's allocation.

In all policies considered here, changes in the UI benefits schedule work by altering job acceptance strategies. Because these are most important for young workers, and because so many worker outcomes are determined by events early on in the life cycle, unemployment benefits have impacts on aggregates like output and unemployment, as well as earnings inequality. To further improve on the welfare gains seen here, an alternative policy would need to be designed that steers workers specifically toward high learning environment jobs. Raising the unemployment benefits simply increases reservation levels in both learning environment and productivity. Nevertheless, these experiments highlight an important function for unemployment insurance design beyond just insuring workers against short-term job loss. The results suggest that age and the role of UI for incentivizing workers to find the right match (not just any match) should be taken into account when designing these policies.

9 Conclusion

In this paper, I demonstrated that heterogeneity in learning environments between firms is a major driver of lifetime earnings inequality across workers. Motivated by the fact that firms offer systematically different earnings trajectories to the workers they employ, I developed a search model in order to disentangle the various sources of earnings growth heterogeneity. In the model, earnings can grow due to differences in worker ability, firm learning environment, and firm productivity.

In my setting, two similar workers can end up with very different levels of human capital due to differences in the firms by which they are employed over their lives. The model also introduced key trade-offs between jobs that drive workers' decisions over the life cycle. Because the ability to accumulate human capital is highest for the young, they highly value a match with a firm with a good learning environment; eventually this firm attribute becomes irrelevant and workers switch to climbing the ladder in productivity. I exploited these age differences in sources of earnings growth in the data to discipline the relevant sources of heterogeneity in the model. I also confirmed that my measures of learning environment are correlated with characteristics of the establishment that are related to on-the-job human capital accumulation.

I showed that heterogeneity in firm learning environments are responsible for 41% of the increase in the cross-sectional earnings variance over the life cycle. Over their lives, workers are exposed to different opportunities for human capital accumulation. In this way, search frictions have a direct impact on worker heterogeneity. This result signifies that firms play an important role for firms in shaping workers' human capital. Their effects are especially important for younger workers. Although workers do eventually catch up to each other by moving to better firms, early labor market experiences persistently impact lifetime earnings.

My results speak to the importance of initial conditions upon labor market entry and offer a channel through which firm/worker matches have long-term impacts. I explored two settings that illustrate the broader importance of these findings. I showed that firms shape some of the estimated profile heterogeneity across workers, suggesting that labor income processes should account more explicitly for temporary firm/worker matches and incorporate matched employer-employee data. The fact that firms matter also means that part of earnings growth is not driven by irreparable, inherent worker heterogeneity. I demonstrated how unemployment insurance policy can balance the tradeoffs between searching for good matches and human capital accumulation, and improve welfare at the same time.

This research points to several avenues for future work. Guvenen (2007) shows that imperfect knowledge of income growth rates has ramifications for the life-cycle profile of consumption. There, agents do not know their income growth rate when they enter the labor market but learn about it after seeing income realizations. I introduce a different type of uncertainty over income growth rates that stems from which firms a worker meets. Future work should further explore the significance of this kind of risk and how to distinguish it from the learning story.

There are other mechanisms in which firms may impact the earnings growth of their employees and have lasting effects. Some firms may offer better connections to other firms. Individuals at these firms may face higher arrival rates or be more likely to contact better employers. This explanation could point to another way in which search frictions impact the long-term outcomes of workers, without directly affecting workers' skills. To fully understand the long-term impacts of temporary matches, this story could be a worthwhile next step.

References

Abowd, John M, and David Card. 1989. "On the covariance structure of earnings and hours changes." *Econometrica: Journal of the Econometric Society*, 411–445.

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

Abowd, John M, Francis Kramarz, Paul Lengermann, and Sébastien Pérez-Duarte. 2004. "Are good workers employed by good firms? A test of a simple assortative matching model for France and the United States." *Unpublished Manuscript*.

Acemoglu, Daron. 2001. "Good jobs versus bad jobs." *Journal of labor Economics*, 19(1): 1–21.

Acemoglu, Daron, and Robert Shimer. 1999. "Efficient unemployment insurance." *Journal of political Economy*, 107(5): 893–928.

Alvarez, Jorge, Niklas Engbom, and Christian Moser. 2015. "Firms and the Decline of Earnings Inequality in Brazil." Working paper.

Andrews, Martyn J, Len Gill, Thorsten Schank, and Richard Upward. 2008. "High wage workers and low wage firms: negative assortative matching or limited mobility bias?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(3): 673–697.

Arellano-Bover, Jaime. 2019. "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size."

Bagger, Jesper, Fran Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. 2014. "Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics." *American Economic Review*, 104(6): 1551–96.

Baker, Michael. 1997. "Growth-rate heterogeneity and the covariance structure of life-cycle earnings." *Journal of labor Economics*, 15(2): 338–375.

Becker, Gary S. 1962. "Investment in human capital: A theoretical analysis." *Journal of political economy*, 70(5, Part 2): 9–49.

Ben-Porath, Yoram. 1967. "The production of human capital and the life cycle of earnings." *Journal of political economy*, 75(4, Part 1): 352–365.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2019. "A distributional framework for matched employer employee data." *Econometrica*, 87(3): 699–739.

Borovičková, Katarína, and Robert Shimer. 2017. "High Wage Workers Work for High Wage Firms." National Bureau of Economic Research.

Bowlus, Audra J, and Huju Liu. 2013. "The contributions of search and human capital to earnings growth over the life cycle." *European Economic Review*, 64: 305–331.

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

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace heterogeneity and the rise of West German wage inequality." *The Quarterly Journal of Economics*, 128(3): 967–1015.

Dey, Matthew S, and Christopher J Flinn. 2005. "An equilibrium model of health insurance provision and wage determination." *Econometrica*, 73(2): 571–627.

Eberle, Johanna, and Alexandra Schmucker. 2017. "Creating Cross-sectional Data and Biographical Variables with the Sample of Integrated Labour Market Biographies 1975-2014." *FDZ-Methodenreport*, 6: 2017.

Engbom, Niklas. 2020. "Labor Market Fluidity and Human Capital Accumulation."

Engbom, Niklas, and Christian Moser. 2020. "Firm Pay Dynamics." *Available at SSRN 3531250*.

Farhi, Emmanuel, and Iván Werning. 2013. "Insurance and taxation over the life cycle." *Review of Economic Studies*, 80(2): 596–635.

Fischer, Gabriele, Florian Janik, Dana Müller, and Alexandra Schmucker. 2009. "The IAB Establishment Panel-things users should know." *Schmollers Jahrbuch*, 129(1): 133.

Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter. 2005. "Imputation rules to improve the education variable in the IAB employment subsample." *ZEW-Centre for European Economic Research Discussion Paper*, , (05-010).

Friedrich, Benjamin, Lisa Laun, Costas Meghir, and Luigi Pistaferri. 2019. "Earnings dynamics and firm-level shocks." National Bureau of Economic Research.

Gartner, Hermann. 2005. "The imputation of wages above the contribution limit with the German IAB employment sample." *FDZ Methodenreport*, 2(2005): 2005.

Gourieroux, Christian, Alain Monfort, and Eric Renault. 1993. "Indirect inference." *Journal of applied econometrics*, 8(S1): S85–S118.

Gregory, Victoria, Guido Menzio, and David G Wiczer. 2020. "Pandemic Recession: L or V-Shaped?" National Bureau of Economic Research.

Guvenen, Fatih. 2007. "Learning your earning: Are labor income shocks really very persistent?" *American Economic Review*, 97(3): 687–712.

Guvenen, Fatih. 2009. "An empirical investigation of labor income processes." *Review of Economic dynamics*, 12(1): 58–79.

Håkanson, Christina, Erik Lindqvist, and Jonas Vlachos. 2015. "Firms and skills: the evolution of worker sorting." Working Paper, IFAU-Institute for Evaluation of Labour Market and Education Policy.

Hause, John C. 1980. "The fine structure of earnings and the on-the-job training hypothesis." *Econometrica (pre-1986)*, 48(4): 1013.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante. 2010. "The macroeconomic implications of rising wage inequality in the United States." *Journal of political economy*, 118(4): 681–722.

Heckman, James J. 1976. "A life-cycle model of earnings, learning, and consumption." *Journal of political economy*, 84(4, Part 2): S9–S44.

Heckman, James J, Lance Lochner, and Christopher Taber. 1998. "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents." *Review of economic dynamics*, 1(1): 1–58.

Heining, Jörg, Wolfram Klosterhuber, Stefan Seth, et al. 2014. "An overview on the Linked Employer-Employee Data of the Institute for Employment Research (IAB)." *Schmollers Jahrbuch*, 134(1): 141–148.

Herkenhoff, Kyle, Jeremy Lise, Guido Menzio, and Gordon M Phillips. 2018. "Production and Learning in Teams." National Bureau of Economic Research.

Hornstein, Andreas, Per Krusell, and Giovanni L Violante. 2011. "Frictional wage dispersion in search models: A quantitative assessment." *American Economic Review*, 101(7): 2873–98.

Huggett, Mark, Gustavo Ventura, and Amir Yaron. 2011. "Sources of lifetime inequality." *American Economic Review*, 101(7): 2923–54.

Jarosch, Gregor. 2015. "Searching for job security and the consequences of job loss." *manuscript, Stanford University*.

Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg. 2019. "Learning from Coworkers." National Bureau of Economic Research.

Kahn, Lisa B. 2010. "The long-term labor market consequences of graduating from college in a bad economy." *Labour Economics*, 17(2): 303–316.

Kambourov, Gueorgui, and Iourii Manovskii. 2009. "Occupational specificity of human capital." *International Economic Review*, 50(1): 63–115.

Karahan, Fatih, Serdar Ozkan, and Jae Song. 2019. "Anatomy of lifetime earnings inequality: Heterogeneity in job ladder risk vs. human capital." *FRB of New York Staff Report*, , (908).

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten. 2019. "Leave-out estimation of variance components." National Bureau of Economic Research.

Klosterhuber, Wolfram, Jörg Heining, Stefan Seth, et al. 2013. "Linked employer–employee data from the IAB: LIAB longitudinal model 1993–2010 (LIAB LM 9310)." *Institute for Employment Research, Nuremberg, FDZ-Datenreport*, 8: 2013.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2018. "Life cycle wage growth across countries." *Journal of Political Economy*, 126(2): 797–849.

Lazear, Edward P. 2009. "Firm-specific human capital: A skill-weights approach." *Journal of political economy*, 117(5): 914–940.

Lise, Jeremy, and Fabien Postel-Vinay. 2015. "Multidimensional skills, sorting, and human capital accumulation." Vol. 386.

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin. 2016. "Matching, sorting and wages." *Review of Economic Dynamics*, 19: 63–87.

Low, Hamish, Costas Meghir, and Luigi Pistaferri. 2010. "Wage risk and employment risk over the life cycle." *American Economic Review*, 100(4): 1432–67.

Luttmer, Erzo GJ. 2014. *An Assignment Model of Knowledge Diffusion and Income Inequality.* Federal Reserve Bank of Minneapolis, Research Department.

MaCurdy, Thomas E. 1982. "The use of time series processes to model the error structure of earnings in a longitudinal data analysis." *Journal of econometrics*, 18(1): 83–114.

Meghir, Costas, and Luigi Pistaferri. 2004. "Income variance dynamics and heterogeneity." *Econometrica*, 72(1): 1–32.

Meghir, Costas, and Luigi Pistaferri. 2011. "Earnings, consumption and life cycle choices." In *Handbook of labor economics*. Vol. 4, 773–854. Elsevier.

Michelacci, Claudio, and Hernán Ruffo. 2015. "Optimal life cycle unemployment insurance." *American Economic Review*, 105(2): 816–59.

Nagypál, Éva. 2007. "Learning by doing vs. learning about match quality: Can we tell them apart?" *The Review of Economic Studies*, 74(2): 537–566.

Rubinstein, Ariel. 1982. "Perfect equilibrium in a bargaining model." *Econometrica: Journal of the Econometric Society*, 97–109.

Rubinstein, Yona, and Yoram Weiss. 2006. "Post schooling wage growth: Investment, search and learning." *Handbook of the Economics of Education*, 1: 1–67.

Schaefer, Daniel, and Carl Singleton. 2016. "Recent changes in British wage inequality: Evidence from firms and occupations."

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. 2015. "Firming up inequality." National Bureau of Economic Research.

Sørensen, Kenneth Lykke, and Rune Vejlin. 2011. "Worker and firm heterogeneity in wage growth: an AKM approach." *Labour*, 25(4): 485–507.

A Construction of main sample

From the raw data, I construct a monthly panel of workers which is used as the basis for all of the analyses in this paper.

The data arrive in spell format which tell me the exact start and end dates of the employment spell, or registration in the unemployment benefits system. Employment spells are always contained within a single calendar year, and therefore do not last longer than one year. Unemployment spells can span more than one year. It is also possible to have gaps in a given worker's employment biography.

To correct the inconsistencies and missing values in the LIAB's education variable, I apply the imputation method of Fitzenberger, Osikominu and Völter (2005). This method looks at a worker's past and future values of the education variable to impute values for the gaps.

I also correct for top-coding in the LIAB's wage variable, which represents the worker's average daily wage throughout the spell. The wage ceiling is based on the contribution limits of social security, which change from year to year and are different in the former East and West Germany. About 7-10% of wage observations per year are top-coded, and these are mainly concentrated among the college-educated group. As do many other studies which use this data source, I implement a Tobit imputation to fill in the top-coded wages. I follow the approach suggested by Gartner (2005). In each year, 12 Tobit models are estimated by education group (6 categories: missing; no qualification; vocational training degree; high-school degree; high-school degree + vocational training degree; university graduate) and gender. Let the log of the wage variable for worker i in year t be w_{it}. The Tobit model for w_{it} has $w_{it} \sim \mathcal{N}(x_{it}'\beta, \sigma)$. To impute a wage for a censored value, compute:

$$w_{it} = x_{it}'\hat{\beta} + \eta_{it}$$

η_{it} is a draw from a truncated distribution, computed as:

$$\eta_{it} = \hat{\sigma}\Phi^{-1}(k_{it} + u_i(1 - k_{it}))$$

where $u_i \sim \mathcal{U}(0,1)$, $k_{it} = \Phi\left(\frac{c_t - x_{it}'\hat{\beta}}{\hat{\sigma}}\right)$, c_t is the censoring point in year t, and $\Phi(\cdot)$ is a standard normal cdf. $\hat{\beta}$ and $\hat{\sigma}$ are estimated from a Tobit regression with age as an explanatory variable. Not that u_i does not depend on t to avoid introducing extra noise which would show up in person-level wage growth, an important component of this paper. All wages are then deflated using the German CPI.

Because the source of the LIAB is worker-level social security records, I need to merge in data from the IAB's BHP (Establishment History Panel) to obtain a richer set of characteristics about the establishment. The BHP contains the industry, size class, and location (federal state), as well as a variety of employment-related variables of all establishments that appear in the LIAB. Although the variables in the BHP are derived from the employment records upon which the LIAB is based, they enable me to observe these establishment-level characteristics in cases where the LIAB does not contain all of the records for the establishment.[66]

After merging in the imputed wages, education levels, and BHP variables, I construct a monthly panel. I record all of the variables (wages, establishment identifier, occupation, etc.) associated with a worker's job spell as long as the spell includes at least one day in a given calendar month and year combination. This is done with the help of programs which convert spells into monthly cross-sections provided by the IAB: see Eberle and Schmucker (2017). I then append these into a monthly panel spanning 1993 - 2014. From the original spell dataset, I also record the previous and subsequent employment states, as well as the number of days between them, so I can better identify job-to-job transitions and employment-to-unemployment flows later on.

Lastly, I apply some restrictions to arrive at the final set of monthly employment records. I drop part-time and marginal part-time workers, workers younger than 16, workers older than 70, and workers who earn less than 10 Euros per day. Tables A.1 and A.2 report basic summary statistics for this panel.

Worker-month observations	203,143,204
Unique workers	1,320,693
German	91.77%
Female	38.05%
Education level	
High school degree or less	9.32%
Vocational degree	73.83%
College degree	16.85%
Age	40.69
	(9.73)
Daily log earnings (2010 Euros)	4.571
	(0.552)
Number of months in sample	154.11
	(81.71)
Number of establishments per worker	3.547
	(3.055)

Table A.1: Worker summary statistics: full sample
Summary statistics for the baseline monthly panel of workers. Statistics on nationality, gender, number of months, and number of establishments are reported at the worker level; statistics on education, age, and earnings are reported at the worker-month level because these are potentially time-varying. Means are reported with standard deviations in parentheses. Time period is 1994-2014.

[66] For example, this occurs when a worker employed in one of the core sample establishments moves to one outside of the core sample. Even though the LIAB does not contain the complete set of employment records for the latter establishment, some of its basic characteristics can be found in the BHP.

Unique establishments	970,286
Worker-months per establishment	209.51
	(13123.72)
Size class	
1-4 employees	23.12%
5-9 employees	19.14%
10-19 employees	18.05%
20-49 employees	16.85%
50-99 employees	7.39%
100-199 employees	3.95%
200-499 employees	2.15%
500+ employees	0.86%
Industry	
Agriculture, hunting, forestry and fishing	2.25%
Mining and quarrying, electricity, gas and water supply	0.49%
Manufacture of food products, beverages and tobacco	1.91%
Manufacture of consumer products	2.40%
Manufacture of industrial goods	2.99%
Manufacture of capital and consumer goods	6.84%
Construction	14.54%
Trade, maintenance and repair of motor vehicles and goods	16.78%
Transport, storage and communication	6.81%
Financial intermediation	1.58%
Hotels and restaurants	5.56%
Education	1.83%
Health and social work	6.03%
Computer and related activities	1.32%
Research and development	0.31%
Legal, accounting; market research; consultancy; advertising	3.72%
Real estate activities	1.49%
Renting of equipment and personal goods; other business activities	8.34%
Other community, social and personal service activities	2.98%
Public administration, defence; private households	3.24%

Table A.2: Establishment summary statistics: full sample
Summary statistics for each establishment connected to a worker in the main monthly panel. Means are reported with standard deviations in parentheses. The size and industry groups are the ones reported by the IAB's Establishment History Panel (BHP). Time period is 1994-2014.

B Details on motivational evidence

B.1 Construction of annual panel

Much of the analysis in this paper is performed on a annual panel of workers and the establishments that they are attached to. I count the number of spells that each worker has by comparing their employment statuses in consecutive months. If they transition between employment and unemployment or between establishments, I count a new spell. To collapse the monthly panel, I record the year of hire for each job spell and calculate the worker's tenure in months. If the worker's education level or the establishment's size class changes at some point during the match, I assign the value at hiring to the entire spell. I collapse at the worker $\times$ job spell ID (which will correspond to a single establishment ID) $\times$ annual tenure level, assigning the average wage observed during each year of employment as the annual wage variable, w_{ijt}. This panel will contain observations that correspond to less than 12 months. For instance a worker with a 2.5 year employment spell will have 3 observations for the spell: years 0 to 1, years 1 to 2, and the last 6 months of the spell.

Earnings growth from year $t-1$ to t is $\Delta \log w_{ijt} = \log w_{ijt} - \log w_{ij,t-1}$. Obviously, spells with less than a year of tenure are excluded from any analysis that relies on this variable. I also trim the top and bottom 2% of this variable. In this annual panel, each worker appears for an average of 13.89 years and has an average of 3.97 employment spells.

B.2 Results from two-way fixed effects models

Detailed results from the two-way fixed effects specification (1) are displayed in Table B.1. The results here are in line with Sørensen and Vejlin (2011)'s study for Denmark which found that both worker and establishment effects had relatively low explanatory power for the variance of wage growth in the population, relative to how much is typically found for wage levels. Complete histograms of the establishment and worker effects from this estimation are in Figures B.1 and B.2.

	All	$\leq$ High School Grad	Vocational Training	College Grad
Earnings growth				
Mean $\Delta \log w_{ijt}$	0.0161	0.0133	0.0148	0.0239
Std. dev. $\Delta \log w_{ijt}$	0.0615	0.0671	0.0611	0.0572
Establishment effects				
Std. dev. ψ_j	0.0242	0.0475	0.0269	0.0271
P90 - P10	0.0443	0.0903	0.0513	0.0458
Worker effects				
Std. dev. α_i	0.0262	0.0466	0.0274	0.0305
P90 - P10	0.0537	0.0900	0.0553	0.0621
corr(α_i, ψ_j)	-0.4900	-0.8340	-0.5663	-0.5541
Std. dev. ε_{ijt}	0.0556	0.0606	0.0553	0.0495
# of person-years	13,620,563	1,517,733	9,903,449	2,121,693
# of establishments	381,191	56,045	315,367	83,669
# of workers	1,114,653	120,479	807,635	206,494
R^2	0.1829	0.1850	0.1822	0.2495

Table B.1: Full results of (1) with $\boldsymbol{\beta} = \mathbf{0}$.

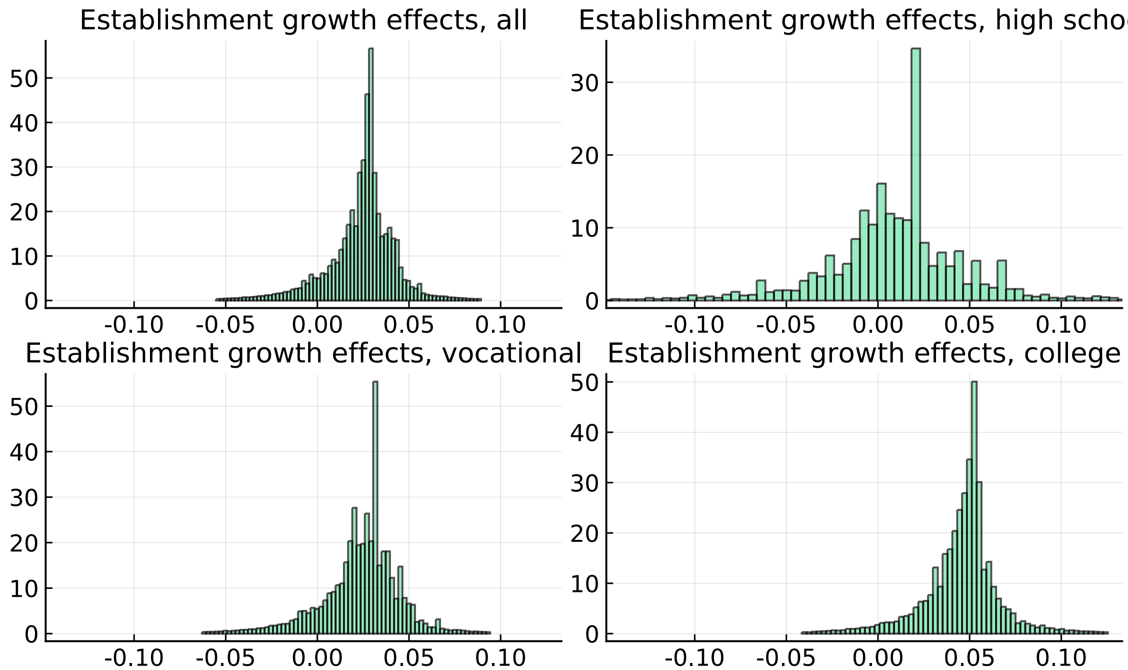

Figure B.1: Distributions of establishment wage growth fixed-effects by education group. Histograms of the estimated fixed effects for establishments, ψ_j, from equation (1) with $\boldsymbol{\beta} = \mathbf{0}$, for the full sample and broken down by education group.

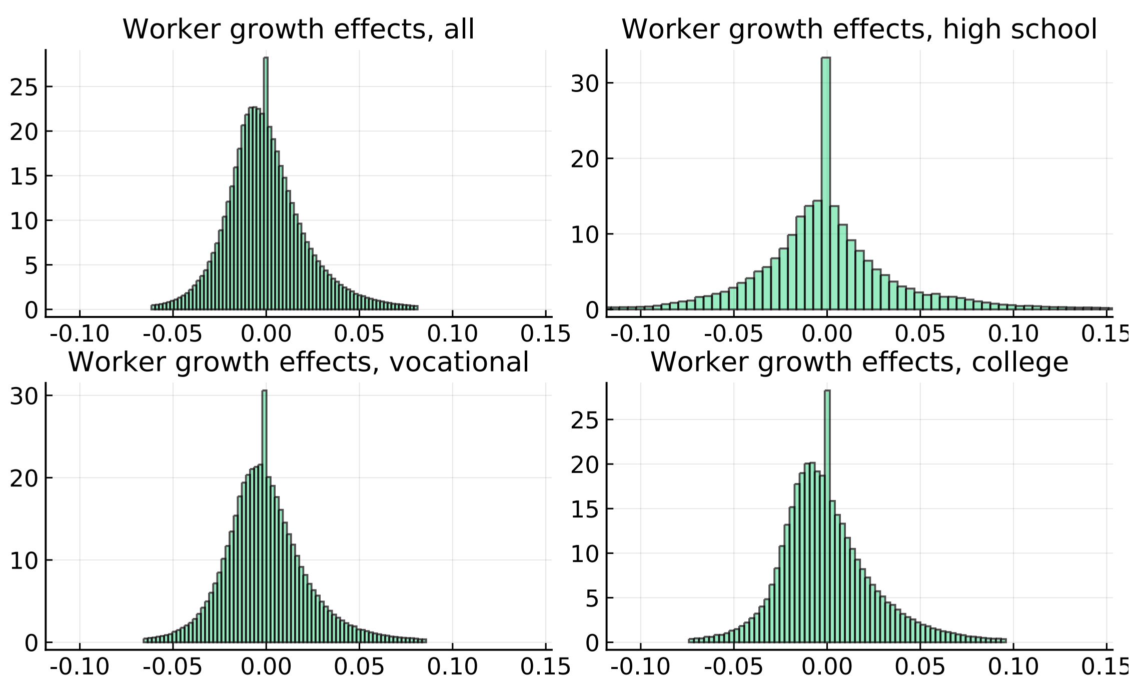

Figure B.2: Distributions of worker wage growth fixed-effects by education group. Histograms of the estimated fixed effects for workers, α_i, from equation (1) with $\boldsymbol{\beta} = \mathbf{0}$, for the full sample and broken down by education group. The α_i were normalized to have mean 0.

The full results from the estimation underlying Figure 2 are in Table B.2. Adding in the age and tenure profiles does not significantly change the dispersion in the fixed effects or fit of the model relative to the results in Table B.1.

	All	$\leq$ High School Grad	Vocational Training	College Grad
Coefficients				
age	-0.0021401^{***}	-0.0030259^{***}	-0.0014828^{***}	-0.0056794^{***}
	(0.0000222)	(0.0000891)	(0.0000270)	(0.0000690)
age^2	0.0000245^{***}	0.0000415^{***}	0.0000176^{***}	0.0000549^{***}
	(0.0000002)	(0.0000008)	(0.0000003)	(0.0000008)
tenure	-0.0030517^{***}	-0.0029908^{***}	-0.0029098^{***}	-0.003654^{***}
	(0.0000166)	(0.0000801)	(0.000200)	(0.0000377)
tenure^2	0.0001321^{***}	0.0001415^{***}	0.0001178^{***}	0.0001669^{***}
	(0.0000008)	(0.0000031)	(0.0000009)	(0.0000019)
Earnings growth				
Mean $\Delta \log w_{ijt}$	0.0161	0.0133	0.0148	0.0239
Std. dev. $\Delta \log w_{ijt}$	0.0615	0.0671	0.0611	0.0572
Establishment effects				
Std. dev. ψ_j	0.0241	0.0452	0.0267	0.0272
P90 - P10	0.0446	0.0855	0.0506	0.0466
Worker effects				
Std. dev. α_i	0.0252	0.0450	0.0268	0.0287
P90 - P10	0.0510	0.0881	0.0533	0.0561
corr(α_i, ψ_j)	-0.4960	-0.8093	-0.5664	-0.5967
Std. dev. ε_{ijt}	0.0554	0.0604	0.0552	0.0493
# of person-years	13,620,563	1,517,733	9,903,449	2,121,693
# of establishments	381,191	56,045	315,367	83,669
# of workers	1,114,653	120,479	807,635	206,494
R^2	0.1864	0.1891	0.1848	0.2575

Table B.2: Full results of (2), which adds age and annual tenure profiles to (1).

I construct establishment-specific earnings profiles depicted in Figure 2 as follows. Let $j(x)$ be the establishment at the xth percentile of the ψ_j distribution. Setting $\alpha_i = 0$, the earnings profile as a function of tenure $p_{ea}^{j(x)}(\tau)$ for a worker at that establishment with education e and hired at age a is:

$$p_{ea}^{j(x)}(\tau) = \sum_{t=1}^{\tau} \Delta \widehat{\log w_{i,j(x)}^{ea\tau}} \tag{B.1}$$

where

$$\Delta \widehat{\log w_{i,j(x)}^{ea\tau}} = \widehat{\psi_{j(x)}^{e}} + \widehat{\beta}_1^e(a+\tau) + \widehat{\beta}_2^e(a+\tau)^2 + \widehat{\beta}_3^e\tau + \widehat{\beta}_4^e\tau^2 \tag{B.2}$$

(B.2) constructs the predicted wage growth for the worker based on the estimation results of (2), which is performed separately by education level. The profiles are then calculated as the cumulative sum of predicted growth in each previous year, as in (B.1).

B.3 Sensitivity to Bias in the Establishment Fixed-Effects

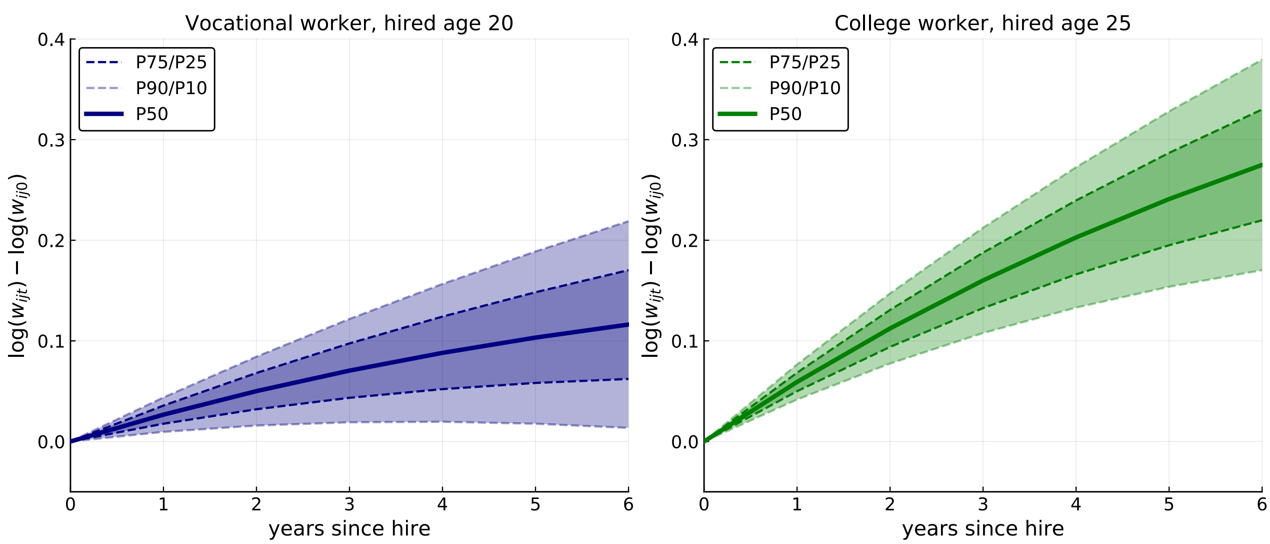

Figure B.3: Establishment-specific earnings growth profiles: $\frac{1}{4}$ variance.
Each panel depicts profiles of cumulative earnings growth as a function of tenure for workers with the same education level, age of hire, and fixed effect α_i. Estimates of the age and tenure profiles come from equation (2). Each profile is constructed by computing the predicted values of earnings growth for each implied tenure and age horizon and taking the cumulative sum. Each series from bottom to top corresponds to the earnings growth profile of the establishment at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of the establishment fixed effect, ψ_j, distribution. For more details, see Appendix B.2.

Methods that correct for limited mobility bias in two-way fixed effects models typically find a much smaller role for establishment/firm fixed-effects than what is implied by AKM (see Borovičková and Shimer (2017), Bonhomme, Lamadon and Manresa (2019), Andrews et al. (2008), and Kline, Saggio and Sølvsten (2019)). I run a senstivity check to assess how this bias impacts the heterogeneity in earnings trajectories in Figure 2. The starting point for this exercise is a Normal distribution with the same mean and variance as the original establishment fixed effect distribution. I then shrink it's variance by one-fourth. This factor is close to what other studies have found for the bias in the variance of the fixed-effects. I then plot the earnings trajectories based on that distribution. The results are presented in Figure B.3. As expected, the fanning out of the earnings profiles is less pronounced compared to Figure 2. However, there are still substantial variation in earnings growth for identical workers hired at different establishments, especially when comparing their outcomes at longer tenures.

C Details on identification method

C.1 Step 1: establishment-specific returns to search capital

The starting point for the sample in Step 1, described in Section 4.1, is the monthly panel. The sample includes only workers who start new jobs at age 50 or above, and who were previously in unemployment. I include spells in which the received unemployment benefits in the previous month, or if the last observed spell is employment in another establishment which ended between 3 weeks and 1 year ago. After keeping all of the relevant spells, I convert the monthly panel to an annual one in the same way as described in Section B.1. I then drop observations for annual wage growth in the top and bottom 2% tails.

I apply some additional refinements to arrive at the sample I use for the regression in (12) because it becomes difficult to construct establishment-level statistics based on the spells from this very specific group of workers. In order to avoid "very small" establishments, I drop establishments with less than 5 spells, pooled across all of the years the establishment is present in the panel. In order to get a more complete measure of each establishment's tenure profile, I then drop establishments with no workers who stay less than 4 years. In the end, I am left with 1,481 establishments. The final summary statistics for this sample are displayed in the first column of Table C.1.

I run a random coefficients model to compute each establishment's (α^j, β_j^1). The procedure estimates via maximum likelihood the coefficients of a bivariate normal distribution for $(\alpha^j, \beta_j^1) \sim \mathcal{N}\left(\begin{bmatrix}\mu_\alpha \\ \mu_\beta\end{bmatrix}, \begin{bmatrix}\sigma_\alpha^2 & \sigma_{\alpha\beta} \\ \sigma_{\alpha\beta} & \sigma_\beta^2\end{bmatrix}\right)$. The full set of parameter estimates with comparison to pooled OLS is reported in Table C.2.

C.2 Step 2: establishment-specific returns to human capital

The sample for step 2 is based on young workers. I start again from the monthly panel and record the date and age at which each individual is hired for each employment spell I observe. Because the panel starts in January 1993, I drop all spells that begin there because I cannot measure the true start date.

In this step, I need a measure of work experience so that I can focus on workers who have just

	Step 1	Step 2
Worker-year observations	62,660	809,356
Number of unique workers	18,718	162,655
Female	27.50%	32.89%
Education level		
High school degree or less	5.56%	36.32%
Vocational degree	72.76%	40.12%
College degree	20.75%	22.99%
Age	56.67	27.48
	(3.23)	(5.90)
Daily earnings (2010 Euros)	82.90	105.29
	(69.58)	(78.91)
Annual log earnings growth	0.011	0.078
	(0.056)	(0.148)
Tenure (years)	3.16	4.85
	(2.30)	(3.74)
Number of unique establishments	1,481	1,058
Number of workers per establishment	13.90	159.39
	(22.99)	(634.24)

Table C.1: Summary statistics for samples of workers used to construct residual growth moments. Construction of residuals growth moments is described in Section 4.1. The number of workers per establishment refers to the number of workers *in that sample* per establishment, not the overall number of employees.

	Step 1			Step 2	
	OLS	**Random Coeffs.**		**OLS**	**Random Coeffs.**
α, μ_α	0.02044	0.02234	γ, μ_γ	0.11943	0.11073
	(0.00061)	(0.00090)		(0.00044)	(0.00205)
β^1, μ_β	-0.00464	-0.00530	δ^1, μ_δ	-0.01067	-0.01207
	(0.00035)	(0.00036)		(0.00015)	(.00025)
β^2	0.00032	0.00037	δ^2	0.00013	0.00031
	(0.00004)	(0.00004)		(0.00001)	(0.00001)
σ_α	-	0.02395	σ_γ	-	0.06076
		(0.00066)			(0.00156)
σ_β	-	0.00331	σ_δ	-	0.00585
		(0.00016)			(0.00017)
$\text{corr}(\alpha, \beta^1)$	-	-0.88767	$\text{corr}(\gamma, \delta^1)$	-	-0.98240
		(0.01490)			(0.00216)
# of worker-years	62,660	62,660		809,356	809,356
# of establishments	-	1,481		-	1,058
min(worker-years/estab.)	-	6		-	8
max(worker-years/estab.)	-	1,293		-	117,546

Table C.2: Parameter estimates for equations (12) and (13). Standard errors in parentheses. $\alpha, \beta, \gamma, \delta$ refer to the OLS estimates; $\mu_\alpha, \mu_\beta, \mu_\gamma, \mu_\delta$ refer to the random coefficients estimates.

entered the labor market. Age is not perfect indicator of this since people start their careers at different ages. Moreover, in the model everyone enters the labor force at the same age, so I need some way to ensure that everyone in the data is starting from a common point corresponding to zero on-the-job human capital. To this end, I apply the following two restrictions to identify each worker's first "career" job. First, I only keep spells that start within an acceptable age range which depends on the worker's education level: the more education they have, the later they are expected to enter the labor force. These ranges are 17-21 for workers with less than a high school degree, 19-23 for workers with a high school degree or vocational degree, 21-27 for workers with both a high school degree and vocational degree, 24-30 for workers with a college degree, 19-23 for workers with a missing education level. Second, I keep only spells that last at least 90 days. Starting with the first job that meets these requirements, I can now count each worker's experience level.

I then apply some further restrictions which are the same as in step 1: dropping the top and bottom 2% tails of annual earnings growth, dropping establishments with less than 5 spells, and dropping establishments with no workers who stay less than 4 years. In the end, the number of establishments to which I am able to apply both step 1 and step 2 is 1,058. The final summary statistics for this sample are displayed in the second column of Table C.1.

Once I have this sample of young workers, I remove the part of earnings growth coming from search capital by using the fitted values for establishment j estimated in step 1: $\Delta \log \widehat{\text{earnings}}_{ijt} = \Delta \log \text{earnings}_{ijt} - \hat{\alpha}_j - \hat{\beta}_j^1 \text{tenure}_{it} - \hat{\beta}_2 \text{tenure}_{it}^2$. I then estimate the random coefficients model specified in (13). The full set of parameter estimates is reported in Table C.2.

The moments displayed in Figure 7 are constructed in the same way as the data in Figure 2. At the establishment level, I first compute the expected cumulative growth in residual earnings at horizon τ:

$$LE_{j\tau} = \sum_{h=0}^{\tau} \hat{\gamma}_j + \hat{\delta}_j^1 \tau + \hat{\delta}^2 \tau^2$$

I then target $\text{mean}_j(LE_{j\tau}), \text{p10}_j(LE_{j\tau})$, and $\text{p90}_j(LE_{j\tau})$ for $\tau = 1, 2, \ldots, 10$ in the estimation, as displayed in Figure 7.

C.3 Step 3: correlation between returns to human capital and search capital

For each establishment, I compute a single measure of the returns to human capital, $LE_j = \text{mean}_\tau(LE_{j\tau})$, and the returns to search capital, $SC_j = \text{mean}_\tau(SC_{j\tau})$ where

$$SC_{j\tau} = \sum_{h=0}^{\tau} \hat{\alpha}_j + \hat{\beta}_j^1 \tau + \hat{\beta}^2 \tau^2$$

The pairwise correlation is then used in the estimation. These LE_j's and SC_j's are also what is reported in Tables D.1 and D.2.

C.4 Robustness to random coefficients

In this section, I discuss the sensitivity of the results to the use of the random coefficients models. These models enable me to estimate individual slope parameters for each establishment without having to impose too many size restrictions on the establishments. The baseline version keeps only establishents with at least 5 worker spells, where at least one of them is a minimum of 5 years in length.

I can get similar results using establishment-by-establishment OLS. However I need to apply stricter establishment size restrictions. In this alternative version, I only keep establishments that have at least 5 workers who each stay longer than 5 years. These restrictions leave me with 251 establishments, rather than 1058 in the baseline.

Figure C.1 compares the two approaches. In either case, the residual growth moments look similar.

In the end, I chose the weak sample selection with the random coefficients model as my baseline because I can assign learning environments to a larger set of establishments.

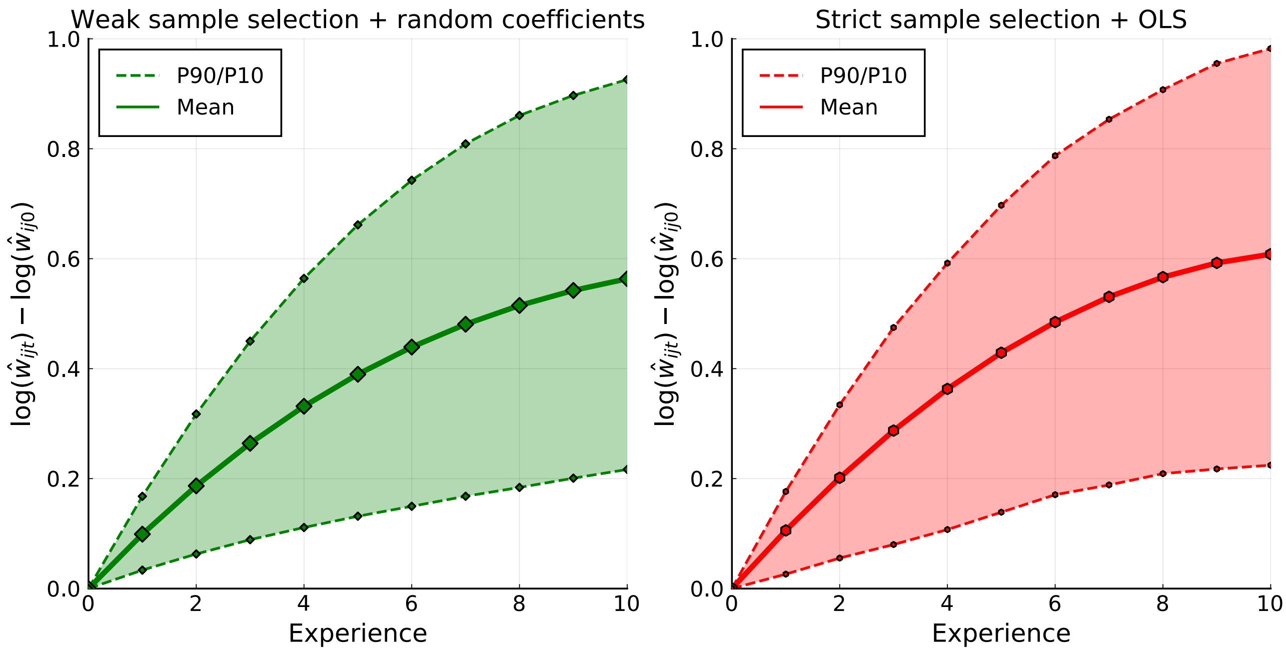

Figure C.1: Residual earnings growth moments: alternative approaches.
The left panel displays the residual growth moments for my baseline specification, the same as in Figure 7. The right panel displays the same moments, but with a stricter size requirement for establishments and using OLS on each establishment separately, as desribed in Appendix C.4.

D Learning environment and establishment characteristics

Establishment size	Productivity	Learning Env.	# Estabs.
1-4 employees	-	-	-
5-9 employees	-	-	-
10-19 employees	-	-	-
20-49 employees	0.010 (0.0047)	0.037 (0.0124)	112
50-99 employees	0.011 (0.0057)	0.038 (0.0126)	155
100-199 employees	0.012 (0.0073)	0.035 (0.0141)	208
200-499 employees	0.012 (0.0061)	0.037 (0.0141)	294
500+ employees	0.013 (0.0074)	0.039 (0.0137)	263
All establishments	0.012 (0.0060)	0.037 (0.0146)	1058

Table D.1: Productivity and learning environment by establishment size
Productivity and learning environment measures are based on the procedure outlined in Section 5.3. Means are reported, with standard deviations in parentheses. Size classes with blanks had less than 20 establishments included in the estimation, which cannot be reported due to data disclosure regulations. The size class categorization is as reported by the IAB's Establishment History Panel (BHP).

Industry	Productivity	Learning Env.	# Estabs.
Agriculture, hunting, forestry and fishing	0.012 (0.0053)	0.037 (0.0111)	35
Mining and quarrying, electricity, gas and water supply	-	-	-
Manufacture of food products, beverages and tobacco	0.009 (0.0067)	0.044 (0.0122)	31
Manufacture of consumer products	0.009 (0.0064)	0.042 (0.0107)	26
Manufacture of industrial goods	0.012 (0.0056)	0.040 (0.0128)	109
Manufacture of capital and consumer goods	0.012 (0.0064)	0.045 (0.0134)	184
Construction	0.010 (0.0052)	0.045 (0.0144)	81
Trade, maintenance and repair of motor vehicles and goods	0.011 (0.0074)	0.042 (0.0137)	47
Transport, storage and communication	0.009 (0.0060)	0.038 (0.0138)	34
Financial intermediation	-	-	-
Hotels and restaurants	-	-	-
Education	0.007 (0.0073)	0.021 (0.0081)	53
Health and social work	0.010 (0.0053)	0.028 (0.0083)	65
Computer and related activities	-	-	-
Research and development	-	-	-
Legal, accounting; tax consultancy; market research; business consultancy; holdings; advertising	-	-	-
Real estate activities	-	-	-
Renting of machinery, equipment, personal, and household goods; other business activities	0.014 (0.0062)	0.028 (0.0092)	178
Other community, social and personal service activities	0.014 (0.0049)	0.037 (0.0119)	32
Public administration, defence; private households	0.014 (0.0047)	0.037 (0.0122)	135
All establishments	0.012 (0.0062)	0.037 (0.0139)	1058

Table D.2: Productivity and learning environment by industry
These measures are based on the procedure outlined in Section 5.3. Means are reported, with standard deviations in parentheses. Industries with blanks had less than 20 establishments included in the estimation, which cannot be reported due to data disclosure regulations. The industry categorization is as reported by the IAB's Establishment History Panel (BHP).

E Details on IAB Establishment Panel link

The raw dataset from the IAB Establishment Panel comes in separate files for each year, 1993 - 2016. With the help of programs provided by the IAB, I merge these together to create a single panel file with a consistent naming system for the variables that appear in multiple years. In this step, only the variable blocks related to general information about the establishment, further on-the-job training, and apprenticeship programs are retained.

For the training module, my main variables of interest are the types of training that the establishment offers, as well as the topics that the training focuses on. First, the establishment is asked if they released staff for the purpose of participating in training courses and covered the expenses in full or in part (the answer to this question is entitled “offers any training" in my tables). If they answer affirmatively, they are then asked to check off items on a list for the types of training they offered in that year. These are indicated under the list “Types of training offered" in the tables. This set of questions is asked in the years 1997, 1999, 2003, 2005, and 2007 - 2016. Additionally, in 2001, there is a special module which asks about what the top two most important topics of the training programs were. These options are under “1^{st} or 2^{nd} most important training topic" in the tables.

In the apprenticeship modules, the most consistent and relevant questions ask how many apprentices complete the program and are retained as full-time employees of the establishment. The establishment is first asked if it is qualified to provide adequate vocational training in compliance with the laws surrounding these arrangements. This variable is “fulfills educational requirements" in the tables. If they answer “yes," they are then asked further questions about their apprenticeship program. The most consistenly available and relevant questions involve completion and retainment. The survey asks how many workers completed apprenticeships this year – I study that number as a fraction of all apprentices. They also ask how many apprentices were retained as regular employees. I am also interested in this as a fraction of all completed apprenticeships. All of these questions are asked every year after 1997.

My learning environment measure is one value per establishment that is derived off of all of employment records associated with the establishment in all of years in which they appear in the LIAB. However, each of the survey questions in the IAB Establishment panel is potentially answered by the establishment in multiple years. Since most of the variables correspond to “yes" or “no" questions, I aggregate them to a single value per establishment by calculating the fraction of years in which they answered “yes." Exceptions are the fraction of employees receiving further training, the fraction of apprentices retained (out of all completed apprenticeships), and the fraction of successfully completed apprenticeships (out of all apprentices). For these, I aggregate by taking the average across years.

The summary statistics for all of the variables are in Tables E.1 and E.2. I report them for both the entire sample and for the sample of establishments that also have estimated learning environments. The latter group is where all of my main analysis is done. Both on-the-job training and apprenticeship programs are more common in this sample. This is a consequence of only being able estimate learning environment for larger establishments, who are more likely to have the infrastructure to offer more formal training.

Variable	Full Sample	Learning Env. Sample
Offers any training	0.609 (0.434)	0.881 (0.22)
Fraction of employees receiving training	0.399 (0.352)	0.302 (0.249)
Types of training offered		
Number of types of training per year	3.31 (1.601)	4.075 (1.431)
Number of types offered all years	4.257 (2.0)	6.076 (1.668)
External courses, seminars, or workshops	0.848 (0.307)	0.911 (0.193)
Internal courses, seminars, or workshops	0.61 (0.43)	0.771 (0.305)
Further training on-the-job (instruction, initial skill adaptation training)	0.602 (0.418)	0.744 (0.288)
Participation in lectures, symposia, fairs, etc.	0.59 (0.423)	0.725 (0.319)
Job rotation	0.11 (0.263)	0.186 (0.271)
Self-directed study	0.239 (0.361)	0.303 (0.328)
Quality circles, workshop circles, continuous improvement teams	0.153 (0.309)	0.216 (0.299)
Other	0.158 (0.301)	0.218 (0.263)
1^{st} or 2^{nd} most important training topic		
Business topics	0.356 (0.479)	0.271 (0.445)
Commercial, scientific, technical, design topics	0.359 (0.48)	0.405 (0.491)
EDP, information/communication technology	0.644 (0.479)	0.662 (0.473)
Soft skills (e.g. ability to work in team, conflict management, work organization)	0.37 (0.483)	0.403 (0.491)
Other	0.302 (0.459)	0.29 (0.454)
Number of establishments	63,670	913

Table E.1: Summary statistics for further on-the-job training.
This table reports the mean of each variable in the two samples, with standard deviation in parentheses. All represent binary variables (taking value 0 or 1) except for the fraction of employees receiving training and the number of types per year and in all years. All variables (except for “Offers any training") are reported only for establishments who offered training in at least one year in the panel. Because the questions in the bottom panel are only asked in 2001, establishments who enter the panel after that or exit before that do not get asked this question. Consequently, the sample size for each mean is lower: 7,705 for the full sample and 462 for the learning environment sample.

Variable	Full Sample	Learning Env. Sample
Fulfills educational requirements	0.622 (0.444)	0.856 (0.266)
All apprentices retained	0.511 (0.419)	0.536 (0.344)
Fraction of apprentices retained	0.456 (0.371)	0.537 (0.313)
Has successfully completed apprenticeships	0.424 (0.428)	0.673 (0.343)
Fraction of successfully completed apprenticeships	0.581 (0.396)	0.469 (0.211)
Number of establishments	64,983	906

Table E.2: Summary statistics for apprenticeship.
This table reports the mean of each variable in the two samples, with standard deviation in parentheses. "Fulfills educational requirements," "All apprentices retained," and "Has successfully completed apprenticeships" represent binary variables (taking value 0 or 1). "All apprentices retained" is reported for all establishments with completed apprenticeships, and for this same group "Fraction of apprentices retained" is reported as a fraction of all completed apprenticeships. Similarly, "Has successfully completed apprenticeships" is reported for all establishments that have apprentices, and for this same group "Fraction of successfully completed apprenticeships" is reported as a fraction of all apprentices.

F Additional decompositions

F.1 Decomposition of earnings growth and variance, shutting down *firm* heterogeneity

G Details on the earnings process estimation

To construct the empirical covariances needed for the estimation of the income process from the LIAB micro data, I need to aggregate my monthly panel to the worker-age level. I calculate the age of the worker as the year recorded in each monthly spell subtracted by their birth year (I do not observe their exact date of birth). If a worker holds multiple jobs in a year, I simply average the monthly earnings in that year because I cannot distinguish the point at which they changed ages within the year. This is my measure of earnings by age for each individual in the panel.

As is common in the literature, I calculate residualized earnings, y_h^i. I regress log earnings on a set of age, education, gender, and year dummies and extract the residuals to use in the cross-covariances. From there, I calculate $\text{cov}(y_h^i, y_{h+n}^i)$ as:

$$\text{cov}(y_h^i, y_{h+n}^i) = \frac{1}{N_{h,h+n}} \sum_{i=1}^{N_{h,h+n}} y_h^i y_{h+n}^i$$

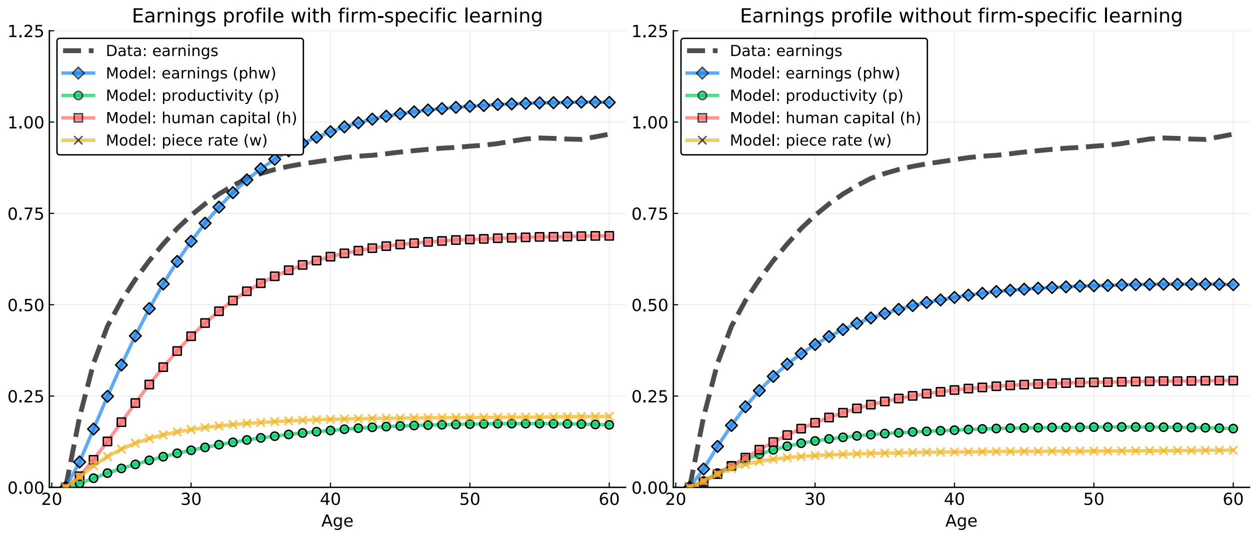

Figure F.1: Life-cycle mean of log earnings and decomposition.
The left panel plots the mean of log earnings in the data and in the model where firms are heterogeneous in q. The right panel plots the corresponding means in the version of the model in which there is no firm-specific component of human capital accumulation: $q = 0$ for all firms so that all growth in human capital solely comes from worker learning abilities. Each series is derived from the profile of mean log earnings by age. Each is normalized to zero at the start by subtracting the value at age 20.

for the set of workers who are observed at both ages h and $h+n$. $N_{h,h+n}$ is the number of these workers. Because of the length of the panel, n is never greater than 20.

When estimating the earnings process on model-simulated data, I apply the exact same restrictions, as long as each worker is employed for at least 1 quarter in each year (so that I have some labor earnings observations for that age). I also residualize earnings based only on age, since there are no education, gender, or year effects in the model.

Figure G.1 compares the autocovariance structure of earnings across ages in the model and data. All of them exhibit a downward sloping pattern because the relationship between earnings at different ages becomes weaker as the horizon increases. This figure also points to the main source of differences in the model and data parameter estimates: the autocovariances decay a lot less gradually in the model. I suspect this is because the model abstracts from forms of heterogeneity that would impact the within-worker patterns of job draws or transition rates. For example, some workers may be ex-ante less likely to separate into unemployment, have longer job durations, and be more likely to draw better matches (see, for instance, the forms of worker heterogeneity in Gregory, Menzio and Wiczer (2020)). Modifications like these would generally bring up the within-person correlation of earnings at different ages. Nevertheless, these autocorrelations were never targeted in the model and adding features that better match them goes beyond what's needed for the main goal of the paper.

To obtain the GMM estimates of the parameters of the income process, I stack all "empirical" (from

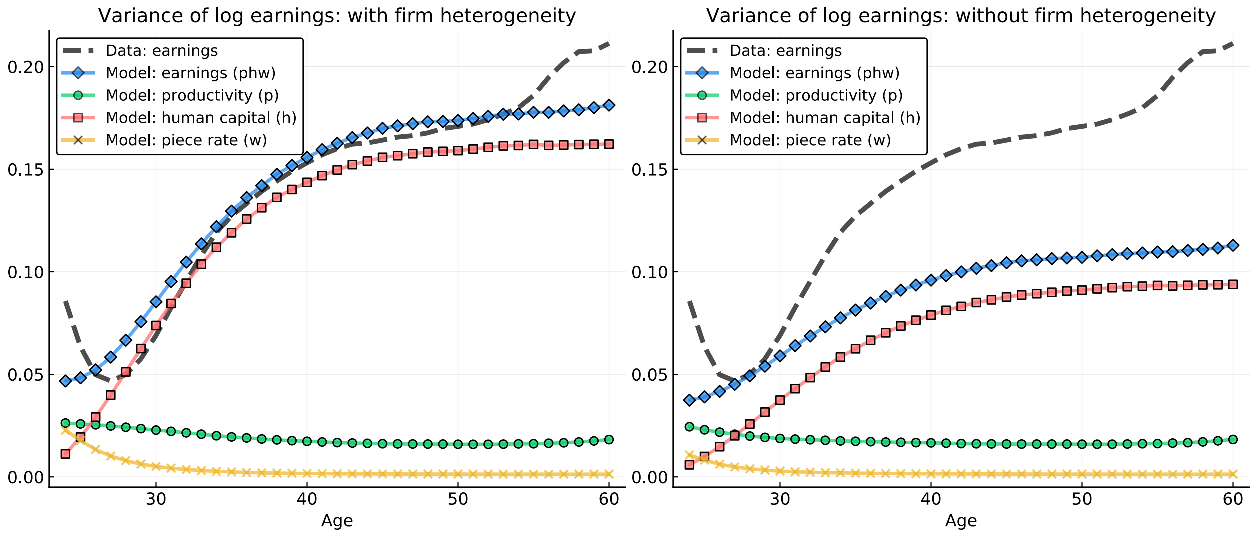

Figure F.2: Life-cycle variance of log earnings and decomposition.
The left panel plots the variance of log earnings in the data and in the model where firms are heterogeneous in q. The right panel plots the corresponding variances in the version of the model without heterogeneity in q – all firms have the median learning environment from the original distribution $F(p, q)$.

the LIAB data or simulated from the model) covariances into a vector $\hat{\mathbf{C}}$ and the corresponding theoretical moments from (18) and (19) into vector $\mathbf{C}(\mathbf{\Theta})$, where $\mathbf{\Theta}$ is the set of parameters for either the RIP of the HIP model. The parameter estimates solve the minimum distance problem:

$$\min_{\mathbf{\Theta}}(\hat{\mathbf{C}} - \mathbf{C}(\mathbf{\Theta}))'\mathbf{W}(\hat{\mathbf{C}} - \mathbf{C}(\mathbf{\Theta}))$$

where $\mathbf{W}$ is a weighting matrix that to takes into account the precision of the empirical moments. It is a diagonal matrix where each element is the inverse of the standard error of the corresponding empirical covariance: $\mathrm{se}_{h,h+n} = \mathrm{sd}(y^i_h, y^i_{h+n}) / \sqrt{N_{h,h+n}}$.

H Details on age-dependent UI schedules

There are multiple combinations of (z_1, z_2) that can achieve the same level of a given outcome. This is because the tax changes to offset the benefits and costs that disproportionately affect workers of different ages. To see this, Figure H.1 shows the contours of different outcomes of the model as a function of (z_1, z_2). The general trend is that workers can be made indifferent between a steep UI schedule with a high intercept and a flatter one with a lower intercept. This multiplicity is what leads me to restrict my search to the policies depicted in Figure 14.

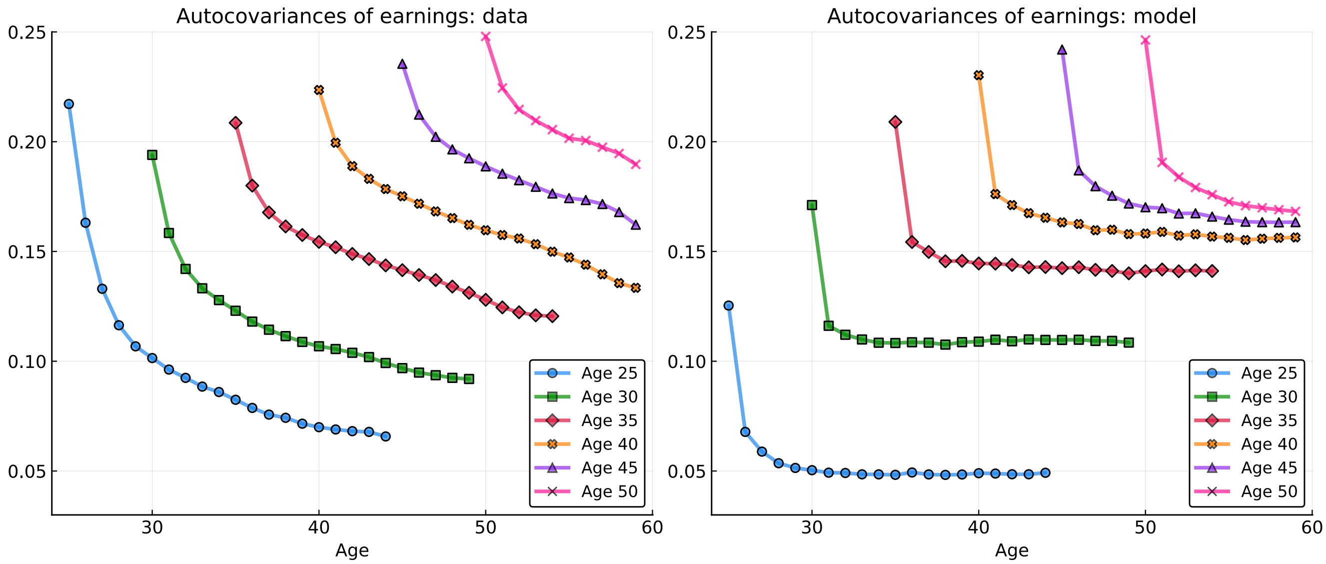

Figure G.1: Comparison of autocovariances of earnings in the data and model.
Each series indicates the covariance of residual earnings at age h and $h + n$, where $h = 25, 30, \ldots, 50$. Residual earnings are calculated as described in Appendix G. The right panel comes from the baseline version of the model with both worker and firm heterogeneity.

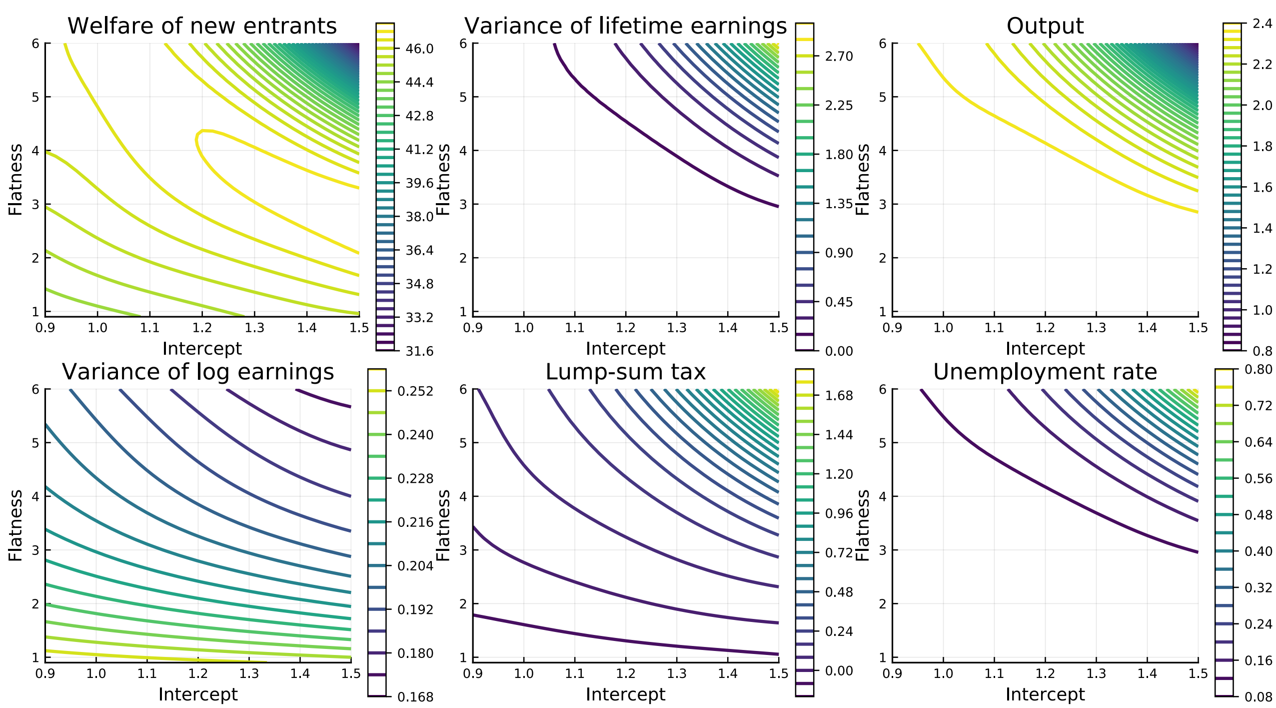

Figure H.1: How the intercept and slope of the benefits function affects different outcomes in the model. “Intercept" refers to z_1 and “flatness" refers to z_2 in $b(t) = \left(b + z_1 t^{-\frac{1}{z_2}}\right) h$, where $b = 0$.