How do Workers Learn? Theory and Evidence on the Roots of Lifecycle Human Capital Accumulation*

Xiao Ma — *Peking University*

Alejandro Nakab — *Universidad Torcuato Di Tella*

Daniela Vidart — *University of Connecticut*

April 2024

Abstract

How do the sources of worker learning change over the lifecycle, and how do these changes affect human capital and wages? We use data from Germany and the US to document that internal learning (learning from coworkers) decreases with workers' experience, whereas external learning (on-the-job training) follows an inverted U-shape pattern. We build a search model featuring multiple learning sources whose benefits evolve as workers accumulate human capital. Quantitative results indicate that internal learning is more important than external learning for early-career human capital, and generates key wage gains later in life due to compensation for learning spillovers among coworkers.

Keywords: On-the-job Learning; Human Capital Accumulation; Lifecycle Wage Growth; Learning Sources

JEL Codes: E24, J24, M53

*Email: xiaoma@phbs.pku.edu.cn, anakab@utdt.edu and daniela.vidart@uconn.edu.

We would like to thank Titan Alon, Edward Glaeser, Victoria Gregory, Zibin Huang, Nir Jaimovich, David Lagakos, Jeremy Lise, Simon Mongey, Marc Muendler, Joseph Mullins, Mariacristina de Nardi, Tommaso Porzio, Valerie Ramey, Victor Rios-Rull, Todd Schoellman, Kjetil Storesletten, Gianluca Violante, and Johannes Wieland for their helpful comments. We are also grateful for the insightful comments of participants in seminars at UCSD, CU Boulder, UH Manoa, Univ. Torcuato di Tella, IADB, NUS, UIUC, PHBS Macro and Finance Workshop, BSE Summer Forum, SED, UMN, ASU, Notre Dame, IZA Macroeconomics of Labor Productivity Workshop, U of Arizona, Harvard, UCSB, U of Houston, U of Toronto, Boston College, Clemson, UCSB-LAEF Labor Markets Conference, CMU, CICM, SITE: Frontiers of Macroeconomic Research, BU, and UC Berkeley.

1 Introduction

Ever since Becker (1962), the economics literature has recognized the importance of on-the-job learning in driving lifecycle wage dynamics (Rubinstein and Weiss (2006)). Work in this literature has identified several key inputs driving on-the-job human capital acquisition, including on-the-job training (Acemoglu (1997); Acemoglu and Pischke (1998), Moen and Rosén (2004), Ma et al. (2020)), learning-by-doing (Bagger et al. (2014), Gregory (2019)), and coworkers (Nix (2017), Akcigit et al. (2018), Herkenhoff et al. (2024), Jarosch et al. (2021)).[1] To date, this literature has focused on studying each input individually, and thus has not yet considered how different inputs interact and jointly shape on-the-job skill acquisition. This paper contrasts with this literature by studying the shared scope of different learning sources to influence lifetime human capital and wage dynamics.

Motivated by the literature and data, we focus on two sources of learning: internal learning (or learning through colleagues), which draws on firms' internal knowledge and thus depends on coworker quality and firm structure; and external learning (or external on-the-job training), which draws on external knowledge and may depend on broader institutional aspects.[2] Distinguishing between these two sources of learning and studying their shared scope to influence lifetime wage dynamics is important for several reasons. First, given that these two sources draw from separate knowledge pools, their relevance varies with worker and firm characteristics. This has important implications for designing policies aimed at enriching worker learning, and understanding the impact of shocks that affect these two sources of learning asymmetrically. An example of a prominent such shock which we examine in this paper, is the post-pandemic rise in remote work which has disrupted face-to-face interactions between coworkers and thus greatly affected internal learning. Second, the two learning sources imply different incentives for human capital acquisition outside of usual productivity gains. Through internal learning, more knowledgeable workers may be compensated for generating valuable learning spillovers to their coworkers. However, through external learning, these knowledgeable workers may have strong incentives to leave their jobs and start their own teaching-focused enterprises. We show in this paper that accounting for these non-own productivity aspects of human capital is important to fully understand the role of human

[1] Other inputs explored in the literature include formal schooling (Ben-Porath (1967)), knowledge hierarchies, (Garicano (2000); Garicano and Rossi-Hansberg (2004, 2006), Caicedo et al. (2019)), materials (Manuelli and Seshadri (2014)), and managerial inputs (Burstein and Monge-Naranjo (2007), Luttmer (2014)).

[2] Our model also includes learning-by-doing since this is another important component of on-the-job human capital accumulation and wage growth. However, since this type of learning is costless and often indistinguishable from work, we do not focus on it as much in the data.

capital acquisition in lifecycle wage growth.

In the empirical section of the paper we first document two novel facts that speak to the importance of our two sources of learning from both firms' and workers' perspectives. First, using firm survey data from Europe, we show that both internal and external sources of learning are widely provided by firms to their workers, and that larger firms offer a greater variety of learning options by providing their workers with more opportunities to engage in both internal and external learning.[3] Second, we use detailed worker qualification data from Germany and the United States to show that both sources of learning are important to workers, and have markedly different lifecycle patterns. In particular, we document that: (1) the prevalence of internal learning decreases with workers' experience; and (2) the prevalence of external learning has an inverted U-shape in workers' experience. These lifecycle patterns are robust to considering alternate definitions of workers' experience, controlling for industry, occupation and demographics, and decomposing the data across several worker- and firm-level characteristics such as education level, gender, and firm size.[4]

We then build a quantitative search model featuring a two-source learning technology to shed light on these findings and assess the importance of internal and external learning for lifecycle wage dynamics. The model features an overlapping generations structure and two sectors: a final-good production sector and a training sector with trainers providing external learning services. The training sector is frictionless, while the production sector is characterized by labor market frictions and firm heterogeneity in productivity. Firms in the production sector meet workers by random search. After matching, workers and firms in the production sector engage in Nash bargaining over the worker's compensation and jointly choose internal and external learning investments to maximize the match value. Human capital follows a ladder structure with a discrete number of steps and determines productivity in each sector. Workers in the production sector allocate their time between work and learning from internal and external sources. Learning from internal (external) sources follows from random meetings with coworkers (trainers), and is contingent on matching with a coworker (trainer) with a higher human capital level than the worker's own. Both forms of learning carry a foregone production cost, but external learning carries an additional cost from the purchase of training services.

[3]We also show evidence in support of this finding in the United States using aggregate data from the Survey of Employer-Provided Training.

[4]We also show evidence in support of these lifecycle findings in other OECD countries using data from the Program for the International Assessment of Adult Competencies (PIAAC).

We calibrate the model to the US economy and find that the stationary equilibrium replicates the lifecycle learning patterns we found empirically. This follows from the fact that the incentives to engage in each source of skill acquisition evolve throughout the workers' lifecycle as they accumulate human capital. In particular, changes in the relative position of the worker in the human capital distribution of the firm mediate the supply of coworkers and trainers that can be learned from and lead to distinct lifecycle patterns of learning. Consistent with our empirical findings, a young worker in the model disproportionately relies on coworkers to learn since internal learning is relatively cheap, and the proportion of coworkers with higher human capital than that of the the young worker is large. As the worker accumulates human capital, the proportion of coworkers with human capital higher than the worker's own declines, inducing a switch to external learning since trainers tend to have higher average human capital levels than production workers. As workers continue to age and human capital continues to increase, the opportunity cost of learning rises while the benefit from learning decreases as the remaining working life shortens, leading external learning to decline.[5]

The calibrated model also highlights the importance of firms' learning environments for human capital formation. At all levels of human capital, workers in more productive firms spend more time on both internal and external learning and thus climb the human capital ladder faster. This finding matches our empirical evidence showing that workers in larger firms spend significantly more hours on both sources of learning, and is also consistent with evidence found by Engbom (2017), Arellano-Bover (2020), and Arellano-Bover and Saltiel (2023) showing that workers in more productive firms exhibit faster rates of skill acquisition. In our model, more productive firms invest more in both types of learning since they exhibit both larger returns to skill acquisition (due to higher efficiency of learning and supermodularity of the production function) and a better pool of coworkers to learn from (due to positive assortative matching between firms and workers).[6]

[5]We also present empirical evidence matching key predictions from this lifecycle theory. First, we provide evidence showing that trainers have higher average human capital levels than production workers and are thus better equipped to teach mid-career workers who have exhausted learning opportunities within their firms. Second, we show that the portion of individuals who learn-by-doing rises with human capital. This is consistent with the idea that as workers age, they put more hours toward working rather than internal or external learning, therefore increasing their chances of learning-by-doing. Third, we show that internal learners exhibit lower levels of task complexity than external learners, and are thus more easily trained by coworkers, who have lower human capital levels than external trainers.

[6]This pattern of positive assortative matching emerges in our framework due to the more favorable learning environments prevalent in more productive firms.

To assess the importance of internal and external learning for the formation of human capital, wage growth, and wage dispersion, we perform counterfactual analyses in which we subsequently shut down each of these two sources of skill acquisition and examine how the stationary equilibrium changes. We have two main findings. First, we find that internal and external learning contribute equally to aggregate human capital: without either external or internal learning, workers' average human capital decreases by 14%. However, when we examine the trajectory of human capital throughout the lifecycle in these two counterfactual scenarios we find that internal learning is relatively more critical to human capital formation during youth, whereas external learning is relatively more critical at older ages.

Second, we find that accounting for the non-own productivity dimension of human capital, which captures gains from human capital acquisition other than the increase in the worker's own production output, is key to understanding lifetime wage dynamics. In particular, we find that changes in learning costs and compensation stemming from learning spillovers among coworkers account for 36% of the wage gains from human capital over 25 years of experience. Moreover, although the lifecycle increase in the dispersion of workers' own productivity levels is largely driven by external learning,[7] the lifecycle increase in wage dispersion is mainly driven by internal learning, as the compensation stemming from coworker learning spillovers explains much of the wage gains for more-experienced workers in the model. These results suggest that accounting for internal learning is critical for understanding wage dynamics, and thus that a model that ignores internal learning may mistakenly attribute wage growth to increases in workers' own productivity, rather than increases in compensated non-own productivity, such as coworker learning spillovers.[8]

The importance of internal learning for wage dynamics is further confirmed when we use the model to evaluate the effects of remote work. To do this, we recalibrate our model using data from recent papers documenting the rise of remote work after the Covid-19 pandemic and the impact of the disruption of face-to-face contact on the opportunities for internal learning. Specifically, we rely on the findings of Barrero et al. (2021) who document a rise in the proportion of remote work days from 5% to 20% before and after the pandemic, together

[7]Without external learning, the dispersion in workers' productivities remains low throughout the lifecycle as workers learn from and catch up fast to colleagues.

[8]We also present empirical evidence supporting the importance of coworker instruction in many occupations. To do this, we use data from the US Department of Labor's O*NET project which aims to characterize the tasks pertaining to each occupation, along with the mix of knowledge, skills, and abilities required to perform these tasks. We find that tutoring coworkers is an important part of the job in many non-teaching occupations and that these occupations are mainly performed by high-skill workers.

with the findings of Emanuel et al. (2023) showing a 15% reduction in the feedback provided by adjacent coworkers following the shift to remote work. We find that the surge in remote work following the pandemic leads to a 0.44% reduction in workers' average human capital, which constitutes 4.73% of the human capital accumulated by internal learning.

In addition, our results suggest a 2.86% decrease in workers' average wage growth after 25 years of experience due to the post-pandemic rise in remote work. At younger ages (0–5 years of experience), the decline in own productivity following the lack of senior mentoring is a key factor driving the reduction in wage growth, accounting for 39% of the decline. At older ages (5–25 years of experience), on the other hand, the decrease in spillover compensation following the lack of junior colleagues to teach becomes more important, accounting for 37% of the decline in wage growth. This implies that remote work not only affects the wages of younger workers through learning, but also affects the wages of more senior workers through compensation tied to their ability to teach and mentor. In addition, our results show that remote work can have important long-lasting spillover effects as the young workers who learn less as a consequence of remote work will be potentially less able to teach younger workers in the future. However, external learning helps alleviate these negative impacts as young workers can shift from internal to external learning.

The paper is organized as follows. In Section 2 we present a literature review. In Section 3 we describe the data and document the main empirical findings. In Section 4, we present the quantitative model, calibration, and properties of equilibrium. In Section 5 we report our counterfactual results that assess the importance of internal and external learning for human capital and wage dynamics. We evaluate the impact of remote work in Section 6. In Section 7 we discuss the robustness of our quantitative results to several model extensions and alternative parameterizations. We conclude in Section 8.

2 Related literature

Our paper is closely related to the literature exploring the importance of on-the-job skill acquisition for human capital and wage growth. Our theory provides a unified structure which jointly considers internal and external sources of learning, and thus relates to different strands within this literature. First, our paper relates to the literature exploring the role of peers in knowledge diffusion within coworker and production teams (Garicano (2000), Garicano and Rossi-Hansberg (2004, 2006), Azoulay et al. (2010), Luttmer (2014), Nix (2017), Akcigit et al. (2018), Herkenhoff et al. (2024), Jarosch et al. (2021), Caicedo et al.

(2019), Sohail (2021), Emanuel et al. (2023), Wallskog (2023)) and within the population at large (Glaeser (1999), Jovanovic (2014), Lucas and Moll (2014), Perla and Tonetti (2014), de la Croix et al. (2016), Benhabib et al. (2021)). Our paper is particularly related to Akcigit et al. (2018), who build a model where inventors can learn both through interacting with others, and from an external exogenous source. Similar to them, our model features learning through both external sources and coworkers. In contrast with their paper, however, our model poses the boundary of the firm as the distinguishing factor between internal and external learning, and endogenizes both of these choices. In addition, our data focuses on general workers rather than inventor teams, and our theory highlights that changes in the worker's relative position in the human capital distribution of the firm affect her returns of learning from coworkers, and can thus explain the observed lifecycle patterns in human capital acquisition. By considering external on-the-job training, this paper also relates to the literature on general training investments first proposed by Becker (1964), and later developed by others (Acemoglu (1997), Acemoglu and Pischke (1998), Acemoglu and Pischke (1999), Autor (2001), Moen and Rosén (2004)).

Our paper also relates to the literature exploring the interaction between learning and lifecycle dynamics. First, our paper relates to studies that examine the effects of work-related human capital acquisition on earnings. Much of this literature has focused on disentangling the role of learning, search dynamics, and shocks on earnings growth (Bunzel et al. (1999), Rubinstein and Weiss (2006), Barlevy (2008), Yamaguchi (2010), Burdett et al. (2011), Huggett et al. (2011), Bowlus and Liu (2013), Bagger et al. (2014), Gregory (2019), Karahan et al. (2022)). This contrasts with our goal, which is to disentangle the contributions of different sources of learning to human capital and earnings growth, and our findings, which highlight the importance of the non-own productivity aspects of human capital in workers' compensation dynamics. Second, since we explicitly consider the time and monetary costs of internal and external learning, our paper relates to the seminal literature highlighting the tradeoff between learning and work (Ben-Porath (1967), Heckman (1976), Rosen (1976)). Thus, our paper contrasts with several recent papers that examine the role of on-the-job human capital accumulation in knowledge diffusion or earnings growth (Lucas (2009), Bagger et al. (2014), Gregory (2019)), and which model on-the-job human capital accumulation via costless learning-by-doing and do not consider multiple sources of learning.

By examining how the provision of different sources of learning shapes firms' learning environments, our paper also relates to the literature that considers the role of firms and firm-level characteristics in shaping workers' lifecycle dynamics (Gregory (2019), Arellano-

Bover (2020), Engbom (2021), Friedrich et al. (2021), Jarosch (2021), Engbom et al. (2022), Arellano-Bover and Saltiel (2023)). This literature shows that there is substantial heterogeneity in firms' promotion of human capital accumulation, and that such heterogeneity is an important determinant of lifecycle earnings dynamics. However, the drivers of firms' learning environments are still poorly understood. In our paper, we contribute in this direction by providing empirical and theoretical evidence of a concrete driver of success in firms' learning environments: the provision of external and internal learning opportunities.

Finally, our paper also relates to the vast labor literature examining the impacts of on-the-job learning opportunities on workers' earnings (see Heckman et al. (1999), Kluve (2010), McKenzie (2017), Card et al. (2018), and What Works - Centre for Local Economic Growth (2016) for a review), and particularly the literature showing that the productivity and earnings gains of on-the-job training can vary greatly depending on the type of learning opportunity provided (Fitzenberger and Völter (2007), What Works - Centre for Local Economic Growth (2016)). One key distinction highlighted in these studies arises from comparing in-firm to classroom-based on-the-job learning opportunities, which broadly match our internal and external categories of learning, respectively. Our paper contributes to this literature by suggesting that the relevance of these two types of learning opportunities, and thus the empirical evidence found in these studies, will crucially depend on the characteristics of the workers studied such as experience and human capital levels.

3 Data and empirical findings

In this section, we empirically explore the importance and lifecycle patterns of internal and external learning using both firm- and worker-level data. We first describe our data sources.

3.1 Data

In Table 3.1 we summarize the data sources used to document each of our facts. This table also provides links to the appendices containing more detailed information about each data source.

Table 3.1: Data sources used in empirical analysis

Data source	Setting	Description	Methodology & size	More info
Data for Fact 1 (provision of learning opportunities by firms)				
EU Continuing Vocational Training Survey (EU-CVT)	European Union & Norway 2005, 2010, 2015	Firm-level survey focusing on firms' investments in continuing vocational training (CVT) for staff	Repeated cross-section ∼ 95,000 firms per wave	Appendix A.1
US Survey of Employer-Provided Training (US-SEPT)	United States 1995	Firm- and worker-level survey focusing on training investments in establishments with 50+ workers	Single cross-section ∼ 1,000 establishiments & 1,000 workers	Appendix A.2
Data for Fact 2 (lifecycle patterns of learning)				
Data source	**Setting**	**Description**	**Methodology & size**	**More info**
German BIBB/BAuA Worker Qualification Survey	Germany 1979, 1985, 1992, 1999, 2006, 2012, 2018	Worker-level survey focusing on on-the-job skill acquisition and occupational skill requirements	Repeated cross-section ∼ 25,000 workers per wave, excludes apprentices	Appendix A.3
US Adult Training & Education Module in National Household Education Survey (NHES)	United States 2016	Worker-level survey focusing on formal education and on-the-job skill acquisition	Single cross-section ∼ 48,000 adults	Appendix A.4
OECD Program for the International Assessment of Adult Competencies (PIAAC)	40 OECD countries 2011–2017 (each country surveyed in a different year)	Worker-level survey focusing on learning investments and skills	Single cross-section per country ∼ 230,000 adults total	Appendix A.5

3.2 Fact 1: Larger firms provide more learning options

First, using the EU-CVT data, we document the importance of internal and external learning for firms, and how the provision of each of these forms of learning varies with firm size. We distinguish between internal and external learning opportunities by relying on information on the location and instructor affiliation of CVT activities. CVT encompasses educational or training activities that are planned in advance, organized, or supported with the specific goal of learning. The survey explicitly distinguishes between "internal CVT courses" and "external CVT courses" by separating courses, seminars or activities that take place inside firms and employ internal trainers from those that occur outside firms or employ external trainers. We consider "internal CVT courses" as reflecting internal learning and "external CVT courses" as reflecting external learning. In addition, the survey also measures "other types of CVT activities," which include four types of activities: participation in conferences and lectures, guided on-the-job training, job rotation, and learning or quality circles. Based on the definitions of these activities, we consider participation in conferences and lectures as external learning, and the remaining types as internal learning. Table A.1 shows that a large portion of firms in each country offer "internal CVT courses" and "external CVT courses," along with "other types of CVT activities," indicating that both internal and external learning sources are part of the learning portfolio offered to workers.[9]

Table 3.2: Share of firms providing internal and external learning activities by firm size

Small firms, 5–19 wk.		External learning 0	External learning 1	**Medium firms, 20–99 wk.**		External learning 0	External learning 1	**Large firms, 100+ wk.**		External learning 0	External learning 1
Internal learning	0	0.41	0.16	Internal learning	0	0.28	0.15	Internal learning	0	0.13	0.09
	1	0.11	0.32		1	0.11	0.46		1	0.07	0.71

Notes: These tables present the proportion of firms of different sizes reporting having employees participating in internal and external learning activities in the EU-CVT data (CVT3, CVT4 and CVT5 surveys). The sample encompasses firms with 5 or more workers, since smaller firms encompass a very small portion of the sample. The results are weighted using the observational weights provided in the surveys.

In Table 3.2 we examine how the proportion of firms that invest in external and internal learning varies with firm size. We find that 32%, 46%, and 71% of firms with 5–19, 20–99, and 100+ workers, respectively, offer both external and internal learning to their workers, indicating that the share of large firms offering both sources of learning is much larger than that of smaller firms.[10] Interestingly, there are much smaller differences in the proportion of

[9] This is confirmed in Table B.1, which aggregates the data across all countries and shows that 41% of all firms surveyed offer both external and internal learning opportunities to their employees, while 26% offer one of these forms of learning.

[10] The sample encompasses firms with 5 or more workers, since smaller firms encompass a very small portion of the sample.

firms offering a single source of learning across firms of different size categories, suggesting that large firms favor learning environments with both sources of learning (and thus a great variety of learning), rather than ones with a single source of learning.

This finding is consistent with evidence found by Engbom (2017), Arellano-Bover (2020), and Arellano-Bover and Saltiel (2023) showing that workers in more productive firms exhibit faster rates of skill acquisition, and with evidence found by Gregory (2019) showing that having different forms of training available is important for firms' learning environments.

We consider the robustness of these patterns in Appendix B. In Table B.2 we show that the positive correlation between firm size and learning opportunities is robust to controlling for industry, socioeconomic, and country-year fixed effects. In Table B.3, we show that in larger firms, workers on average spend more hours engaging in both sources of learning.[11]

In Table B.5 we further show the robustness of these patterns using the US-SEPT data from the United States. The results suggest that the share of employees exposed to informal and formal training (which roughly though not perfectly map onto internal and external learning),[12] along with the number of hours each employee spends on each of these training activities, generally increase with firm size. Nevertheless, these results should be interpreted with caution since the results stem from aggregate statistics.

3.3 Fact 2: Workers' learning sources change over the lifecycle

Using the German BIBB/BAuA and US NHES data, we now document the importance of the two different sources of learning for workers, and how this importance changes over the lifecycle. We rely on variables about both human capital accumulation and potential work experience to conduct our analysis. First, we construct a measure of internal learning that captures workers who have recently engaged in learning from colleagues or superiors. In the German data, internal learning indicates whether an individual has acquired the skills or knowledge necessary to complete the tasks in their current job through colleagues or superiors. In the US data, internal learning captures workers who reported receiving instruction or training from a coworker or supervisor in their last work-experience program, which is defined

[11] This positive correlation between learning hours and firm size is robust to controlling for industry, socioeconomic, and country-year fixed effects (see Table B.4), and is also confirmed when we show the histograms of the share of working hours spent on each type of learning source for firms of each size category (see Figure B.1).

[12] In particular, formal training is defined as training that is planned in advance and has a structured format and defined curriculum (such as classes or seminars), while informal training is defined as training that is unstructured and unplanned (such as having a coworker or supervisor teach a new skill).

as a job with learning attributes.[13] Second, we construct a measure of external learning that captures workers who have recently engaged in learning from companies or instructors outside the current firm. In the German data, external learning indicates whether an individual received external on-the-job training in the previous 2–5 years, or acquired the skills/knowledge necessary to complete the tasks in their current job through external training. In the US data, external learning captures workers who reported taking classes or training from a company, association, union, or private instructor in their last work-experience program, or ever earned a training certificate from an employment-related training program.[14] In Appendix A.3 and Appendix A.4 we provide further details on the questions and answers used to construct each variable in the German and US surveys, respectively.[15]

We construct potential years of work experience using age and educational level. Specifically, we construct $Potential\ experience = Age - Years\ of\ \ schooling - 6$ for both Germany and the US. We limit our sample to individuals who are currently employed and have between 1 and 45 years of potential experience, given that the number of observations outside this range is very small. In Table A.2 we present some key summary statistics of our samples. This table shows that in both Germany and the US the proportions of individuals reporting each of the two sources of learning are sizeable, suggesting that both of these sources are important to workers. In Germany, 31% and 68% of workers report internal and external learning, respectively, while in the US, these percentages are 23% and 44%, respectively.

We now document the prevalence of each learning source over the lifecycle. In Figure 3.1 we plot how the prevalence of workers reporting engaging in internal and external learning changes with workers' potential experience in Germany and the US. We find that in both

[13] In particular, a work-experience program is defined as a job with learning attributes such as an internship, co-op, practicum, clerkship, externship, residency, clinical experience, apprenticeship, or other learning components. About 25% of the surveyed sample in the US reported having been part of such a program. In Figure C.8 we show that our results are robust to limiting the sample only to individuals reporting participating in a work-experience program, and to decomposing across learning components that involve a "work-experience" program and those that do not.

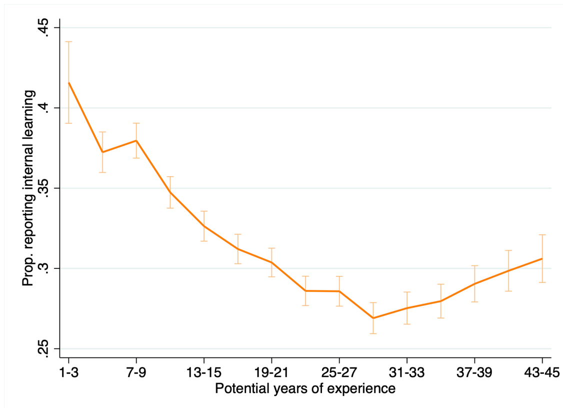
[14] In principle, formal schooling also fits the characterization of external learning as it draws from knowledge outside the firm. However, since less than 10% of adult education corresponds to schooling in the EU, while over 90% corresponds to on-the-job learning (Ma et al. (2020)), we abstract from formal schooling.

[15] It is important to note that both internal and external learning generally capture flows and not stocks of learning investments, since they refer to skill acquisition in the current job or work-experience program (which changes every few years as workers climb the career ladder), or training incurred in the last few years. The exception to this is the part of the external learning definition in the US which involves ever earning a training certificate. This variable is still informative of workers' flow of external learning, since a steeper increase in this variable along a specific portion of the lifecycle denotes larger positive flows at certain ages. In panel (c) of Figure C.8 we show that the inverted U-shape pattern for external learning holds when we limit its definition to exclude this part, though the result is noisier.

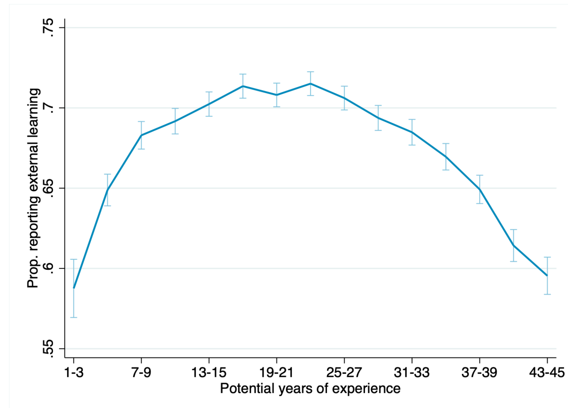
Figure 3.1: Prevalence of internal and external learning throughout workers' lifecycles

Germany

(a) Internal learning

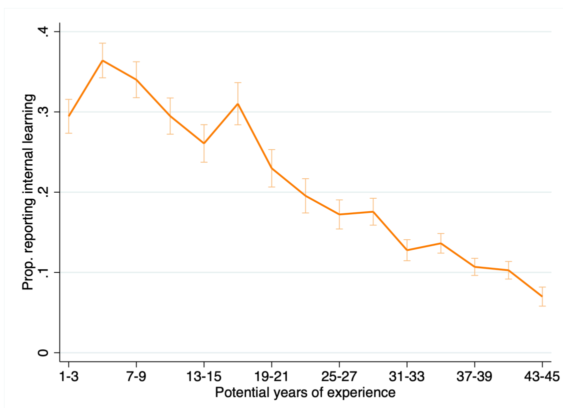


(b) External learning

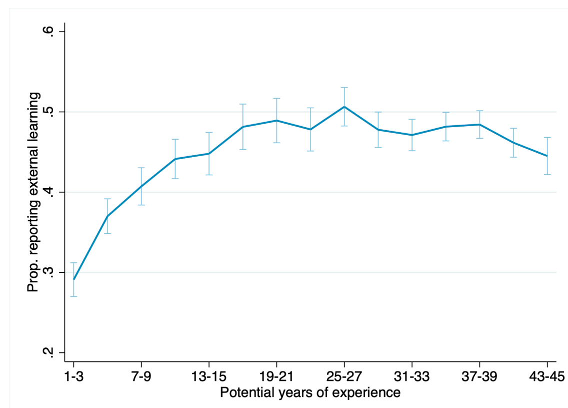


United States

(c) Internal learning



(d) External learning



Notes: These figures plot the proportion of workers reporting engaging in internal and external learning across different potential experience bins in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have between 1 and 45 years of potential experience in both settings. The results are weighted using the observational weights provided in the surveys. 95% confidence intervals are plotted.

countries, the prevalence of internal learning decreases with workers' potential experience, while the prevalence of external learning has an inverted U-shape in workers' potential experience. These patterns are robust to decomposing the data along many dimensions[16] and to controlling for several demographic variables and firm characteristics, along with occupation

[16]These include using age as an alternate measure of experience (Figure C.1), and decomposing the data by one-year experience bins (Figure C.2), gender (Figure C.3), educational level (Figure C.4), survey wave in Germany (Figure C.5), cohort in Germany (Figure C.6), and firm size in Germany (Figure C.7).

and industry fixed effects (see Table C.1).[17] The results are also robust to considering workers' tenure rather than potential experience (see Figure C.9). This suggests that the patterns we document are not solely a consequence of the aging process, but that work experience and human capital matter. This is further confirmed when we formally explore the correlations between current tenure and the sources of learning in Table C.2, given that the patterns of interest hold even when we include age fixed effects.

In Figure C.10, we further consider the robustness of these lifecycle patterns in other OECD countries using data from PIAAC.[18] These binned scatterplots use pooled data from all countries considered and present the proportion of workers reporting internal and external learning, respectively, across different experience bins after controlling for country fixed effects.[19] The results suggest similar lifecycle patterns to the ones described above. In Table C.4 we show that these results are robust to controlling for several demographic variables and firm characteristics, along with occupation and industry fixed effects, though the effects become noisier since the number of individuals interviewed in each country is relatively small.

The lifecycle patterns found in this section are consistent with the evidence found by several papers suggesting that younger workers are more sensitive to peer learning than older workers. For example, Nix (2017) finds that coworker learning spillovers are larger for younger workers, with no impact for workers over 40. Jarosch et al. (2021) find that the positive effects of peers' wages on future wages are substantially stronger for younger workers. Using data from software engineers, Emanuel et al. (2023) find that sitting proximity increases how much junior engineers learn from senior colleagues. While focusing on academic networks, Azoulay et al. (2010) find that academics who had collaborators who died unexpectedly experience a decline in their quality-adjusted publication rates, and that this decline is larger for younger academics. When comparing learning opportunities between urban and rural

[17] We also examine the correlation between internal and external learning in both settings, and find a negative correlation in Germany, and a positive correlation in the US. We discuss these results in Appendix C.1.2.

[18] The countries considered do not include Germany and the US due to data restrictions, but do include 27 other countries. For a full list of the countries considered in this analysis please see Appendix A.5.

[19] External learning is measured in this data through participation in seminars, workshops, or other courses led by outsiders in the last 12 months. Internal learning is measured through participation in courses led by coworkers or supervisors in the last 12 months, or reporting learning skills from coworkers more than once a month in the current job. Thus, these variables capture flows of learning investments, since they refer to skill acquisition in the last 12 months, or the current position. In addition, the internal learning variable captures both formal instruction from coworker-led courses, along with informal instruction arising from coworker interaction. Please note that years of experience are measured directly in this data through a variable capturing realized years of paid work up to date of measurement. To increase comparability with the previous results, we also include the analogous plots with potential years of experience in Figure C.11 showing the same patterns. Summary statistics for this data can be found in Table A.3.

environments, Glaeser (1999) finds that young workers are particularly attracted to cities due to the increased probability of meeting skilled workers to learn from.

4 A model of human capital with two sources of learning

We now develop a model featuring a two-source learning technology to evaluate the importance of internal and external learning for lifecycle wage and human capital dynamics. The model features two primary mechanisms connecting to the firm-level and lifecycle learning patterns presented above. The first mechanism is that the returns from workers' human capital investments change across firms of different productivity levels, leading to differences in learning investments and the speed of skill acquisition across firms. The second mechanism is that as workers age and acquire human capital, their potential to learn from coworkers decreases since the number of colleagues with higher human capital than their own narrows. Consequently, workers gradually transition from internal to external learning sources, as the latter is more costly but involves highly skilled training experts. These mechanisms have important implications for lifecycle wage dynamics, especially since highly skilled workers are compensated for teaching their colleagues, and this compensation may vary across firms. Moreover, this structure has implications for the effects of shocks that asymmetrically impact both forms of learning. In Section 6 we use the model to assess the impact of one such shock, remote work, which disrupts face-to-face interactions between coworkers and thus affects internal learning.

4.1 Model setup

The model features an overlapping generations structure where each worker has a finite lifetime. Workers are endowed with one unit of time in each period and accumulate human capital through internal and external learning, as well as learning-by-doing. The economy consists of two sectors: a final-good production sector, and a training (or external learning) sector through which trainers provide external training services to production workers. The training sector is frictionless, while the production sector is characterized by heterogeneous firms and labor market frictions à la Cahuc et al. (2006).[20] In the production sector, firms

[20]From a theoretical perspective, labor market frictions are essential for both firms and workers to benefit from on-the-job training targeting general skills (Acemoglu (1997); Moen and Rosén (2004); Acemoglu and Pischke (1999)). Without frictions, human capital gains are immediately priced into wages, eliminating incentives for firms to provide them. This contradicts the data, which shows that 80% of all training investments are at least partially sponsored by firms (Ma et al. (2020)).

post vacancies to attract unemployed individuals and meet their matches through random search. After matching, workers and firms engage in Nash bargaining over the worker's compensation and choose internal and external learning investments to maximize the match value. Learning from each source stems from random meetings with coworkers and trainers, respectively, and depends on matching with a coworker or trainer who has a higher human capital level than the worker's own. Both forms of learning carry a foregone production cost, but external learning carries an additional cost from the purchase of training services. In what follows, we focus on the model's stationary equilibrium, in which firm-level distributions of workers' age and human capital levels remain constant. We index workers by i and firms by their productivity level z.

4.1.1 Workers

The model economy consists of overlapping generations of workers who participate in the labor market at ages $a = 1, 2, \ldots, J$. Workers of age J retire and are replaced by a cohort of incoming workers of age 1. Each cohort's size is normalized to 1. Throughout their lives, workers accumulate human capital h, which determines their labor productivity. Workers can be employed in either the production or training sector. In the production sector, workers allocate their time to working, internal learning, and external learning. Conversely, in the training sector, they spend all their time working. Workers derive linear utility from consuming a homogeneous final good and discount the future at a rate of ρ.

4.1.2 Human capital accumulation

We assume that human capital evolves throughout workers' lives in a ladder-like fashion:

$$h \in \{h_1, h_2, ..., h_M\}, \text{ where } h_M > ... > h_2 > h_1, \text{ and } \log(h_{m+1}) - \log(h_m) = \gamma_h \ \forall m.$$

Newborn workers (age 1) are endowed with the lowest human capital level, h_1. Workers ascend this ladder through various forms of learning. All learning methods contribute to general human capital, which can be transferred if workers change jobs.[21]

[21] We focus on general human capital for tractability, and also because, as documented by Altonji and Shakotko (1987), Lazear (2009), and Kambourov and Manovskii (2009), the truly firm-specific components of human capital are much less important for wage growth than the general component. Additionally, we concentrate on the differences between learning sources arising from the pool of knowledge each of them taps into, rather than differences in the "transferability" of learning for two reasons. First, the organizational literature on workplace learning (see Manuti et al. (2015) for a review) suggests that both internal and external learning (often labeled informal and formal learning) are vital dimensions of workplace learning, and both can contribute to forming new and transferable skills and competences for workers. Second,

The probability of climbing the human capital ladder depends on internal and external learning investments, as well as learning-by-doing. We represent worker i's human capital level as h_{m_i}, where m_i indexes the worker's current position on the human capital ladder. We assume that for worker i in firm z, the probability of climbing up the human capital ladder is given by

$$s(h_{m_i}, \mathbb{H}(z)) = \min\left(s^E(h_{m_i}, \mathbb{H}(z))l^\gamma + \epsilon, 1\right). \tag{1}$$

$s^E(h_{m_i}, \mathbb{H}(z))$ captures the increase in human capital per unit of time spent on learning, where state variable $\mathbb{H}(z) = \{h_{m_i}\}_{i\in\mathbb{I}(z)}$ captures the human capital distribution of firm z's workforce and thus its internal learning environment, and $\mathbb{I}(z)$ describes the workforce of firm z. l represents the total time dedicated to learning, and the parameter $0 < \gamma < 1$ captures its diminishing marginal benefits. $\epsilon > 0$ is the exogenous probability of climbing the human capital ladder, resembling learning-by-doing.[22]

The human capital increment, $s^E(h_{m_i}, \mathbb{H}(z))$, depends on the time spent on each learning source and the probability of matching with a colleague or trainer who can effectively teach the worker. Specifically, we assume that workers can only successfully learn from coworkers and trainers with a higher human capital level than their own. This aligns with the findings of Herkenhoff et al. (2024), who show that workers learn from more knowledgeable coworkers rather than less knowledgeable ones, and with the findings of Jarosch et al. (2021), who document that peers higher up in a team's wage distribution have a greater impact on workers' future wage outcomes than peers below. Thus, we assume that the functional form of the human capital increment is given by

$$s^E(h_{m_i}, \mathbb{H}(z)) = \left[(A(z)p(h_{m_i}, \mathbb{H}(z))g)^{\frac{\sigma-1}{\sigma}} + (A_e(z)p_e(h_{m_i})(1-g))^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}. \tag{2}$$

The probability that worker i in firm z matches with a colleague possessing higher human capital is denoted by $p(h_{m_i}, \mathbb{H}(z))$. This probability is determined by the share of coworkers above i in the firm's human capital distribution, and can be expressed as $p(h_{m_i}, \mathbb{H}(z)) = 1 - F(h_{m_i}; \mathbb{H}(z))$, where $F(\cdot; \mathbb{H}(z))$ represents the cumulative distribution of human capital of the workforce in firm z. Similarly, the probability that worker i matches with a trainer possessing higher human capital is denoted as $p_e(h_{m_i})$ and can be calculated as $p_e(h_{m_i}) = 1 - F_t(h_{m_i})$, with $F_t(\cdot)$ representing the cumulative distribution of trainers' human capital (weighted by the number of training units each trainer produces) within the training sector.

although other differences between the two forms of learning may exist, our theory generates several testable predictions that match key features of the data (see Appendix G).

[22]This probability also ensures the presence of high-skill workers in the equilibrium.

The proportions of learning time spent on internal and external learning are denoted as g and $(1-g)$ respectively, with $\sigma > 0$ representing the elasticity of substitution between the two learning modes. Additionally, $A(z)$ and $A_e(z)$ represent the efficiencies of internal and external learning respectively within firm z. By modeling $A(z)$ and $A_e(z)$ as functions of firm productivity z, we account for the possibility that learning investments lead to faster human capital accumulation in more productive firms, which could be due to factors such as scale effects of learning or differences in the availability of complementary physical capital investments.[23] This model setting aids in matching the empirical differences in learning time by firm size.

When engaging in internal or external learning, the worker has to forego production. Moreover, in the case of external learning, the worker must also pay a price q per unit of time for purchasing training services. This price is only paid if the worker meets a trainer with higher human capital than her own. Furthermore, we assume that workers incur no cost when colleagues learn from them, and can have multiple colleagues learning from them simultaneously. This assumption is based on the qualitative observation that workers who are learning are often included in projects with senior colleagues to observe and learn.[24]

Finally, we also allow the human capital gains from on-the-job learning to depreciate following the findings of Mincer (1989) and Blundell et al. (2021) by considering that a worker with human capital level h_{m_i} has a probability $\delta_h(h_{m_i} - h_1)/(h_{m_i} - h_{m_i-1})$ of descending the human capital ladder by one step after finishing learning in each period. δ_h thus captures the depreciation rate of returns from on-the-job learning.[25]

4.1.3 Production sector

The final-good production sector features frictional labor markets and a unit measure of heterogeneous firms that produce a homogeneous good. Firms post vacancies, meet workers through random search, and differ in their productivity level z which is Pareto-distributed: $z \sim \Phi(z)$. We follow Acemoglu and Hawkins (2014) and consider convex costs of posting v vacancies $c_v v^{1+\gamma_v}/(1+\gamma_v)$, which are denoted in units of the final good. We assume $\gamma_v > 0$ to ensure that more productive firms are larger and that firms of varying sizes coexist. $v(z)$

[23] The idea that workers may learn more in certain jobs dates back to Rosen (1972) and has been recently explored in Gregory (2019), Monge-Naranjo (2019), and Engbom (2021).

[24] In Appendix F.2 we allow mentoring to incur a time cost for internal trainers and find that the properties of equilibrium and quantitative results are very similar to those of the baseline model.

[25] This model setting also ensures that human capital cannot decrease below h_1.

represents the optimal number of vacancies in firm z. The total number of vacancies is then given by $V = \int v(z) \mathrm{d}\Phi(z)$.

At the start of each period, existing labor matches are terminated at an exogenous rate κ. These separated workers enter the unemployment pool along with newly-born workers. Before job search occurs, unemployed individuals can choose to search for a job in the production sector or switch to the training sector. Meanwhile, trainers can decide to stay in the training sector or switch back to the production sector and search for a job alongside other unemployed workers. We denote the number of unemployed workers at the point of job search as U.[26] Unemployed workers randomly meet job vacancies at the meeting rate λ_U, which is endogenously determined by $\lambda_U = \chi(V/U)$. The function $\chi(\cdot)$ governs the matching process. When an unemployed worker encounters a job vacancy, she accepts the job with certainty, assuming that unemployment is equivalent to employment at the least productive firm.[27] Unemployment continues if the unemployed person does not encounter a vacancy.

When a job match is formed, worker i's production in firm z follows from the product of firm productivity and worker human capital: $y = zh_{m_i}$. Consequently, the production function is supermodular: a firm with higher productivity generates more revenue per unit of labor, and human capital and firm productivity serve as complements, as illustrated in Acemoglu and Pischke (1998) and Bagger et al. (2014). The worker and the firm collaboratively choose learning investments to maximize the match value. They also engage in Nash bargaining to determine the distribution of the match surplus and the corresponding wage rate. These choices and results are characterized below.

4.1.4 Training sector

We assume that the training sector is frictionless, and thus that unemployed workers can freely transition to this sector and utilize the training technology without any restrictions. We also assume that the amount of training services provided by a trainer is proportional to her human capital level, as high-skill individuals can typically teach multiple students simultaneously. Trainers randomly meet workers who engage in external learning. After observing

[26] Our baseline model does not incorporate on-the-job search in order to focus on the role of human capital in driving wage growth and due to the added complications of this extension. In Appendix F.1 we incorporate on-the-job search into our baseline model and find that the properties of equilibrium and quantitative results are very similar to those of the baseline model.

[27] Following Bagger et al. (2014), this assumption avoids the complication of heterogeneous reservation wages for workers with different human capital levels and ages.

the trainer's human capital, external learners (or trainees) decide whether to participate in external training or not. Trainers can only effectively train external learners with human capital lower than the trainers' own.[28] Thus, the expected earnings of trainer i with human capital h_{m_i} are $h_{m_i} q F_e(h_{m_i-1})$, where $F_e(\cdot)$ is the cumulative distribution of human capital levels (weighted by external learning time) for external learners. In equilibrium, the price of training services q clears the market for training services such that the total demand for external training services (external training time aggregated across all production workers) equals the total supply (training units provided by all external trainers).

4.2 Solving the model

We now characterize learning, wages, and vacancy postings in the production sector.

4.2.1 Solving for learning investments

Firms and workers in the production sector jointly select the total learning time and the proportion of this time allocated to internal and external learning, respectively, to maximize the match value. For a worker i of age a_i and with human capital h_{m_i} in firm z, the match value can be expressed as

$$
\begin{aligned}
M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) = \max_{l,g} &\Bigg[\underbrace{z h_{m_i}(1-l)}_{\text{output value}} - \underbrace{(1-g)l \times q p_e(h_{m_i})}_{\text{pay for external trainers}} \Bigg] \\
&\underbrace{\int_{j \in \mathbb{I}(z), j \neq i} \frac{\partial s}{\partial p} \left[p(h_{m_j}, \mathbb{H}(z)) - p(h_{m_j}, \mathbb{H}(z) \backslash \{i\}) \right] \times O^{a_j+1}(h_{m_j}, z, \mathbb{H}'(z)) dj}_{\text{spillover effects to other workers in firm}} \\
&+ \underbrace{s(h_{m_i}, \mathbb{H}(z)) \times \rho \mathbb{E} \left[(1-\kappa) M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\} \right]}_{\text{future value if successfully climb human capital ladder}} \\
&+ \underbrace{\left[1 - s(h_{m_i}, \mathbb{H}(z))\right] \times \rho \mathbb{E} \left[(1-\kappa) M^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i}), V_{TR}^{a_i+1}(h_{m_i})\} \right]}_{\text{future value if don't successfully climb human capital ladder}} .
\end{aligned}
\tag{3}
$$

The first line comprises the current output value net of training costs stemming from foregone production and payments to external trainers. The second line describes the value of worker i's spillover effects to its coworkers in firm z stemming from worker i's influence on firm z's internal learning environment. These spillover effects capture worker i's impact on his cowork-

[28]The assumption that workers only pay external trainers if they exhibit higher human capital than her own increases the returns to human capital in the training sector, and causes the average human capital level of trainers to rise above that of production workers.

ers' match values by comparing the match value each other worker j has when worker i is in the workforce of firm z to the match value worker j would have if worker i was absent from the workforce of firm z in the current period, taking all else as given. In particular, the expression $\frac{\partial s}{\partial p}\left[p(h_{m_j}, \mathbb{H}(z)) - p(h_{m_j}, \mathbb{H}(z)\backslash\{i\})\right]$ reflects the change in worker j's learning probability due to the presence of worker i in the workforce of firm z, and $O^{a_j+1}(h_{m_j}, z, \mathbb{H}'(z)) = \rho\mathbb{E}\left[(1-\kappa)\Delta M^{a_j+1}(h_{m_j+1}, z, \mathbb{H}'(z)) + \kappa\Delta\max\{V_U^{a_j+1}(h_{m_j+1}), V_{TR}^{a_j+1}(h_{m_j+1})\}\right]$ captures the gains in worker j's future match value due to one-step increment in human capital.[29] The value of these spillover effects will be positive for workers with high human capital levels, while the opposite will be true for workers with low human capital levels.[30] The remaining two lines capture the discounted future values contingent on whether worker i climbs the human capital ladder in the current period or not. Appendix D.1 provides details on the derivation of the match value.

We can solve for the total time spent on learning and the proportion of this time allocated to internal learning to maximize the match value by taking the first-order conditions of Equation (3). The following proposition summarizes these results.

Proposition 1 (Optimal learning investments)**.** *For worker i at firm z, the optimal learning investments (if they are internal solutions) satisfy:*

(i) The optimal total learning time l satisfies:

$$l = \left(\frac{\gamma s^E(h_{m_i}, \mathbb{H}(z)), \mathbb{H}(z)) O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z))}{z h_{m_i} + (1-g) q p_e(h_{m_i})}\right)^{1/(1-\gamma)}. \tag{4}$$

(ii) The optimal share of learning time spent on internal learning g satisfies:

[29] $\mathbb{H}(z)\backslash\{j\}$ is the human capital distribution of firm z's workforce absent worker j. The expressions $\Delta M^{a+1}(h_{m_i+1}, z, \mathbb{H}'(z)) = M^{a+1}(h_{m_i+1}, z, \mathbb{H}'(z)) - M^{a+1}(h_{m_i}, z, \mathbb{H}'(z))$ and $\Delta\max\{V_U^{a+1}(h_{m_i+1}), V_{TR}^{a+1}(h_{m_i+1})\} = \max\{V_U^{a+1}(h_{m_i+1}), V_{TR}^{a+1}(h_{m_i+1})\} - \max\{V_U^{a+1}(h_{m_i}), V_{TR}^{a+1}(h_{m_i})\}$ capture the benefits from climbing the human capital ladder when the worker stays in the current firm, and when the worker becomes unemployed, respectively. The expectation is taken with regard to uncertainty about realizations of human capital depreciation, which may lead the worker to descend the human capital ladder. $\mathbb{H}'(z) = \Gamma(\mathbb{H}(z))$ is the law of motion for the human capital distribution in firm z.

[30] When deriving this match value, and in order to avoid double-counting, we also subtract the spillover effects generated by the other workers in firm z to worker i, which conceptually capture additional costs from internal learning as workers directly compensate internal mentors when they engage in internal learning. This term cancels out for every worker, however, because the impacts from all other workers in the firm to the learning prospects of worker i offset each other. We explain this in detail when we derive the match value in Appendix D.1.

$$
\begin{aligned}
lqp_e(h_{m_i}) = &\left[s^E(h_{m_i}, \mathbb{H}(z))^{\frac{1}{\sigma}} \left(\frac{(A_e(z)p_e(h_{m_i}))^{\frac{\sigma-1}{\sigma}}}{(1-g)^{\frac{1}{\sigma}}} - \frac{(A(z)p(h_{m_i}, \mathbb{H}(z)))^{\frac{\sigma-1}{\sigma}}}{g^{\frac{1}{\sigma}}} \right) \right] \\
&\times O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z))l^{\gamma} = 0.
\end{aligned} \tag{5}
$$

Proof: See Appendix D.3.

Result (i) presents the optimal total learning time, which balances the benefits and costs of human capital accumulation. The numerator of the right-hand side of Equation (4) represents the benefits from learning, which include gains in match value (if the worker stays) and unemployment value (if the worker is separated from the firm) when the worker successfully accumulates human capital. The denominator captures the marginal cost of increasing the probability of accumulating human capital. Since there are diminishing returns to learning time as $\gamma < 1$, we can determine the optimal learning time from this equation.

Result (ii) characterizes the share of learning time spent on internal learning. The left-hand side of Equation (5) captures the benefit of increasing internal learning, namely reducing learning costs as learning internally is cheaper than learning externally. The right-hand side captures the costs of increasing internal learning, corresponding to a potentially lower learning efficiency if the success rate of external learning ($p_e(h_{m_i})$) is high. All else being equal, a higher success rate of internal learning $p(h_{m_i}, \mathbb{H}(z))$ compared to external learning $p_e(h_{m_i})$, or higher additional costs of external learning q, lead to a greater share of learning time allocated to internal learning.

This latter result is key for understanding the lifecycle switch between learning modes. Initially, when workers are young and possess low human capital, they join the production sector and encounter a large group of coworkers with higher human capital than their own. This results in a high probability of success for internal learning, along with a relatively low cost of internal learning compared to external learning as their productivity is low relative to the training fees. These factors lead young workers to spend a significant portion of their time in internal learning. As workers age and their human capital increases, the number of coworkers with higher human capital than the workers' own decreases, and the relative cost of internal learning rises. This leads workers to spend an increasingly larger portion of their time in external learning, which incurs a cost but involves matching with a better pool of mentors, thereby increasing the likelihood of climbing the human capital ladder.

While the level of learning is chosen to maximize the match value, it is important to note that this learning choice is inefficiently low compared to the social optimum, as it does not internalize the gains to future employers and coworkers after workers leave the firm.[31] In Appendix E.5 we evaluate the impact of subsidies that cover a portion of learning costs to correct this inefficiency. We find that subsidizing learning can lead to sizeable increases in human capital and GDP, particularly if it targets external learning.

4.2.2 Wage determination

Following Cahuc et al. (2006) and Bagger et al. (2014), we set the wage of a worker i of age a_i and human capital h_{m_i} to be proportional to the current-period revenue she generates, which includes net output value and spillovers to and from other workers:

$$w^{a_i}(h_{m_i}, z, \mathbb{H}(z)) = r\left[z h_{m_i}(1-l) - (1-g)l \times q p_e(h_{m_i}) + X^{a_i}(h_{m_i}, z, \mathbb{H}(z)) \right].$$

$X^{a_i}(h_{m_i}, z, \mathbb{H}(z))$ captures the spillover effects generated by worker i to other coworkers in firm z, as shown in the second line of Equation (3).[32] r represents the contractual piece rate that determines the worker's wage, and captures the share of revenue attributed to worker i. This share is determined by bargaining upon hiring and remains constant for the duration of the match. Specifically, when the job match is formed, firm z and worker i engage in Nash bargaining to determine the division of match surplus:

$$\max_{r} \; [V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z)) - V_U^{a_i}(h_{m_i})]^{\beta} \, J^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))^{1-\beta}. \tag{6}$$

$V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))$ denotes worker i's value from the match, and $V_U^{a_i}(h_{m_i})$ captures her reservation value (unemployment value). $J^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))$ denotes firm z's value from the match, and the reservation value for the firm is zero. We present the formulas for the worker's value $V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))$ and the firm's value $J^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))$ in Appendix D.2. The worker's and firm's values add up to the total match value, $M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) = J^{a_i}(r, h_{m_i}, z, \mathbb{H}(z)) + V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))$. β governs the negotiation power of the worker. According to the first-

[31] Another potential inefficiency arises from the pattern of positive assortative matching generated in equilibrium, which can result in low-skill individuals having too few opportunities to learn from more knowledgeable coworkers. Recent papers focusing on coworkers' learning have documented the importance of this inefficiency (Herkenhoff et al. (2024), Jarosch (2021)). We concentrate on inefficiencies stemming from underinvestment in skills instead.

[32] In other words, $X^{a_i}(h_{m_i}, z, \mathbb{H}(z)) = \int_{j \in \mathbb{I}(z), j \neq i} \frac{\partial s}{\partial p} \left[p(h_{m_j}, \mathbb{H}(z)) - p(h_{m_j}, \mathbb{H}(z) \backslash \{i\}) \right] \times O^{a_j+1}(h_{m_j}, z, \mathbb{H}'(z)) dj$.

order condition, the piece rate r satisfies

$$\frac{V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z)) - V_U^{a_i}(h_{m_i})}{\beta} = \frac{J^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))}{1-\beta} = M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) - V_U^{a_i}(h_{m_i}). \quad (7)$$

We can solve for the specific piece rate that delivers the bargaining outcome in Equation (7). The resulting piece rate $r^{a_i}(h_{m_i}, z, \mathbb{H}(z))$ depends on the worker's age and human capital level upon hiring, and varies across firms.[33]

4.2.3 Solving for optimal vacancy posting

Each firm z will determine the optimal number of vacancies $v(z)$ to post to maximize its profits from hiring workers:

$$\max_{v(z)} \frac{v(z)}{V} \sum_{m=1}^{M} \sum_{a=1}^{J} \lambda_U (1-\beta) \left[M^a(h_m, z, \mathbb{H}(z)) - V_U^a(h_m) \right] D_U^a(h_m) - c_v \frac{v(z)^{1+\gamma_v}}{1+\gamma_v}, \quad (8)$$

where $D_U^a(h_m)$ denotes the measure of unemployed workers of age a with human capital h_m. The first-order condition yields

$$c_v v(z)^{\gamma_v} = \sum_{m=1}^{M} \sum_{a=1}^{J} \frac{\lambda_U(1-\beta)}{V} \left[M^a(h_m, z, \mathbb{H}(z)) - V_U^a(h_m) \right] D_U^a(h_m). \quad (9)$$

The left-hand side of this equation captures the marginal cost of posting a vacancy, which increases with the number of vacancies. The right-hand side captures the aggregate value per vacancy from hiring unemployed workers, with $(1-\beta)$ governing the firm's share of the increment in surplus from hiring, as shown in Equation (7).

Equilibrium. In Appendix D.4 we define the general equilibrium of the model.

4.3 Calibration

We calibrate the above framework to the United States to evaluate the importance of internal and external learning for workers' human capital acquisition and wage growth. We use a Cobb-Douglas job matching function between searchers and vacancies (Shimer (2005)),

[33] Cahuc et al. (2006) analytically solve for the piece rate and show that it increases with the worker's current outside option. Our model with endogenous human capital formation does not yield an analytical solution for the piece rate. We therefore rely on numerical methods to first solve for the value of employment at any piece rate r and then find the specific piece rate that delivers the value determined by the worker's bargaining outcome.

which implies a meeting rate between the unemployed and firms of $\chi(x) = c_M x^{\phi}$, where $0 < \phi < 1$ is the elasticity of the number of job matches to the number of vacancies. We parameterize the efficiencies of learning internally and externally as $A(z) = \bar{A}z^{\alpha}$ and $A_e(z) = \bar{A}_e z^{\alpha}$, respectively, where $\alpha > 0$ captures potential differences in learning efficiency in more productive firms. We parameterize firm productivity to be Pareto-distributed $\Phi(z) = 1 - z^{-\zeta}$, where ζ is the shape parameter.

4.3.1 Externally calibrated parameters

We first draw some common parameters directly from the literature. These externally calibrated parameters are presented in panel A of Table 4.1. A period in the model is one quarter. We calibrate the discount rate to be $\rho = 0.99$ and consider that each worker stays in the labor force for $J = 160$ periods, corresponding to 40 years of working life. Without loss of generality, we let the step size of the human capital ladder be $\gamma_h = 0.05$ such that climbing a step implies a 5% increase in human capital. We normalize the lower bound of the human capital ladder h_1 to 1. We calibrate the elasticity of the number of matches to the number of vacancies, $\phi = 0.3$ following Shimer (2005).

Due to the lack of US estimates, we calibrate two other parameters using evidence from other countries. There is a wide range of estimates on bargaining power β in labor search models with wage negotiation and human capital accumulation. For example, Bagger et al. (2014) find β to be between 0.29–0.32 using Danish employer-employee data, while Gregory (2019) estimates this to be 0.66 using data from Germany. We set $\beta = 0.5$ following these estimates, such that workers and firms have even bargaining power.[34] Finally, we obtain the curvature of vacancy costs $\gamma_v = 1$ from Dix-Carneiro et al. (2019)'s estimate for Brazilian firms, implying that vacancy costs are quadratic in the number of vacancies.

4.3.2 Internally calibrated parameters

We are left with 11 parameters to estimate: the parameters governing the efficiencies of learning from internal and external sources $\{\bar{A}, \bar{A}_e, \alpha\}$, the elasticity of substitution between the two modes of learning σ, the exogenous learning-by-doing probability of moving up the human capital ladder ϵ, the degree of diminishing returns of learning time in producing new human capital γ, the depreciation rate of human capital δ_h, the constant in the matching function c_m, the constant in vacancy costs c_v, the shape parameter of the firm productivity

[34] In Appendix F.3 we show that our quantitative findings are robust to considering different values of β.

distribution ζ, and the exogenous separation rate of workers κ.

Table 4.1: Parameter values

Label	Description	Value	Source
	Panel A: Externally calibrated parameters		
ρ	Discount rate	0.99	Annual interest rate of 0.04
J	Working life	160	Working life of 40 years
γ_h	Step size of human capital ladder	0.05	Authors' choice
h_1	Lower bound of human capital ladder	1	Normalization
ϕ	Elasticity of matches to vacancies	0.3	Shimer (2005)
β	Workers' bargaining power	0.5	Estimates in literature
γ_v	Curvature of vacancy cost	1	Dix-Carneiro et al. (2019)
	Panel B: Internally calibrated parameters		
$\bar{A}$	Constant in efficiency of internal learning	0.59	Joint estimation
$\bar{A}_e$	Constant in efficiency of external learning	0.27	Joint estimation
α	Elasticity of learning efficiency to firm productivity	0.89	Joint estimation
σ	Elasticity of substitution between internal/external learning	3.75	Joint estimation
ϵ	Exogenous human capital gain	0.02	Joint estimation
γ	Degree of diminishing returns in learning	0.39	Joint estimation
δ_h	Depreciation rate of human capital	0.03	Joint estimation
c_m	Constant in matching function	0.45	Joint estimation
c_v	Constant in vacancy costs	0.33	Joint estimation
ζ	Shape parameter of firm productivity distribution	4.36	Joint estimation
κ	Exogenous separation rate	0.03	Joint estimation

We jointly estimate these parameters using the method of moments to minimize the squared differences between the model and data moments. To do this, we target 11 data moments of which the first 7 are: the unemployment rate from 1994 to 2007 from FRED; the labor market tightness captured by the ratio of the number of vacancies to the number of unemployed people from 2000 to 2007 from FRED; the tail shape parameter of the firm employment distribution as estimated by Axtell (2001); the share of workers that remain employed in the next quarter from Donovan et al. (2020); the ratio of new (1 year of experience) to all workers' average external learning time from the NHES data;[35] and the average quarterly returns to experience within 0–40 years of experience and within 0–25 years of experience from Lagakos

[35] Since the NHES data only provides extensive margin measures of learning, we use the average participation rate in external learning of new workers relative to all workers to construct the relative share of time spent on external learning by these two groups. Given the importance of this moment to estimating the elasticity of substitution between learning modes, and since the lack of the intensive margin in the data may lead to an underestimate of this share, in Appendix F.3 we show that our quantitative results are robust to considering a higher value of this elasticity.

et al. (2018), and specifically the returns stemming from human capital.[36]

Because the NHES data does not contain information on the time spent on learning nor employers' information, we compute the remaining 4 moments using the 1995 US-SEPT described in Section 3.[37] These moments correspond to the shares of total time spent on internal and external learning respectively, the ratio of workers' average learning time in firms with 100+ workers to that in firms with 50–99 workers, and the ratio of old workers' average total learning time to all workers' average total learning time.[38]

Our choice of moments to target is motivated by the link of the parameters to the moments. For instance, the constant in vacancy costs, c_v, directly influences labor market tightness, and given this tightness, the constant in the job matching function, c_m, determines the unemployment rate. The efficiencies of internal and external learning, $\bar{A}$ and $\bar{A}_e$, are informed by the shares of time spent on internal and external learning, respectively. A higher elasticity of learning efficiency to firm productivity, α, suggests a larger gap in the learning environments of different firms. A higher elasticity of substitution between learning modes, σ, indicates a more substantial shift between learning sources from young to old. The degree of diminishing returns to learning, γ, informs how much larger the total learning time of young workers is relative to that of older workers. Finally, given the strength of internal and external learning, the lifetime and youth wage growth from human capital are informative of exogenous human capital gains ϵ and human capital depreciation δ_h, respectively. In Table E.1, we report the Jacobian matrix showing how changes in each parameter affect the relevant chosen moments, further illustrating the aforementioned mapping between the parameters and the moments.

Panel B of Table 4.1 displays the values of the internally calibrated parameters. Overall, the parameter values are reasonable and consistent with other studies. For example, the elasticity of learning efficiency to firm productivity is $\alpha = 0.89$, indicating higher learning efficiency in more productive firms, consistent with estimates in papers that use similar

[36]We focus on the wage returns stemming from human capital since in our baseline model, we concentrate on human capital and exclude other factors of lifetime wage growth, such as job ladders. To determine the share of experience returns stemming from human capital, we follow the literature (Altonji et al. (2013), Bagger et al. (2014)), which has found that human capital contributes to approximately 50% of wage growth. As such, we aim to target 50% of the observed wage growth in our baseline model. In Appendix F.1 we integrate on-the-job search and job ladders into the model and target the full observed wage growth. This extension results in similar quantitative results.

[37]See Appendix A.2 for details on this data and learning definitions.

[38]We consider old workers as workers aged 55 or more in the data, and as workers in the last 10 years of working life in the model.

assumptions (e.g., Engbom, 2021). The elasticity of substitution between the two modes of learning is $\sigma = 3.75$, suggesting moderate substitutability. The degree of diminishing returns to learning is $\gamma = 0.39$, which is similar to that estimated by Manuelli and Seshadri (2014). Each person has a 2% chance to climb the human capital ladder exogenously in each period. Our calibrated quarterly depreciation rate of human capital from learning, $\delta_h = 0.03$, aligns with the depreciation rate of training returns estimated by Blundell et al. (2021) using British labor surveys. Our calibrated quarterly job destruction rate, $\kappa = 0.03$, is close to the employment-unemployment transition probability estimated by Shimer (2012). With these calibrated parameters, our model matches the targeted data moments well, as shown in Table 4.2.

Table 4.2: Targeted moments in the data and model

Description	Model	Data
1. Unemployment rate	0.06	0.06
2. Labor market tightness (#vacancies/#unemployed)	0.57	0.55
3. Shape parameter of firm employment distribution	1.23	1.10
4. Share of workers that remain employed in next quarter	0.96	0.94
5. Average wage growth from HC (per quarter) within 0–40 years of experience	0.0024	0.0023
6. Average wage growth from HC (per quarter) within 0–25 years of experience	0.0039	0.0040
7. Share of total time spent on external learning	0.032	0.033
8. Share of total time spent on internal learning	0.013	0.014
9. Ratio of total learning time in 100+ worker firms to 50–99 worker firms	1.15	1.13
10. Ratio of new to all workers' average time spent on external learning	1.49	1.51
11. Ratio of old to all workers' average learning time	0.49	0.51

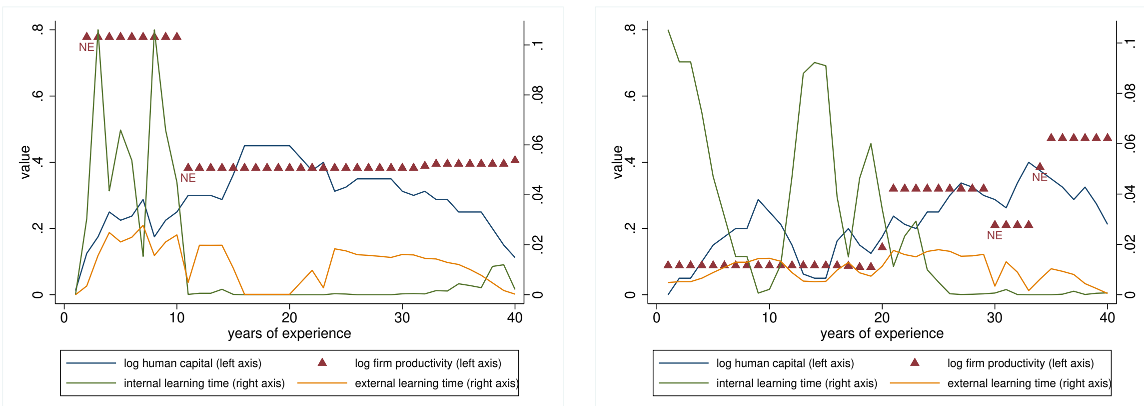
4.4 Properties of equilibrium: workers' learning patterns

We now investigate some key properties of our model's stationary equilibrium. Specifically, we show how learning is influenced by workers' lifecycles, their relative positions on the human capital ladder, and employers' characteristics. To accomplish this, we begin by presenting two examples of workers with distinct employment paths in Figure 4.1. The plots illustrate the progression of internal and external learning, human capital, and employer productivity throughout workers' lifecycles. Despite the different employment paths of these two workers (for instance, the worker in panel (a) encounters many more separation shocks than that in panel (b)), they share two characteristics that illustrate key patterns in the model.

First, when these workers are young, they predominantly learn from internal sources. However, as they age, these workers transition away from internal sources and toward external sources, eventually leading to less learning overall as they ascend the human capital ladder. We label this pattern the *lifecycle pattern of worker learning*. Second, when paired with more productive firms, these workers allocate more time to both internal and external learning, resulting in a quicker increase in human capital. We refer to this pattern as the *firm productivity pattern of worker learning*.

Figure 4.1: Examples of workers' lifecycles

(a) (b)



Notes: These figures illustrate the lifetime progression of internal and external learning, human capital, and employer (firm) productivity for two workers in the model. "NE" denotes that the worker partly experiences unemployment or becomes a trainer during the current year. If the worker spends the whole year being unemployed or a trainer, then this year will have no observation of employer productivity.

4.4.1 The lifecycle pattern of worker learning

The *lifecycle pattern of worker learning* arises from the evolving incentives to engage in different sources of skill acquisition throughout the lifecycle. When workers are young and possess low levels of human capital, they have a lengthy future career to reap the benefits of human capital accumulation and also encounter a large group of coworkers with higher human capital than their own. This encourages learning investments, particularly through internal sources. As workers climb the human capital ladder, the share of coworkers with higher human capital than their own narrows, reducing the likelihood of learning internally. Consequently, firms and workers opt to invest more in external learning, enabling the worker to connect with trainers who, on average, possess higher levels of human capital than their coworkers.[39] Eventually, however, the prevalence of external learning decreases as workers

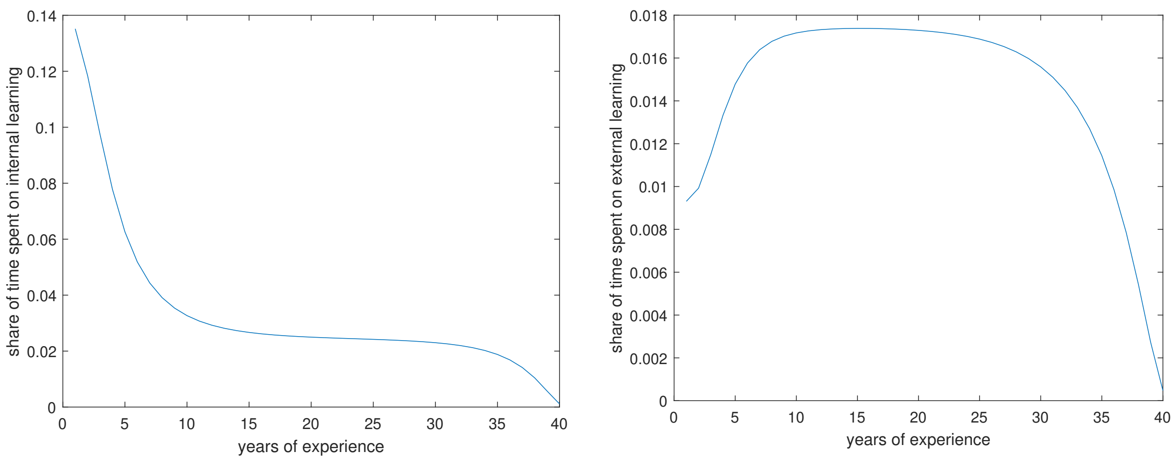
[39] This is illustrated in Figure E.2, which plots the distribution of human capital for production workers and trainers, and shows that the latter is skewed left relative to the former. This follows from the fact that the

age (which shortens the time to enjoy learning benefits) and become more productive (which raises the opportunity costs of learning while reducing its success rate).[40]

Figure 4.2 illustrates these results by plotting the average shares of time spent on internal and external learning throughout the lifecycle. The model predicts that as workers age, the time spent on internal learning decreases, while the time spent on external learning initially increases and then declines. This matches the empirical evidence presented in Section 3.2.

Figure 4.2: Lifecycle patterns of internal and external learning

(a) Internal learning (b) External learning



Notes: These figures plot the share of time spent on internal and external learning, namely $l \times g$, and $l \times (1-g)$, respectively, for workers with different years of experience in the model.

In Appendix G.1 we present evidence supporting three testable predictions that validate and offer empirical backing for these findings. Specifically, we provide empirical evidence showing that: (1) trainers possess higher average human capital levels than production workers, implying that they are better suited to teach more-experienced workers; (2) the proportion of individuals engaging in learning-by-doing increases with human capital, aligning with the rise in work hours and the decline in external and internal learning associated with seniority; and (3) internal learners display lower levels of task complexity compared to external learners,

returns to human capital are higher in the training sector.

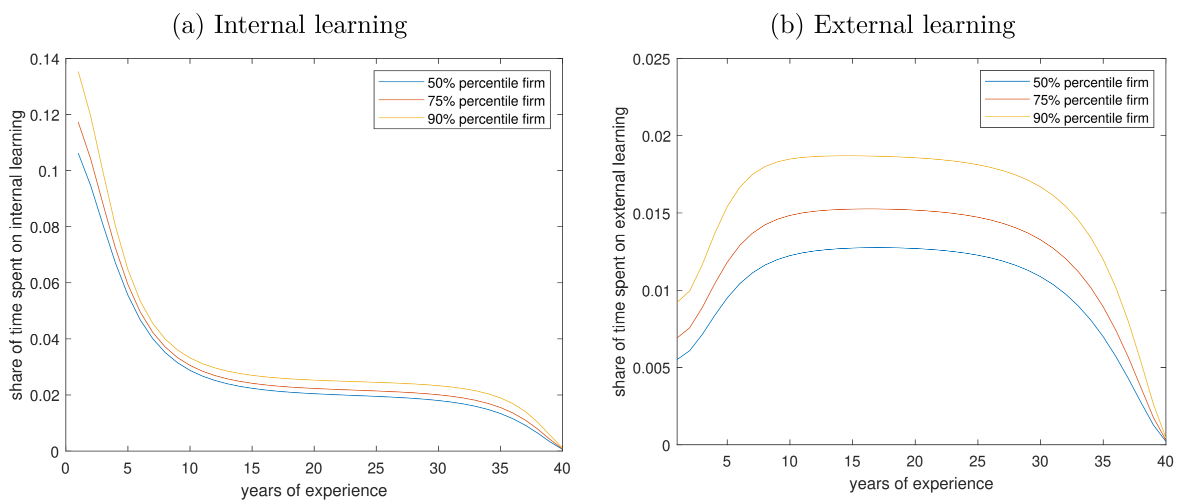
[40]Further, in Figure E.1 we show how the distribution of human capital changes as workers age by plotting the distribution of human capital levels for a given cohort of workers observed at different ages. We find that workers' human capital growth is rapid during the first few years after entering the labor force, and slows down in later years, consistent with the evidence on the lifecycle returns to experience (Rubinstein and Weiss (2006)). In our model, this slowdown stems not only from the depreciation of human capital or aging, but also from the reduction in the scope of learning that occurs as workers climb up the human capital ladder and have fewer colleagues and trainers to learn from.

which is consistent with internal learners having lower levels of human capital and being more easily trained by coworkers.[41]

4.4.2 The firm productivity pattern of worker learning

The *firm productivity pattern of worker learning* arises from the higher returns to skill acquisition prevalent in more productive firms. Figure 4.3 illustrates these results by plotting the average shares of time allocated to internal and external learning for workers of different ages, for firms at the 50th, 75th, and 90th percentiles of the firm productivity distribution. We find that more productive firms provide better learning environments by offering a greater variety of learning options. This finding matches our empirical evidence of Section 3.2 showing that workers in larger European firms are presented with more learning opportunities, and the evidence found by Engbom (2017), Arellano-Bover (2020), and Arellano-Bover and Saltiel (2023) showing that workers in more productive firms exhibit faster rates of skill acquisition. This finding is also consistent with the evidence found by Gregory (2019) showing that having

Figure 4.3: Lifecycle patterns of internal and external learning by firm productivity



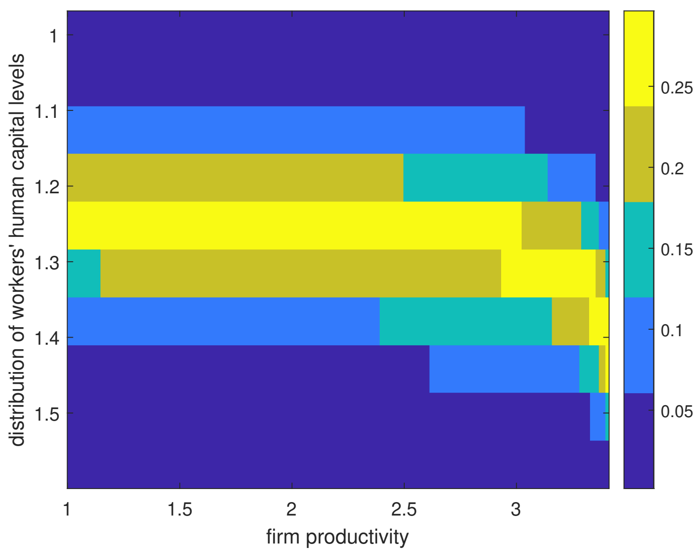
Notes: These figures plot the share of time spent on internal and external learning, namely $l \times g$, and $l \times (1 - g)$, respectively, for workers with different years of experience and working in firms with different productivity levels in the model.

[41] In addition, to validate our modeling of external learning, we cross-check the size of the external learning costs suggested by our model with the data. The US-SEPT data reveals that the overall external training payment amounts to 13% of trainees' wage costs due to lost production time. In this data, external training payments include tuition reimbursements, direct payments to outside trainers, contributions to outside training funds, and subsidies for training received from outside sources. In our model, the payments to external trainers amount to 15% of trainees' wage costs due to lost production time, bearing a notable similarity to the data.

different forms of training available is important for firms' learning environments.[42,43]

The higher returns to skill acquisition prevalent in more productive firms stem from the supermodularity of the production function, as well as greater learning efficiency and an improved pool of coworkers due to positive assortative matching. This latter result is illustrated in Figure 4.4, which presents the human capital distribution within different firms by using different colors to indicate the share of workers of each human capital level in the workforce of a firm of each productivity level. We observe that more productive firms have relatively larger shares of high-skill workers. This aligns with the positive sorting between employers and employees documented in the US (Barth et al. (2016), Abowd et al. (2018), Song et al. (2019)). In our model, this positive sorting pattern is driven by the larger learning investments and more favorable learning environments prevalent in more productive firms, which enable workers to ascend the human capital ladder more rapidly.

Figure 4.4: Human capital distribution within firms



Notes: This figure illustrates the human capital distribution within each firm by using the different colors in the color scale referenced to the right of the plot to indicate the share of workers of each human capital level in the workforce of a firm of each productivity level.

[42] This finding is also consistent with the evidence documented in Figure C.7 which shows that the lifecycle patterns of learning documented in Germany are robust to considering different firm sizes, and that the levels of both internal and external learning vary positively with firm size.

[43] In Figure E.3 we plot how wages evolve with human capital for firms with varying productivity levels. Our findings reveal that wages increase with human capital, and that at all human capital levels, more productive firms offer higher wages. Consequently, despite providing greater learning (partially funded by reduced wages), more productive firms still pay higher wages due to their higher productivity levels.

5 Counterfactual analysis

To evaluate the contribution of each learning source to explaining lifecycle wage and human capital dynamics, we now conduct counterfactual exercises that progressively shut down each of the two learning sources, and observe how the stationary equilibrium shifts. To achieve this, we subsequently set the constant efficiency terms of external and internal learning, $\bar{A}$ and $\bar{A}_e$, to zero. As a result, the time devoted to learning from internal and external sources, respectively, has no impact on the likelihood of climbing the human capital ladder.

5.1 Role of each learning source in aggregate human capital

Table 5.1 summarizes the average share of time spent on each source of learning along with the average level of human capital in the stationary equilibrium of our baseline model, and the stationary equilibria of the models without different forms of learning. We find that internal and external learning contribute equally to aggregate human capital: without either external or internal learning, workers' average human capital decreases by 14%. Without both sources of learning, average human capital decreases by 25%. In this case, there is still some human capital accumulation due to the exogenous learning-by-doing probability. Absent all learning sources, average human capital remains fixed at the initial human capital level $h_1 = 1$, and would thus be 28% lower than that in the calibrated economy.

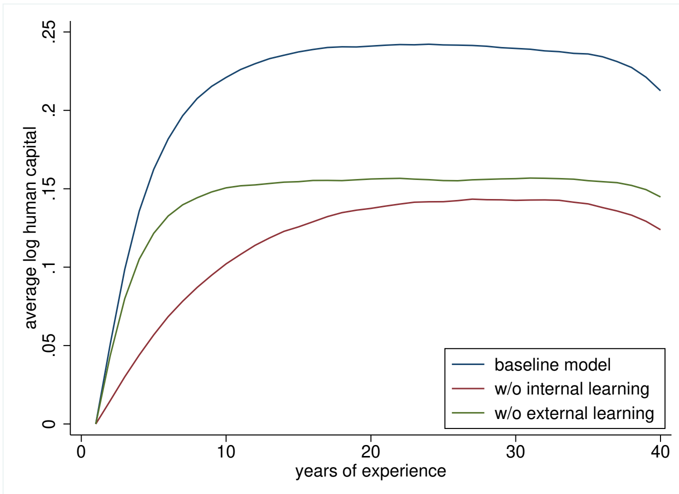
Table 5.1: Learning and human capital in the baseline and counterfactual scenarios

	Workers' share of time spent on learning		
	External learning	Internal learning	Avg human capital
Calibrated economy	1.35%	3.15%	1.39
W/o external learning	0	2.21%	1.20
W/o internal learning	2.43%	0	1.20
W/o both sources of learning	0	0	1.04
W/o any form of learning	0	0	1

Notes: This table presents the share of time workers spend on internal and external learning along with the average level of human capital in the baseline calibrated model economy and the counterfactual scenarios that respectively shut down: internal learning, external learning, both sources of learning (internal and external learning), and all sources of learning (internal and external learning and learning-by-doing).

In Figure 5.1 we further investigate the role of internal and external learning in human capital formation by plotting the trajectory of human capital throughout the lifecycle in the two counterfactual scenarios. We find that although lifetime human capital growth is considerably lower in the absence of either internal or external learning compared to the baseline model, internal learning is relatively more critical to human capital formation during youth, whereas external learning is relatively more critical at older ages.

Figure 5.1: Lifecycle human capital growth in the baseline and counterfactual scenarios



Notes: This figure illustrates the lifetime progression of human capital by plotting the average log human capital level for workers with different years of experience in the baseline model and counterfactual scenarios that shut down internal and external learning, respectively.

In addition, we find that the lifecycle patterns of worker learning also change in the two counterfactual scenarios. As depicted in Figure E.4, the absence of external learning results in lower levels of internal learning relative to the baseline case at all levels of experience, owing to inadequate within-firm knowledge pools as workers cannot learn from training experts. In contrast, in the absence of internal learning workers spend significantly more time learning externally due to the lack of coworker mentoring. This is particularly marked during youth, as the returns to learning are higher then. As a consequence, we find that in the absence of internal learning, the time spent on external learning declines with worker experience. This does not align with our data and thus emphasizes the importance of modeling the two learning modes to accurately reproduce our empirical findings.

5.2 Role of learning sources in wage growth and dispersion

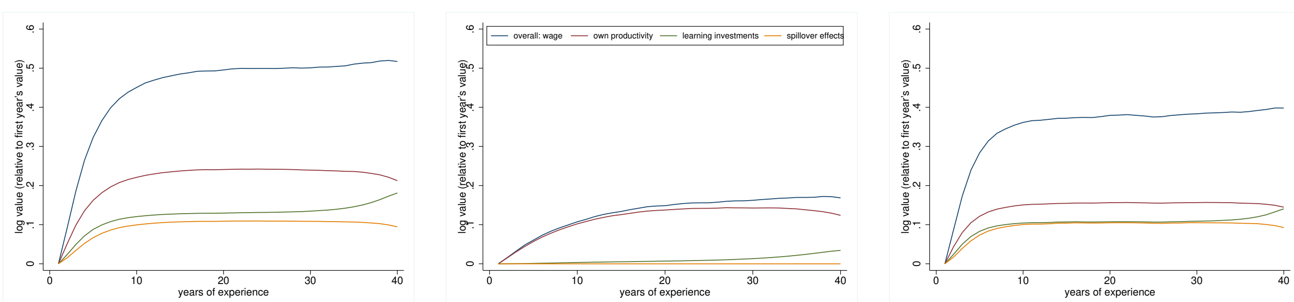
We now examine the changes in wages over the lifecycle in our benchmark and counterfactual scenarios. To understand the factors influencing wage dynamics, we break down the wage into several components:

$$
\log w^{a_i}(h_{m_i}, z, \mathbb{H}(z)) = \underbrace{\log r}_{\text{piece rate}} + \underbrace{\log z}_{\text{firm productivity}} + \underbrace{\log h_{m_i}}_{\text{worker's own productivity}} + \underbrace{\log\left(\frac{zh_{m_i} + X^{a_i}(h_{m_i}, z, \mathbb{H}(z))}{zh_{m_i}}\right)}_{\text{spillover effects to other workers in the firm}}
$$
$$
+ \underbrace{\log\left(\frac{zh_{m_i}(1-l) - (1-g)lqp_e(h_{m_i}) + X^{a_i}(h_{m_i}, z, \mathbb{H}(z))}{zh_{m_i} + X^{a_i}(h_{m_i}, z, \mathbb{H}(z))}\right)}_{\text{learning costs}}. \tag{10}
$$

In Figure 5.2 we plot lifetime wage growth in the baseline and counterfactual scenarios, along with the different human capital components driving it.[44] The contribution of human capital to lifetime wage growth in the baseline model encompasses three components. First, as workers accumulate human capital throughout their lives, they become more productive in the production process, resulting in higher wages (red line). This represents the standard gain of human capital examined in the literature. Second, as workers ascend the human capital ladder, they are more likely to function as "teachers" within the firm, generating spillover effects for their colleagues, which they are compensated for (yellow line). Third, learning investments and thus learning costs tend to decrease with human capital and age, further boosting wages (green line). We find that the latter two components account for 36% of the wage gains from human capital over 25 years of experience, with changes in learning

Figure 5.2: Lifecycle wage growth and its components in the baseline and counterfactual scenarios

(a) Baseline model (b) Without internal learning (c) Without external learning



Notes: These figures illustrate the lifetime progression of wage growth and its different components (own productivity, spillover effects to coworkers, and learning investments and associated costs) by plotting the average log value of each of these (normalized to the first year's value) for workers with different years of experience in the baseline model and counterfactual scenarios that shut down internal and external learning, respectively.

[44] Since our model abstracts from job ladders, the first two components in Equation (10), namely the average share of revenue attributed to workers, and the average productivity of the firms workers are matched with do not change significantly over workers' lifetimes. Thus, we omit these two components for the sake of clarity.

investments accounting for 22 percentage points of this, and changes in spillover effects accounting for the remaining 14 percentage points. This highlights the importance of including spillover effects and learning investments when analyzing lifetime wage growth.

In the model without external learning, wage growth from all components is lower than that in the baseline, yet the patterns remain qualitatively similar. On the other hand, when we disable internal learning, spillover effects are absent as colleagues cannot serve as "teachers," and learning costs also vary more gradually throughout the lifecycle. In this scenario, wage growth follows predominantly from increases in workers' own productivity. Hence, incorporating internal learning is crucial for understanding lifecycle wage dynamics: a model that excludes internal learning might inaccurately attribute wage growth to increases in workers' own productivities, rather than increases in compensated non-own productivity, such as coworker learning spillovers.

In Appendix G.2 we present empirical evidence supporting the importance of coworker instruction in many occupations in the US. To achieve this, we use data from the Department of Labor's O*NET project, which describes the tasks pertaining to each occupation, along with the mix of knowledge, skills, and abilities required to perform these tasks. Our findings reveal that coworker tutoring plays a vital role in numerous non-teaching occupations, and that highly skilled workers primarily perform these occupations.

In addition, in Appendix E.3.3 we consider how wage dispersion changes over the lifecycle in the baseline and counterfactual scenarios. We find that the dispersion in workers' own productivity levels is more driven by external learning than by internal learning, since without external learning, workers catch up to colleagues relatively fast. Nevertheless, we also find that even with the broader own-productivity dispersion, wage dispersion is actually lower when we shut down internal learning compared to when we shut down external learning. This follows from the spillover effects linked to internal learning, which contribute significantly to the wage gains of more-experienced workers. This further reinforces our finding that incorporating internal learning is key for understanding lifetime wage dynamics.

6 Impact of remote work on human capital and wages

We now use the model to examine the effects of remote work on human capital and wage dynamics. To accomplish this, we use data from recent papers documenting the rise of remote work after the Covid-19 pandemic and the impact of the disruption of face-to-face contact on the opportunities for internal learning. Specifically, we rely on the findings of Barrero et al. (2021) who document a rise in the share of remote work days from 5% to 20% before and after the pandemic, together with the findings of Emanuel et al. (2023) showing a 15% reduction in the feedback provided by adjacent coworkers following the shift to remote work. Using these two observations, we infer a 2.25% decrease in the rate of contact between colleagues (15% × 15%) in the post-pandemic era. We incorporate this information into our model by interpreting it as a 2.25% decrease in the success rate of internal learning. In addition, we also examine an extreme scenario where the share of remote working days rises to 100%, implying a 14.25% drop (95% × 15%) in the success rate of internal learning.

We summarize the results of the impact of remote work on average human capital and lifecycle wage growth in Table 6.1. Our findings suggest that the post-pandemic surge in remote work leads to a 0.44% reduction in workers' average human capital, which constitutes 4.73% of the human capital accumulated through internal learning.[45] In the extreme scenario where all work is performed remotely, average human capital drops by 2.71%, and the gains from internal learning fall by 28.82%. Given that internal learning concentrates among younger workers, these results also imply that remote work can have significant and enduring spillover effects, as these younger workers who learn less due to remote work will be less equipped to mentor future young employees when they are more senior.

Table 6.1: Impact of remote work on human capital and wage growth

		Impact of remote work (rel. to baseline)	
	Baseline	20% remote work	100% remote work
Avg human capital	1.39	-0.44%	-2.71%
Human capital gains from internal learning	0.19	-4.73%	-28.82%
Lifecycle wage growth (25 years of experience)	0.49	-2.86%	-21.29%
Wage growth from spillover effects	0.07	-3.80%	-27.98%

Notes: This table presents the following in both the baseline model and the scenarios with 20% and 100% remote work: the average human capital level, the average human capital gains from internal learning (built by subtracting the average human capital in the scenario where internal learning is shut down from the average human capital in the baseline scenario), the average wage growth after 25 years of experience, and the average wage growth stemming from coworker spillover effects. The remote work results report the changes in the absolute values relative to the corresponding baseline results.

The surge in remote work also leads to a decrease in lifetime wage growth. Workers' average

[45] We determine the human capital gains of internal learning by subtracting the average human capital in the scenario where internal learning is shut down from the average human capital in the baseline scenario.

wage growth from human capital at 25 years of experience decreases by 2.86% due to the post-pandemic rise in remote work, and by 21.29% in the scenario where all work is performed remotely. As illustrated in Figure E.7, this drop is not only driven by lower own-productivity, but also by the decline in the compensation of high-skill workers who can mentor their colleagues. In particular, wage growth from spillover effects, as defined in Equation (10), declines by 3.80% and 27.98% in each of the two remote work scenarios, respectively.

Table 6.2: Impact of remote work on wage growth components

		Decomposition of wage main components		
	Δ log wage	Own productivity	Spillover to colleagues	Others
		Panel A: Baseline		
0–5 years of experience	0.326	0.208	0.045	0.073
5–25 years of experience	0.202	0.13	0.029	0.043
		Panel B: 20% remote work (relative to baseline)		
0–5 years of experience	-0.011	-0.004 (39%)	-0.003 (24%)	-0.004 (37%)
5–25 years of experience	-0.002	0.002 (-110%)[46]	-0.001 (37%)	-0.003 (147%)
		Panel C: 100% remote work (relative to baseline)		
0–5 years of experience	-0.066	-0.025 (37%)	-0.014 (22%)	-0.027 (41%)
5–25 years of experience	-0.027	0.001 (-3%)	-0.010 (36%)	-0.018 (67%)

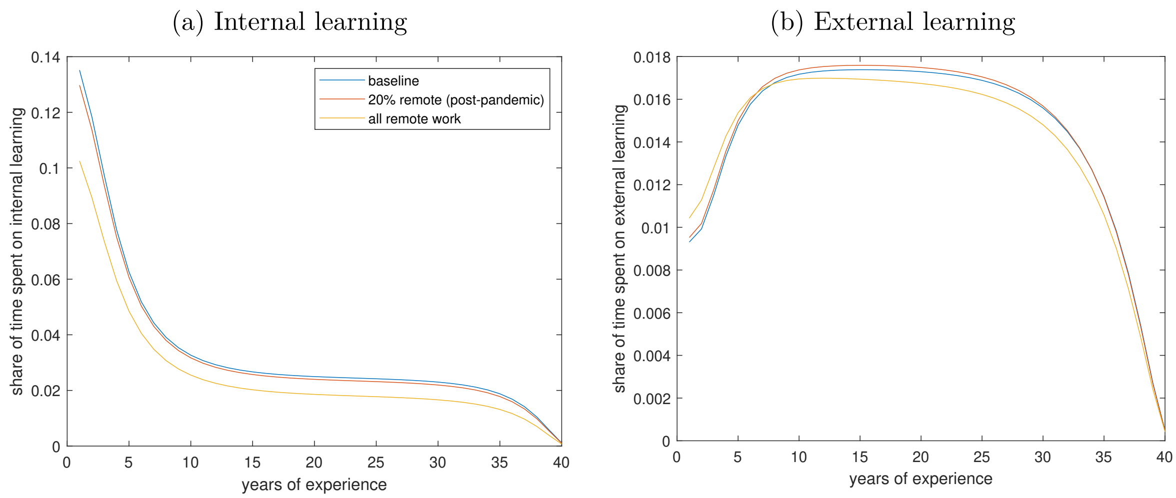
Notes: This table presents the impact of remote work on lifetime wage growth and its different components (own productivity, spillover effects to coworkers, and others) at 0–5 years of experience and 5–25 years of experience in both the baseline model (panel A) and the scenarios with 20% and 100% remote work (panels B and C, respectively). The remote work results of panels B and C report the changes in the absolute values relative to the corresponding baseline results. Δ log wage denotes the rise in average wages. The "Others" component of wage growth mainly reflects learning investments and associated costs which also change with remote work. The values in parentheses report the percentage of wage growth losses in the remote work scenarios explained by each component.

However, the relative importance of the loss of these compensated spillover effects varies throughout the lifecycle, as documented in Table 6.2, which summarizes the change in wage growth and its components in the baseline and counterfactual scenarios at different levels of experience. At younger ages (0–5 years of experience), the decline in own productivity following the lack of senior mentoring after the move to remote work accounts for a larger

[46] Despite a slow start, workers' own productivity growth at older ages is slightly faster in the case of remote work than in the baseline model. This is partly driven by substitution toward external learning and slower depreciation of human capital (due to the lower absolute level of human capital). Because the overall change in wage growth at 5–25 years of experience is small in the case of 20% remote work, the faster workers' own productivity growth becomes relatively important in this scenario.

portion (37–39%) of the decline in wage growth than changes in spillovers (22–24%). This matches the findings of Emanuel et al. (2023), who find that the decrease in feedback from coworkers after remote work particularly affects the learning prospects of young workers who were the main beneficiaries of close collaboration with more senior coworkers. At older ages (5–25 years of experience), the decline in spillover compensation following the lack of junior colleagues to teach after the move to remote work becomes more important, accounting for 36–37% of the overall decline in wage growth. This implies that remote work not only affects the wages of younger workers through learning, but also the wages of more senior workers through compensation tied to their ability to teach and mentor.[47]

Figure 6.1: Impact of remote work on lifecycle learning patterns



Notes: These figures plot the share of time spent on internal and external learning, namely $l \times g$, and $l \times (1 - g)$, respectively, for workers with different years of experience in the baseline model and 20% and 100% remote work scenarios.

Finally, we show that the decrease in internal learning after transitioning to remote work is partially offset by an increase in external learning for younger workers. In Figure 6.1 we present the shares of time allocated to internal learning and external learning in the baseline and remote work scenarios. Upon shifting to 20% or 100% of remote work, young workers slightly replace internal learning with external learning.[48] This suggests that external learning can help alleviate the negative impact of remote work on human capital as young workers can shift from internal to external learning.

[47] Another key driver of the decline in wage growth at all ages is the change in learning investments, captured under "Others." Since remote work reduces the opportunities to learn, changes in learning investments over the lifecycle are attenuated, thus reducing their contribution to lifetime wage growth.

[48] Older workers' external learning remains relatively stable in the 20% remote work scenario, but decreases in the 100% remote work scenario due to the overall deterioration of knowledge pools in the economy.

7 Model extensions and robustness

We consider several model extensions and alternative parameterizations to evaluate the robustness of our quantitative results. We present the details of each of these model extensions in Appendix F. We summarize the quantitative results of the importance of each learning source for aggregate human capital in each scenario in Table F.1, and the results of the impacts of remote work on human capital and wage growth in Table F.2.

In Appendix F.1 we integrate on-the-job search into the model since job ladders are an important driver of lifetime wage growth (Cahuc et al., 2006). We find that our main quantitative results remain very consistent, with the properties of equilibrium and impacts of internal and external learning on aggregate human capital closely matching the baseline results. However, there are two noteworthy differences in this case. First, overall wage growth and dispersion rise due to the introduction of job-to-job transitions, though the role of spillovers and learning costs in wage growth and dispersion driven by human capital is unaltered. Second, we find a larger decline in human capital in the remote work exercise relative to the baseline model. By hindering coworker learning, remote work disproportionately reduces the incentives of highly productive firms to post job vacancies since these firms have more skilled workers and thus higher intensity of coworker learning than unproductive firms. This disincentive is more severe with on-the-job search since through poaching, skilled workers tend to be more concentrated in highly productive firms than in the baseline model.

In Appendix F.2 we incorporate time costs of mentoring for the colleagues workers are learning from (internal trainers). We target the share of internal trainers' wage costs relative to learners' wage costs to calibrate this time cost, and find that our results remain very consistent. In particular, we find that even after accounting for the time costs of mentoring, 12.5% of wage growth over 25 years of experience can be attributed to spillover effects from teaching colleagues, which is similar to the baseline result (14%). We also show that the contributions of external and internal learning to aggregate human capital and the impacts of remote work are very similar to the baseline results in this model extension.

Finally, in Appendix F.3, we consider the robustness of our results to considering different values for two key parameters driving wage bargaining and human capital formation in the model, namely workers' bargaining power (β), and the elasticity of substitution between learning modes (σ). Our quantitative results remain very consistent in both cases.

8 Conclusions

In this paper, we study how different sources of on-the-job learning shape lifecycle human capital and wage dynamics. We document two novel facts that speak to the importance of internal and external sources of learning from both firms' and workers' perspectives. First, both internal and external sources of learning are widely provided by firms to their workers, and larger firms provide better learning environments by offering a greater variety of learning options. Second, both sources of learning are important to workers, and have markedly different lifecycle patterns: the prevalence of internal learning decreases with workers' experience, whereas the prevalence of external learning has an inverted U-shape in workers' experience. To shed light on the mechanisms behind these facts and quantify the importance of the two learning sources, we develop a quantitative search model featuring a two-source learning technology. In this framework, the incentives to engage in each source of skill acquisition evolve throughout the lifecycle due to aging and shifts in the relative position of the worker in the human capital distribution. Using the calibrated model, we find that internal learning is critical to early-career human capital formation, and also plays a key role in driving the lifecycle increase in wage growth and dispersion due to compensation for the learning spillovers high-skill workers trigger for their colleagues. In particular, we find that compensation stemming from coworker learning spillovers accounts for 14% of the wage growth spurred by human capital over 25 years of experience; and that the disruption to internal learning triggered by remote work not only affects the wages of younger workers through learning, but also the wages of more senior workers through compensation tied to their ability to teach and mentor. We also show that remote work can have important long-lasting spillover effects as the young workers who learn less due to remote work will be less equipped to mentor future young employees when they are more senior. However, the availability of external learning helps alleviate the negative impact of remote work on human capital as young workers can shift from internal to external learning.

Our results have several implications. First, our results suggest that the role of human capital in lifecycle wage growth is multifaceted, and thus that interpreting wage growth as primarily reflecting productivity growth may be mistaken due to the importance of non-own productivity compensation stemming from learning costs and coworker spillovers, particularly for more-experienced workers. Second, our results suggest that policies and technologies that address the communication gap created by remote work or provide alternate learning options for workers who work from home can be beneficial in both the short and long terms since they will enhance the learning of workers who will eventually become mentors themselves. Finally, our results suggest that other sources of human capital, such as schooling, may also be important for incentivizing on-the-job human capital accumulation since they improve the knowledge pool in the economy. Studying the interactions between sources of learning occurring at different stages of the lifecycle is an interesting avenue for future research.

References

Abowd, J. M., McKinney, K. L., and Zhao, N. L. (2018). Earnings Inequality and Mobility Trends in the United States: Nationally Representative Estimates from Longitudinally Linked Employer-Employee Data. *Journal of Labor Economics*, 36. 32

Acemoglu, D. (1997). Training and Innovation in an Imperfect Labour Market. *The Review of Economic Studies*, 64(3):445. ISBN: 00346527. 2, 7, 15

Acemoglu, D. and Hawkins, W. (2014). Search with Multi-worker Firms. *Theoretical Economics*, 9. 18

Acemoglu, D. and Pischke, J.-S. (1998). Why Do Firms Train? *The Quarterly Journal of Economics*, 113(1):79–119. 2, 7, 19

Acemoglu, D. and Pischke, J.-S. (1999). Beyond Becker: Training in Imperfect Labour Markets. *The Economic Journal*, 109(453):112–142. ISBN: 1468-0297. 7, 15

Akcigit, U., Caicedo, S., Miguelez, E., Stantcheva, S., and Sterzi, V. (2018). Dancing with the Stars: Innovation Through Interactions. *NBER Working Paper Series*. 2, 6, 7

Altonji, J. G. and Shakotko, R. A. (1987). Do Wages Rise with Job Seniority? *Review of Economic Studies*, 54(3):437–459. 16

Altonji, J. G., Smith, A. A., and Vidangos, I. (2013). Modeling Earnings Dynamics. *Econometrica*, 81(4). 27

Arellano-Bover, J. (2020). Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size. 4, 7, 11, 31

Arellano-Bover, J. and Saltiel, F. (2023). Differences in On-the-Job Learning Across Firms. *Working Paper*. 4, 8, 11, 31

Autor, D. H. (2001). Why do temporary help firms provide free general skills training? *Quarterly Journal of Economics*, 116(4):1409–1448. ISBN: 9780470343890. 7

Axtell, R. L. (2001). Zipf Distribution of U.S. Firm Sizes. *Science*, 293(5536):1818–1820. 26

Azoulay, P., Wang, J., and Zivin, J. S. G. (2010). Superstar Extinction. *The Quarterly Journal of Economics*, 125(2):549–589. 6, 14

Bagger, J., Fontaine, F., Postel-Vinay, F., and Robin, J. M. (2014). Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review*, 104(6):1551–1596. 2, 7, 19, 23, 25, 27, 88

Barlevy, G. (2008). Identification of search models using record statistics. *The Review of Economic Studies*, 75(1):29–64. 7

Barrero, J. M., Bloom, N., and David, S. J. (2021). Why Working from Home Will Stick. *NBER Working Paper Series*. 5, 37

Barth, E., Bryson, A., Davis, J. C., and Freeman, R. (2016). Its Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States. *Journal of Labor Economics*, 34. 32

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70(5, Part 2):9–49. 2

Becker, G. S. (1964). *Human Capital*. University of Chicago Press, Chicago. 7

Ben-Porath, Y. (1967). The Production of Human Capital Over the Life Cycle. *Journal of Political Economy*, 75(4-1):352–365. 2, 7

Benhabib, J., Perla, J., and Tonetti, C. (2021). Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier. *Econometrica*, 89(5):2261–2301. 7

Blundell, R. W., Costa Dias, M., Goll, D., and Meghir, C. (2021). Wages, Experience and Training of Women. *Journal of Labor Economics*, 39(1):275–315. 18, 28

Bowlus, A. J. and Liu, H. (2013). The contributions of search and human capital to earnings growth over the life cycle. *European Economic Review*, 64:305–331. 7

Bunzel, H., Christensen, B. J., Kiefer, N. M., and Korsholm, L. (1999). *Equilibrium search with human capital accumulation*, volume 99. Centre for Labour Market and Social Research. 7

Burdett, K., Carrillo-Tudela, C., and Coles, M. G. (2011). Human capital accumulation and labor market equilibrium. *International Economic Review*, 52(3):657–677. 7

Burstein, A. and Monge-Naranjo, A. (2007). Foreign Know-How, Firm Control, and the Income of Developing Countries. *NBER Working Paper Series*. arXiv: 1011.1669v3 ISBN: 9788578110796. 2, 96

Cahuc, P., Postel-Vinay, F., and Robin, J. M. (2006). Wage Bargaining with On-the-job Search: Theory and Evidence. *Econometrica*, 74. 15, 23, 24, 40, 81

Caicedo, S., Lucas, R. E., and Rossi-Hansberg, E. (2019). Learning, career paths, and the distribution of wages. *American Economic Journal: Macroeconomics*, 11(1):49–88. 2, 6

Card, D., Kluve, J., and Weber, A. (2018). What works? A meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3):894–931. 8

de la Croix, D., Doepke, M., and Mokyr, J. (2016). Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Pre-Industrial Economy. Technical report. 7

DiNardo, J. E. and Pischke, J. S. (1997). The returns to computer use revisited: Have pencils changed the wage structure too? *Quarterly Journal of Economics*, 112(1):291–303. 102

Dix-Carneiro, R., Goldberg, P., Meghir, C., and Ulyssea, G. (2019). Trade and Informality in the Presence of Labor Market Frictions and Regulations . 25, 26

Donovan, K., Lu, J., and Schoellman, T. (2020). Labor Market Flows and Development. 26

Emanuel, N., Harrington, E., and Pallais, A. (2023). The Power of Proximity Office Interactions Affect Online Feedback and Quits, Especially for Women and Young Workers. *NBER Working Paper Series*. 6, 7, 14, 37, 39

Engbom, N. (2017). Worker Flows and Wage Growth over the Life-Cycle: A Cross-Country Analysis. Publication Title: Mimeo. 4, 11, 31

Engbom, N. (2021). Labor market fluidity and human capital accumulation. 8, 18, 28

Engbom, N., Moser, C., and Sauermann, J. (2022). Firm Pay Dynamics. *NBER Working Paper Series*. 8

Faberman, R. J., Mueller, A. I., Sahin, A., and Topa, G. (2017). Job Search Behavior among the Employed and Non-Employed. *NBER Working Paper Series*. 84

Fitzenberger, B. and Völter, R. (2007). Long-run effects of training programs for the unemployed in East Germany. *Labour Economics*, 14(4):730–755. 8

Friedrich, B., Laun, L., Meghir, C., and Pistaferri, L. (2021). Earnings Dynamics and Firm-Level Shocks. *NBER Working Paper Series*. 8

Garicano, L. (2000). Hierarchies and the Organization of Knowledge in Production. 108(5):874–904. 2, 6

Garicano, L. and Rossi-Hansberg, E. (2004). Inequality and the Organization of Knowledge. *The American Economic Review*, 94(2):197–202. ISBN: 0002-8282. 2, 6

Garicano, L. and Rossi-Hansberg, E. (2006). Organization and Inequality in a Knowledge Economy. *Quarterly Journal of Economics*, (November):1383–1435. 2, 6

Glaeser, E. (1999). Learning in Cities. *Journal of Urban Economics*, 46(2):254–277. 7, 15

Gregory, V. (2019). Firms as Learning Environments: Implications for Earnings Dynamics and Job Search. pages 1–57. 2, 7, 11, 18, 25, 31, 88

Heckman, J. J. (1976). A Life-Cycle Model of Earnings, Learning, and Consumption. *Journal of Political Economy*, 84(4):S11–S44. 7

Heckman, J. J., Lalonde, R. J., and Smith, J. A. (1999). The economics and econometrics of active labor market programs. *Handbook of Labor Economics*, 3(1):1865–2097. 8

Herkenhoff, K., Lise, J., Menzio, G., and Phillips, G. M. (2024). Production and Learning in Teams. *Econometrica*, Forthcoming. 2, 6, 17, 23

Huggett, M., Ventura, G., and Yaron, A. (2011). Sources of lifetime inequality. *American Economic Review*, 101(7):2923–2954. ISBN: 00028282. 7

Jarosch, G. (2021). Searching for Job Security and the Consequences of Job Loss. *NBER Working Paper Series*. 8, 23

Jarosch, G., Oberfield, E., and Rossi-Hansberg, E. (2021). Learning from Coworkers. *Econometrica*, 89(2):647–676. 2, 6, 14, 17, 70

Jovanovic, B. (2014). Misallocation and growth. *American Economic Review*, 104(4):1149–1171. 7

Kambourov, G. and Manovskii, L. (2009). Occupational Specificity of Human Capital. *International Economic Review*, 50(1):63–115. 16

Karahan, F., Ozkan, S., and Song, J. (2022). Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job Ladder Risk vs. Human Capital. *FRB St. Louis Working Paper Series*. 7

Kluve, J. (2010). The effectiveness of European active labor market programs. *Labour Economics*, 17(6):904–918. Publisher: Elsevier B.V. 8

Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2018). Life Cycle Wage Growth across Countries. *Journal of Political Economy*, 126(2). 26, 78

Lazear, E. P. (2009). Firm-Specific Human Capital: A Skill-Weights Approach. *Journal of Political Economy*, 117(5):699–726. 16

Lucas, R. E. (2009). Ideas and growth. *Economica*, 76(301):1–19. 7

Lucas, R. E. and Moll, B. (2014). Knowledge Growth and the Allocation of Time. *Journal of Political Economy*, 122(1):1–51. arXiv: 1011.1669v3 ISBN: 0022-3808. 7

Luttmer, E. G. J. (2014). An Assignment Model of Knowledge Diffusion and Income Inequality. 2, 6, 96

Ma, X., Nakab, A., and Vidart, D. (2020). Human Capital Investment and Development: The Role of On-the-job Training. 2, 12, 15

Manuelli, R. E. and Seshadri, A. (2014). Human capital and the wealth of nations. *American Economic Review*, 104(9):2736–2762. 2, 28

Manuti, A., Pastore, S., Scardigno, A. F., Giancaspro, M. L., and Morciano, D. (2015). Formal and informal learning in the workplace: a research review: Formal and informal learning in the workplace. *International Journal of Training and Development*, 19(1):1–17. 16

McKenzie, D. (2017). How effective are active labor market policies in developing countries? A critical review of recent evidence. *World Bank Research Observer*, 32(2):127–154. 8

Mincer, J. (1989). Human Capital and the Labor Market: A Review of Current Research. *Educational Researcher*, 18(1). 18

Moen, E. and Rosén, A. (2004). Does Poaching Distort Training? *The Review of Economic Studies*, 71(4):1143–1162. 15

Moen, E. R. and Rosén, . (2004). Does poaching distort training? *Review of Economic Studies*, 71(4):1143–1162. 2, 7

Monge-Naranjo, A. (2019). Markets, Externalities, and the Dynamic Gains of Openness. *International Economic Review*, 60(3):1131–1170. 18

Nix, E. (2017). Learning Spillovers in the Firm. 2, 6, 14

Perla, J. and Tonetti, C. (2014). Equilibrium Imitation and Growth. *Journal of Political Economy*, 122(1):1–21. ISBN: 00223808. 7

Rosen, S. (1972). Learning and experience in the labor market. *The Journal of Human Resources*, 7(3):326–342. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System]. 18

Rosen, S. (1976). A Theory of Life Earnings. *Journal of Political Economy*, 84(4):S45–S67. 7

Rubinstein, Y. and Weiss, Y. (2006). Chapter 1 Post Schooling Wage Growth: Investment, Search and Learning. *Handbook of the Economics of Education*, 1:1–67. ISBN: 9780444513991. 2, 7, 30

Shimer, R. (2005). The Cyclical Behavior of Equilibrium Unemployment and Vacancies. *American Economic Review*, 95(1):25–49. 24, 25, 26

Shimer, R. (2012). Reassessing the Ins and Outs of Unemployment. *Review of Economic Dynamics*, 15(2). 28

Sohail, F. (2021). From employee to entrepreneur: Learning, employer size, and spinout dynamics. *Journal of Economic Dynamics and Control*, 133:104270. 7

Song, J., Price, D., Guvenen, F., Bloom, N., and Wachter, T. V. (2019). Firming Up Inequality. *Quarterly Journal of Economics*. 32

Wallskog, M. (2023). Entrepreneurial Spillovers Across Coworkers. 7

What Works - Centre for Local Economic Growth (2016). Evidence Review 1: Employment Training. Technical report, London School of Economics. 8

Yamaguchi, S. (2010). Job search, bargaining, and wage dynamics. *Journal of Labor Economics*, 28(3):595–631. 7

A Appendix: Data

A.1 EU Continuing Vocational Training Survey

The European Union's Continuing Vocational Training Enterprise Survey (EU-CVT) collects information from enterprises (firms) in the European Union, and focuses on their investments in continuing vocational training of their staff. The survey provides information on the types, content, costs, and volume of continuing training, and firms' use of own and external training providers and resources. CVT surveys have been carried out for the reference years 1993, 1999, 2005, 2010 and 2015. However, due to data availability, we rely on the three waves in 2005, 2010, and 2015, labeled as CVT3, CVT4, and CVT5. These three surveys provide a sample of 78,000, 101,000, and 111,000 firms, respectively, from all EU member states and Norway.

A.1.1 Internal and external learning

To build our measures of internal and external learning, we combine information measuring the provision of external, internal, and other types of CVT activities from the EU-CVT survey manuals. In particular, firms are considered to offer internal learning if they offer either "internal CVT courses" or "other forms of CVT" that draw on the internal knowledge pool. Similarly, firms are considered to offer external learning if they offer "external CVT courses" or "other forms of CVT" that draw on external knowledge. We now explore each of these categories closely.

- CVT courses refer to education or training activities that are planned in advance, organized, or supported with the specific goal of learning, and are financed at least partially by the enterprise. These activities aim to generate "the acquisition of new competences or the development and improvement of existing ones" for firms' employees. These courses are typically separated from the active workplace (for example, they take place in a classroom or at a training institution), show a high degree of organization, and include content explictly designed for a group of learners (e.g., a curriculum exists). We consider two types of CVT courses:
  - Internal CVT courses: Courses, seminars or activities that take place inside firms and employ internal trainers.
  - External CVT courses: Courses, seminars or activities that take place outside firms or employ external trainers.

- Other forms of CVT are geared toward learning and are typically connected to active work and the active workplace, but can also include participation in conferences, trade fairs, etc. These activities are often characterized by self-organization by the individual learner or by a group of learners and are typically tailored to the workers' needs. Specifically, we use the following types of "other forms of CVT:"[49]

  - Guided on-the-job training: "It is characterized by planned periods of training, instruction or practical experience in the workplace using the normal tools of work, either at the immediate place of work or in the work situation. The training is organized (or initiated) by the employer. A tutor or instructor is present. It is an individual-based activity, i.e. it takes place in small groups only (up to five participants)." This is categorized as internal learning.

  - Job rotation, exchanges, secondments, or study visits: "Job rotation within the enterprise and exchanges with other enterprises as well as secondments and study visits are other forms of CVT only if these measures are planned in advance with the primary intention of developing the skills of the workers involved. Transfers of workers from one job to another which are not part of a planned developmental programme should be excluded." This is categorized as internal learning.

  - Learning or quality circles: "Learning circles are groups of persons employed who come together on a regular basis with the primary aim of learning more about the requirements of the work organisation, work procedures and workplaces. Quality circles are working groups, having the objective of solving production and workplace-based problems through discussion. They are counted as other forms of CVT only if the primary aim of the persons employed who participate is learning." This is categorized as internal learning.

  - Participation in conferences, workshops, trade fairs, and lectures: "Participation (instruction received) in conferences, workshops, trade fairs and lectures are considered as training actions only when they are planned in advance and if the primary intention of the person employed for participating is training/learning." This is categorized as external learning.

[49] Among the "other types of CVT activities," the survey also contemplates self-directed learning, which is more akin to learning-by-doing and thus not considered here.

A.1.2 Summary statistics in EU-CVT data

Table A.1: Share of firms providing CVT courses and other types of CVT activities in EU-CVT

	CVT courses		Other types of CVT activities			
Country	Internal CVT courses	External CVT courses	Conferences, workshops or lectures	Guided on-the-job training	Job rotation	Learning and quality circles
Germany	0.436	0.532	0.438	0.524	0.085	0.151
France	0.329	0.666	0.193	0.253	0.094	0.082
United Kingdom	0.418	0.532	0.399	0.655	0.204	0.218
Italy	0.263	0.403	0.263	0.260	0.100	0.035
Spain	0.186	0.564	0.198	0.354	0.108	0.127
Poland	0.134	0.219	0.123	0.189	0.054	0.022
Romania	0.117	0.157	0.084	0.139	0.069	0.051
Belgium	0.457	0.605	0.364	0.437	0.159	0.161
Portugal	0.212	0.360	0.216	0.389	0.064	0.099
Czech Rep.	0.391	0.611	0.243	0.367	0.048	0.088
Hungary	0.171	0.307	0.211	0.198	0.034	0.057
Sweden	0.600	0.724	0.499	0.611	0.355	0.103
Bulgaria	0.163	0.177	0.130	0.227	0.057	0.087
Denmark	0.485	0.640	0.513	0.452	0.157	0.189
Slovak Rep.	0.352	0.531	0.427	0.326	0.091	0.188
Finland	0.352	0.671	0.346	0.391	0.119	0.119
Norway	0.676	0.694	0.481	0.704	0.327	0.221
Latvia	0.124	0.271	0.145	0.475	0.059	0.056
Estonia	0.326	0.552	0.279	0.439	0.163	0.103
Cyprus	0.185	0.447	0.266	0.328	0.089	0.166
Luxembourg	0.473	0.580	0.390	0.444	0.170	0.197
Malta	0.284	0.322	0.320	0.435	0.142	0.136
Total	0.275	0.433	0.242	0.328	0.101	0.091

Notes: This table presents the share of firms in which workers participate in different types of CVT courses and activities for each country in the EU-CVT data. The results are the simple averages of the respective proportions from three different CVT survey waves: CVT3, CVT4 and CVT5. The “Total” in the last row is an average for all waves and all countries sampled. The sample encompasses firms with 5 or more workers, since smaller firms encompass a very small portion of the sample. The results are weighted using the weighting factors provided in the surveys.

A.2 US Survey of Employer Provided Training

The 1995 Survey of Employer-Provided Training (US-SEPT) was conducted by the Bureau of Labor Statistics (BLS), and collected information on training from firms and randomly selected employees in establishments with 50 or more workers in the United States.[50] The employer portion of the survey focuses on the intensity and costs of employer-provided formal training. The employee portion of the survey focuses on the time that employees spent on both formal and informal training. This survey provides a sample of 1,062 establishments and over 1,000 employees covering all nine major industry classifications across all 50 states.

The micro-level data for this survey is not available for researchers outside the BLS. In addition, information about both informal and formal training, which we map into internal and external learning, is only available in the employee portion of the survey. Thus, we rely on aggregate statistics on the share of employees exposed to different types of training with the current employer, along with the number of hours that employees spent on these different sources of learning across firms of different sizes to support the results in Section 3.2. We also use this data for the calibration of our model in Section 4.

A.2.1 Internal and external learning

To build our measures of external and internal learning in this data, we rely on information about formal and informal training investments, respectively, which roughly, but not perfectly, map onto our two learning sources of interest. We now describe each of these two forms of training in detail:

- Formal training is defined in the survey as training that is planned in advance and has a structured format and defined curriculum. Examples include attending a class conducted by an employee of the company, attending a seminar given by a professional trainer, or watching a planned audio-visual presentation.
- Informal training is defined in the survey as training that is unstructured, unplanned, and easily adapted to situations or individuals. Examples include having a coworker show a colleague how to use a piece of equipment or having a supervisor teach a skill related to the job.[51]

[50] The BLS also conducted a similar survey in 1994, which focused on the existence and types of formal training programs provided or financed by establishments. Due to data availability, we focus only on the 1995 wave of the survey.

[51] Please note that in the survey informal training is different from self-learning activities such as learning-

A.3 German Worker Qualification Survey

The Worker Qualification Survey is conducted by the BIBB (Bundesinstitut fr Berufsbildung, Bonn), a federal agency devoted to vocational education, in conjunction with the IAB (until 1999) and BAuA (after 1999). These BIBB/IAB/BAuA surveys span 7 waves conducted in 1979, 1985, 1992, 1999, 2006, 2012, and 2018, and cover a representative sample of 20,000 to 35,000 members of the labor force (excluding apprentices) in each wave. The survey provides comprehensive data to analyze both the cross-sectional and temporal evolution of the qualifications and working conditions of the German workforce (excluding apprentices).

This data has two important limitations. First, there is variation in the questions asked across survey waves. This partially compromises the comparability of our measures of skill acquisition across waves, and thus the longitudinal nature of the data. However, these changes do not matter for the aggregate lifecycle patterns of on-the-job learning if the age distribution of respondents is constant across waves. More importantly, Table C.1 indicates that the results we find are robust to controlling for survey wave fixed effects, while Figure C.5 shows the results are robust to considering each wave separately. Second, the response rate to the survey is relatively low, reaching levels as low as 44%. To address this issue, we adjust all of our results using the weighting schemes provided by BIBB to account for the selection probabilities of both households and targeted persons caused by the sample design and the selective failures due to refusals.

A.3.1 Internal learning, external learning, and experience

The questions regarding human capital accumulation change considerably throughout survey waves in this data. Therefore, to construct the variables that capture whether the worker engages in internal or external learning, different questions (and variables) have to be used as indicators. In this section we provide the guidelines used for the construction of the skill acquisition variables of interest, along with the measures of experience.

- Internal learning is constructed as a binary variable that indicates whether an individual has acquired the skills/knowledge necessary to complete the tasks required in their current job through colleagues or superiors. This question remains relatively stable throughout the surveys, except for (1) the 1979 survey, which did not distinguish between learning-by-doing and internal learning (and is thus excluded); (2) the 2006

by-doing and learning through hobbies.

survey, which asked about having received professional development through coaching from superiors; and (3) the 2011/2012 and 2017/2018 surveys, when no related question was asked. It is also important to note that the skill acquisition questions in the 1979–1999 surveys had a slight change after the 1986 survey. In the 1979 and 1986 surveys, individuals could list all the sources through which they acquired the skills needed for their current jobs, whereas in 1992 and 1999, they only listed the two main ones.

- External learning is constructed as a binary variable that indicates whether an individual received external on-the-job training in the previous 2–5 years, or acquired the skills/knowledge necessary to complete the tasks in their current job through external training.
  - 1979, 1985/1986: For these two waves, external learning corresponds to (1) reporting that the sources of professional knowledge/skills for the job include on-the-job training or continued training; and/or (2) attending any courses with the purpose of training in the 5 years that preceded the survey.
  - 1991/1992, 1998/1999: For these two waves, external learning corresponds to (1) reporting that the two main sources of professional knowledge/skills for the job are on-the-job training or continued training; and/or (2) attending any courses with the purpose of training in the 5 years that preceded the survey, specifically: visiting trade fairs, congresses, or technical lectures; instruction by external agents, or reading circles at the workplace; and reading of trade journals, or specialist literature.
  - 2005/2006: For this wave external learning corresponds to (1) attending any courses with the purpose of training in the 2 years that preceded the survey, specifically: visiting trade fairs, congresses, or technical lectures; instruction by external agents, or reading circles at the workplace; reading of trade journals, or specialist literature; and learning from computer-based or internet sources; and/or (2) claiming it is important to attend seminars or courses to perform one's occupational activity.
  - 2011/2012, 2017/2018: For these waves, external learning corresponds (1) attending any courses with the purpose of training in the 2 years that preceded the survey (no specific types); and/or (2) claiming it is important to attend seminars

or courses to perform one's occupational activity.

- Potential experience is constructed as: $Age - Years\ of\ schooling - 6$
  - $Years\ of\ schooling$ is constructed using: $Year\ of\ graduation\ of\ highest\ degree - Birth\ year - 6$.
- Number of years with current employer is constructed directly from the corresponding variable in the survey for the years 1979, 1985/1986 and 1991/1992, and from $Current\ year - Year\ start\ with\ current\ employer$ for 1998/1999, 2005/2006, 2011/2012 and 2017/2018. Promotions within the same company are not considered employer switches. For self-employed workers or business owners, this variable captures the years since the start of running the current business or occupation.

A.3.2 Other variables

- Hourly wages are constructed using data on the monthly wage and regular hours. In the first few surveys, monthly wage was captured in an ordinal fashion, with interviewees picking among different wage ranges. In more recent surveys, the answer is given in exact numbers. Thus, individual wages in the early waves are imputed by using the mid-point of the reported wage range. Wages are deflated using the German CPI with the base year 2015 and currency is adjusted to account for the change to the Euro.

A.4 US Adult Training and Education Module in National Household Education Survey

The National Household Education Survey (NHES) has been deployed in 1993, 1995, 1996, 1999, 2001, 2003, 2005, 2007, 2012 and 2016. The Adult Training and Education (ATES) module was not included in every survey, however, and limited to 1991, 1995, 1999, 2001, 2003, 2005, and 2016. Moreover, information on internal learning was only first included in the 2016 wave. The ATES in 2016 was collected via telephone surveys. The survey focuses on non-institutionalized adults (aged 16–65) who are not enrolled in grade 12 or below and comprises 47,744 individuals who are representative of the US population at large.

A.4.1 Internal learning, external learning, and experience

In this section, we provide further information about the construction of the skill acquisition variables of interest, along with the measure of potential experience.

- Internal learning is constructed as a binary variable that takes a value of one for workers who reported receiving instruction or training from a coworker or supervisor in their last work-experience program, and a value of zero for all other workers surveyed. As such, both workers who reported participating in a work-experience program but did not receive instruction from coworkers or supervisors, and workers who did not report having recently participated in a work-experience program are assumed to not have this source of learning. This follows from the definition of work-experience program, which is defined as a job with learning attributes, such as an internship, co-op, practicum, clerkship, externship, residency, clinical experience, apprenticeship, or other learning components.[52]

- External learning is constructed as a binary variable that takes a value of one for workers who either reported taking classes or training from a company, association, union, or private instructor in their last work-experience program; or ever earned a training certificate from an employment-related training program. The variable takes a value of zero for all other workers. Therefore, workers who failed to report either participating in a work-experience program and receiving training, having recently participated in a work-experience program, or ever receiving an employment-related training certificate are assumed to not have this source of learning.[53]

- Potential experience is constructed as: $Age - Years\ of\ schooling - 6$
  - $Years\ of\ schooling$ is constructed by mapping the educational attainment to the corresponding years of schooling. We omit workers with an educational attainment of less than secondary since we cannot directly map this into years of schooling.

A.4.2 Other variables

- Hourly wages are constructed using data on yearly work earnings, weeks worked, and regular hours. Yearly work earnings and weeks worked are answered in an ordinal

[52]In panel (a) of Figure C.8, we show that our documented lifecycle pattern of internal learning is robust if we only consider individuals reporting participating in a work-experience program.

[53]In panels (c) and (d) of Figure C.8, we show that our documented lifecycle pattern of external learning is robust to decomposing across the two learning components of external learning: taking classes during a last work-experience program, or ever receiving an employment-related training certificate. In panel (b) of Figure C.8, we show that our documented lifecycle pattern of external learning is robust if we only consider individuals reporting participating in a work-experience program.

fashion. Thus, yearly wage earnings and weeks worked are imputed by the mid-point of the reported range.

A.4.3 Summary statistics in German BIBB/BAuA and US NHES data

Table A.2: Summary statistics in German BIBB/BAuA and US NHES data

Germany					
	Mean	Std. dev.	Minimum	Maximum	# Obs.
Reports internal learning	0.31	0.46	0	1	109478
Reports external learning	0.68	0.47	0	1	173391
Woman	0.42	0.49	0	1	174647
Age	40.15	11.17	15	74	174647
Years of education	10.75	2.67	0	25	174647
Potential years of experience	23.40	11.59	1	45	174647
Years with current employer	11.34	9.91	0	70	166964
Hourly wage (Euros of 2015)	8.96	9.07	0	207.15	117293
Firm size: 1–9 workers	0.23	0.42	0	1	165770
Firm size: 10–99 workers	0.37	0.48	0	1	165770
Firm size: 100+ workers	0.40	0.49	0	1	165770
United States					
	Mean	Std. dev.	Minimum	Maximum	# Obs.
Reports internal learning	0.23	0.42	0	1	29399
Reports external learning	0.44	0.5	0	1	29399
Woman	0.52	0.5	0	1	29399
Age	41.03	12.48	16	66	29399
Years of education	14.58	2.12	12	20	29399
Potential years of experience	20.72	12.49	1	45	29217
Hourly wage (Dollars of 2016)	27.34	37.93	1.23	2307.69	27767

Notes: This table presents summary statistics for key variables in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have between 1 and 45 years of potential experience in both settings. The results are weighted using the observational weights provided in the surveys.

A.5 OECD Program for the International Assessment of Adult Competencies (PIAAC)

The Program for the International Assessment of Adult Competencies (PIAAC) is an international survey conducted by the Organization for Economic Cooperation and Development (OECD). The survey aims to assess and compare the learning environments, skills, and competencies of adults aged 16 to 65 in more than 40 OECD countries. 24 countries participated in round 1 of the survey, which collected data from 1 August 2011 to 31 March 2012. Round 2 of the assessment included 9 participating countries, with data collection taking place from April 2014 to the end of March 2015. Finally, round 3 included participation from 6 countries, with data collection taking place from July to December 2017. In total, this survey covers a sample of around 230,000 individuals.

PIAAC collects information about workers' learning investments in skills, along with information on how adults utilize these skills in various settings, namely home, work, and the wider community. In addition, PIAAC measures workers' proficiency in three key domains: literacy, numeracy, and problem-solving in technology-rich environments. In every country, PIAAC provides methodological documents and guidelines to facilitate the proper collection of data and ensure harmony in the definitions and concepts across countries.

This survey contains rich information allowing us to construct measures of internal and external learning for workers in participating countries, along with potential years of experience and a measure of realized years of experience. However, this data has some important limitations. First, the sample of individuals interviewed in each country is relatively small, with the number of observations suitable for our analysis ranging from about 700 to 2,500 in each country. This is further complicated when we divide the samples across experience bins. Second, several important variables are excluded in some countries due to privacy laws, further limiting our analysis. For example, the US excludes data on workers' age and years of experience from the publicly available files, which precludes the construction of both the realized and potential years of experience variables, and therefore restricts our use of data from this source.

As before, we limit our sample to individuals who are currently employed and have 1–45 years of both potential and realized experience. In addition, we exclude military personnel from our data. After these refinements and data construction for our main variables of interest (internal and external learning and realized and potential years of experience), the countries included in our analysis from each round of the survey encompass:

- Round 1 (2011–2012): Czech Republic, Denmark, Estonia, Finland, France, Ireland, Italy, Japan, Korea, Netherlands, Norway, Poland, Russia, Slovakia, Spain, Sweden, and the UK.
- Round 2 (2014–2015): Chile, Greece, Israel, Lithuania, Slovenia, and Turkey.
- Round 3 (2017): Ecuador, Kazakhstan, Mexico, and Peru.

A.5.1 Internal learning, external learning, and experience

In this section, we provide further information about the construction of the skill acquisition variables of interest, along with the measures of potential and realized years of experience.

- Internal learning is constructed as is a binary variable that takes a value of one for workers who either reported participating in courses led by coworkers or supervisors in the last 12 months, or reported learning skills from coworkers more than once a month in their current work.
- External learning is constructed as a binary variable that takes a value of one for workers who reported participating in seminars, workshops, private lessons or other courses led by outsiders in the last 12 months.
- Realized years of experience captures the number of years of paid work up to date of measurement. Only years when workers spent 6 months or more on full- or part-time work are included.
- Potential experience is constructed as: $Age - Years\ of\ schooling - 6$
  - $Years\ of\ schooling$ is constructed by mapping the educational attainment to the corresponding years of schooling (this mapping is directly constructed and provided by PIAAC). We omit workers with an educational attainment of less than secondary, since we cannot directly map this into years of schooling.

A.5.2 Other variables

- Hourly wages are constructed using data on monthly work earnings (including bonuses) for wage, salary, and self-employed workers, along with data on hours worked last week. Wages are in US dollars of the year of the survey in each country, and PPP corrected.

A.5.3 Summary statistics in PIAAC data

Table A.3: Summary statistics in PIAAC data

	Mean	Std. dev.	Minimum	Maximum	# Obs.
Reports internal learning	0.75	0.43	0	1	55112
Reports external learning	0.31	0.46	0	1	60244
Woman	0.46	0.5	0	1	60274
Age	39.87	11.06	19	65	60275
Years of education	14.48	2.41	12	23	60275
Realized years of experience	16.56	10.81	1	45	60275
Potential years of experience	19.4	11.25	1	45	60275
Hourly wage (Dollars of survey year, PPP)	18.76	156.24	0	55411.45	49624
Firm size: 1–10 workers	0.3	0.46	0	1	53620
Firm size: 11–250 workers	0.52	0.5	0	1	53620
Firm size: 251+ workers	0.18	0.39	0	1	53620

Notes: This table presents summary statistics for key variables in the OECD PIAAC data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of both potential and realized experience. The results are weighted using the observational weights provided in the survey.

B Appendix: Robustness of Fact 1

B.1 EU-CVT data

Table B.1: Share of firms providing internal and external learning activities

		External learning	
		0	1
Internal learning	0	0.33	0.15
	1	0.11	0.41

Notes: This table presents the proportion of firms reporting having employees participating in internal and external learning activities in the EU-CVT data (CVT3, CVT4 and CVT5 surveys). The sample encompasses firms with 5 or more workers, since smaller firms encompass a very small portion of the sample. The results are weighted using the weighting factors provided in the surveys.

Table B.2: Correlation between different types of CVT provision and firm size

	CVT			Internal CVT			External CVT		
Firm size: 20–49 wks.	0.116***	0.117***	0.117***	0.110***	0.109***	0.109***	0.118***	0.119***	0.121***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Firm size: 50–99 wks.	0.200***	0.199***	0.199***	0.205***	0.200***	0.200***	0.227***	0.227***	0.228***
	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Firm size: 100–250 wks.	0.282***	0.280***	0.280***	0.312***	0.304***	0.304***	0.324***	0.324***	0.325***
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Firm size: 251+ wks.	0.297***	0.293***	0.293***	0.381***	0.372***	0.372***	0.352***	0.349***	0.350***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Observations	286,321	286,321	286,321	286,321	286,321	286,321	286,321	286,321	286,321
R-squared	0.171	0.179	0.179	0.154	0.160	0.160	0.157	0.168	0.170
Country-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE		Y	Y		Y	Y		Y	Y
Socioec. controls			Y			Y			Y

Notes: This table shows the coefficients from regressing a variable indicating whether firms in the EU-CVT data (CVT3, CVT4 and CVT5 surveys) report having employees participating in any kind of CVT, internal CVT, or external CVT activities on different firm size categories, where the omitted category encompasses firms with 5–19 workers. The sample encompasses firms with 5 or more workers, since smaller firms encompass a very small portion of the sample. Regressions are weighted using the using the observational weights provided in the surveys. Year fixed effects correspond to the year of the CVT survey. Industry categories are at the 1-digit level (NACE). Socioeconomic controls encompass the logarithm of per-capita GDP. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3: Average number and share of hours spent by each worker on CVT courses by firm size

	CVT courses	Internal CVT courses	External CVT courses
	Panel A: Average share of working hours		
Small firms, 5–19 workers	0.0027	0.0011	0.0016
Medium firms, 20–99 workers	0.0037	0.0015	0.0022
Large firms, 100+ workers	0.0058	0.0028	0.0030
Total	0.0040	0.0018	0.0023
	Panel B: Average number of hours per worker		
Small firms, 5–19 workers	4.861	2.006	2.855
Medium firms, 20–99 workers	6.077	2.451	3.626
Large firms, 100+ workers	9.491	4.505	4.987
Total	6.550	2.834	3.716

Notes: This table shows the average share and number of hours spent by each worker in any kind of CVT courses, internal CVT courses, and external CVT courses in the last calendar year in firms of different sizes in the EU-CVT data. Shares and numbers of hours are calculated for all firms in each size category, and are presented in panels A and B, respectively. Results are simple averages of the respective calculations across firms from three different CVT survey waves: CVT3, CVT4 and CVT5. The sample encompasses firms with 5 or more workers, since smaller firms encompass a very small portion of the sample. The results are weighted using the weighting factors provided in the surveys.

Table B.4: Correlation between hours spent by each worker on CVT courses and firm size

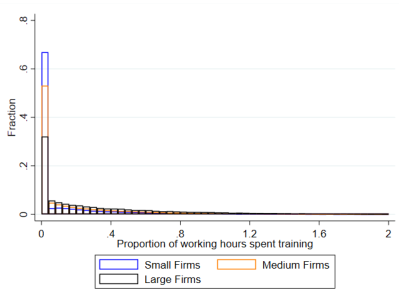
	CVT courses			Internal CVT courses			External CVT courses		
Firm size: 20–49	0.703**	0.737**	0.759**	0.432*	0.424*	0.429*	0.270*	0.313*	0.330**
	(0.314)	(0.316)	(0.317)	(0.247)	(0.247)	(0.249)	(0.164)	(0.165)	(0.165)
Firm size: 50–99	1.476***	1.432***	1.450***	0.773***	0.710***	0.714***	0.703***	0.723***	0.736***
	(0.280)	(0.284)	(0.284)	(0.192)	(0.193)	(0.193)	(0.169)	(0.173)	(0.172)
Firm size: 100–250	3.111***	3.034***	3.047***	1.704***	1.605***	1.608***	1.407***	1.429***	1.439***
	(0.319)	(0.321)	(0.321)	(0.204)	(0.204)	(0.204)	(0.200)	(0.202)	(0.202)
Firm size: 251+	3.615***	3.427***	3.441***	2.567***	2.413***	2.417***	1.048***	1.013***	1.024***
	(0.266)	(0.285)	(0.283)	(0.176)	(0.182)	(0.181)	(0.149)	(0.163)	(0.162)
Observations	273,870	273,870	273,870	273,870	273,870	273,870	273,870	273,870	273,870
R-squared	0.015	0.021	0.022	0.008	0.010	0.010	0.016	0.023	0.023
Country-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE		Y	Y		Y	Y		Y	Y
Socioec. controls			Y			Y			Y

Notes: This table shows the coefficients from regressing the number of hours spent by each worker on any kind of CVT courses, internal CVT courses, and external CVT courses in the last calendar year in the EU-CVT data (CVT3, CVT4 and CVT5 surveys) on different firm size categories, where the omitted category encompasses firms with 5–19 workers. The sample encompasses firms with 5 or more workers, since smaller firms encompass a very small portion of the sample. Regressions are weighted using the using the observational weights provided in the surveys. Year fixed effects correspond to the year of the CVT survey. Industry categories are at the 1-digit level (NACE). Socioeconomic controls encompass the logarithm of per-capita GDP. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

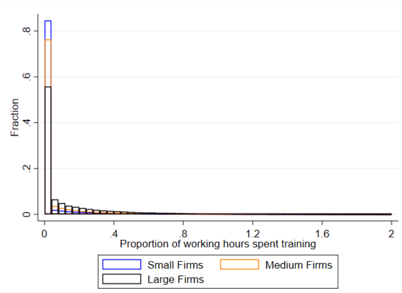
Figure B.1: Histograms of share of hours spent on CVT courses by firm size

All firms

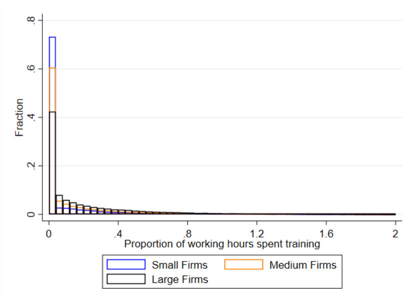
(a) CVT courses



(b) Internal CVT courses

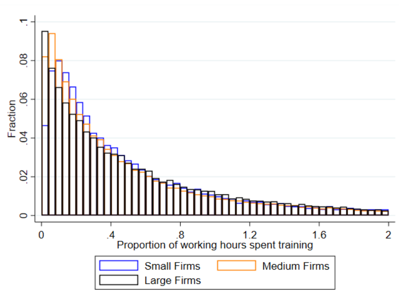


(c) External CVT courses

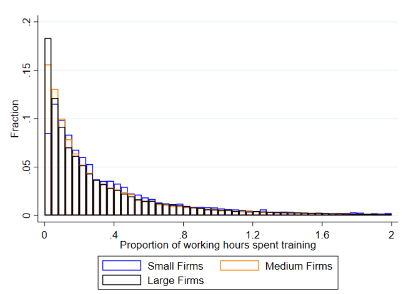


All firms, with zero share of learning hours not plotted

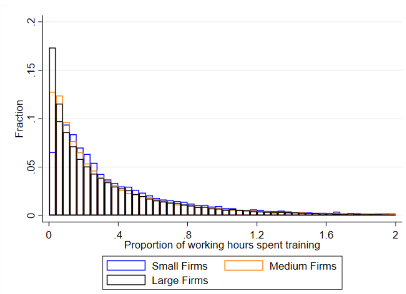
(d) CVT courses



(e) Internal CVT courses



(f) External CVT courses



Notes: These figures plot the histograms of the share of working hours spent by each worker on any kind of CVT courses, internal CVT courses, and external CVT courses by firm size categories for all firms in the EU-CVT data (CVT3, CVT4 and CVT5 surveys), and for all firms in the EU-CVT data (CVT3, CVT4 and CVT5 surveys) but with zero learning hours not plotted. The sample encompasses firms with 5 or more workers, since smaller firms encompass a very small portion of the sample. The results are weighted using the weighting factors provided in the surveys.

B.2 US-SEPT Data

Table B.5: Share of workers trained and average number of training hours by firm size

	Informal training	Formal training
	Panel A: Share of workers who have received training with current employer	
Small firms, 50–99 workers	78.9	97.1
Medium firms, 100–499 workers	84.7	95.0
Large firms, 500+ workers	87.7	96.1
Total	84.4	95.8
	Panel B: Average number of hours per worker	
Small firms, 50–99 workers	8.2	31.9
Medium firms, 100–499 workers	13.5	34.5
Large firms, 500+ workers	16.6	26.0
Total	13.4	31.1

Notes: This table shows the share of workers who ever received formal or informal training with the current employer, along with the average number of hours spent by each worker in these activities in the last 6 months in the 1995 US-SEPT data. Shares of workers and numbers of hours are calculated for all firms in each size category, and are presented in panels A and B, respectively. Results follow from aggregate statistics presented by the BLS.

C Appendix: Robustness of Fact 2

C.1 German BIBB/BAuA and US NHES data

Table C.1: Correlations between internal and external learning and potential experience

	Germany					
	Internal learning			External learning		
Potential years of experience	-0.0104***	-0.00630***	-0.00296***	0.0105***	0.00536***	0.00212***
	(0.000672)	(0.000781)	(0.00103)	(0.000552)	(0.000600)	(0.000722)
Potential years of experience^2	0.000167***	8.37e-05***	4.66e-05**	-0.000247***	-0.000126***	-7.69e-05***
	(1.37e-05)	(1.61e-05)	(2.16e-05)	(1.14e-05)	(1.23e-05)	(1.49e-05)
Constant	0.439***	0.337***	0.359***	0.602***	0.674***	0.678***
	(0.00734)	(0.0141)	(0.0178)	(0.00592)	(0.0109)	(0.0138)
Observations	109,478	69,495	36,813	173,391	126,129	85,280
R-squared	0.006	0.129	0.077	0.005	0.193	0.205
Year FE		Y	Y		Y	Y
Demographic controls		Y	Y		Y	Y
Worker type FE		Y	Y		Y	Y
Industry FE		Y	Y		Y	Y
Occupation FE		Y	Y		Y	Y
Firm size FE		Y	Y		Y	Y
Wage controls			Y			Y

	United States					
	Internal learning			External learning		
Potential years of experience	-0.00489***	-0.00943***	-0.00999***	0.0153***	0.0110***	0.0114***
	(0.00120)	(0.00103)	(0.00105)	(0.00141)	(0.00139)	(0.00142)
Potential years of experience^2	-4.13e-05*	9.76e-05***	0.000110***	-0.000277***	-0.000180***	-0.000188***
	(2.49e-05)	(2.14e-05)	(2.17e-05)	(3.05e-05)	(3.02e-05)	(3.08e-05)
Constant	0.352***	0.270***	0.262***	0.289***	0.278***	0.276***
	(0.0119)	(0.0141)	(0.0146)	(0.0129)	(0.0176)	(0.0180)
Observations	29,217	29,217	27,585	29,217	29,217	27,585
R-squared	0.040	0.228	0.224	0.013	0.073	0.075
Demographic controls		Y	Y		Y	Y
Worker type FE		Y	Y		Y	Y
Industry FE		Y	Y		Y	Y
Occupation FE		Y	Y		Y	Y
Wage controls			Y			Y

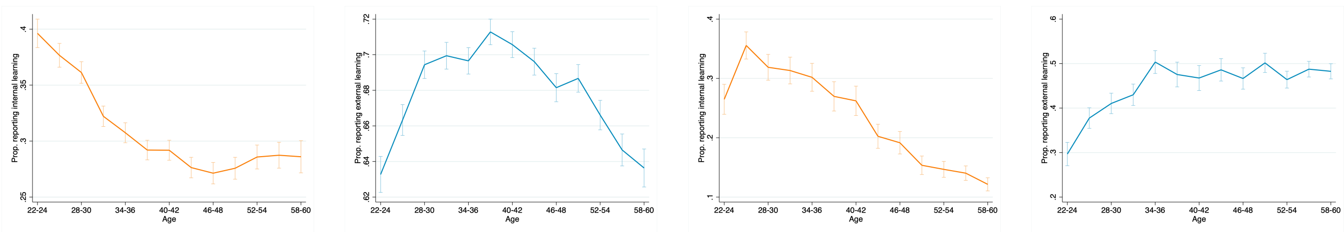
Notes: This table shows the coefficients from regressing internal and external learning on potential experience and potential experience squared in the German BIBB/BAuA data and US NHES data. The samples encompass individuals who are currently employed, and have between 1 and 45 years of potential experience. Regressions are weighted using the using the observational weights provided in the surveys. *Germany:* Year fixed effects correspond to the year of the survey. Demographic controls include educational attainment level and gender. Worker type categories include laborer, private employee, government employee, self-employed, freelancer, or family caregiver. Industry categories are at the 1-digit level. Occupation categories are at the 2-digit level (ISCO 88). Firm size is a categorical variable indicating whether the firm where the worker works has fewer than 4 workers, 5–9 workers, 10–49 workers, 50–99 workers, 100–499 workers, 500–999 workers, and 1000 or more workers. Wage controls include the current hourly wage. *US:* Demographic controls include educational attainment level, race, census region, and gender. Worker type categories include private employee, government employee, self-employed, or working without pay. Industry and occupation categories are at the 2-digit level (ACS 2015). Wage controls include the current hourly wage. We do not include age fixed effects due to high collinearity between potential experience, education, and age. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C.1.1 Decomposition and alternate specifications

Figure C.1: Sources of learning by age

Germany **United States**

(a) Internal learning (b) External learning (c) Internal learning (d) External learning

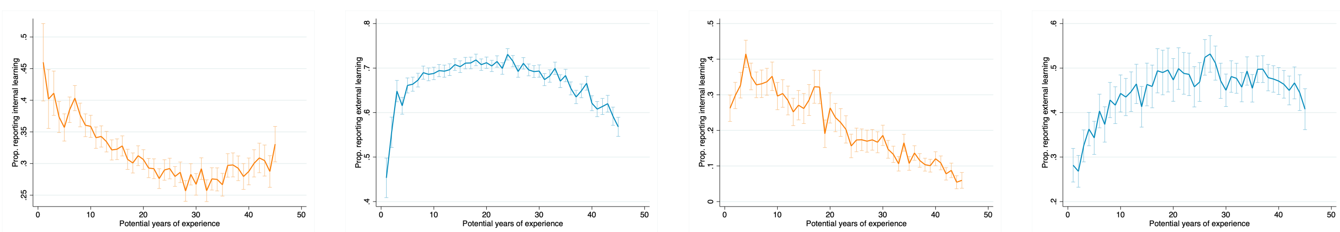


Notes: These figures plot the proportion of workers reporting engaging in internal and external learning across different age bins in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed, have between 1 and 45 years of potential experience, and are 21–60 years of age in both settings. The results are weighted using the observational weights provided in the surveys. 95% confidence intervals are plotted.

Figure C.2: Sources of learning throughout workers' lifecycles by one-year experience bins

Germany **United States**

(a) Internal learning (b) External learning (c) Internal learning (d) External learning

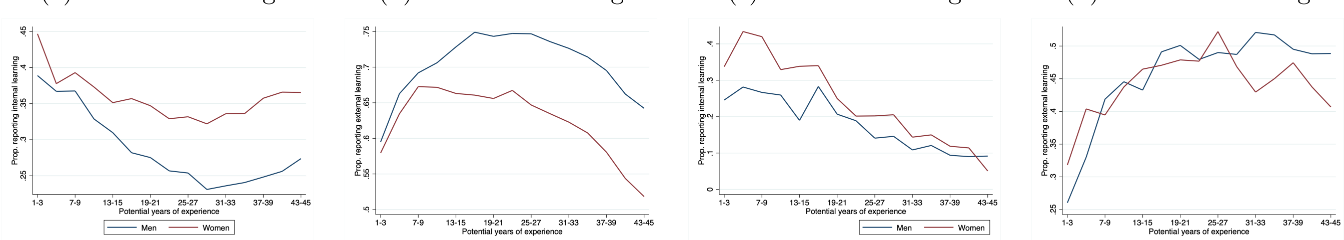


Notes: These figures plot the proportion of workers reporting engaging in internal and external learning across different levels of potential experience in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have between 1 and 45 years of potential experience in both settings. The results are weighted using the observational weights provided in the surveys. 95% confidence intervals are plotted.

Figure C.3: Sources of learning throughout workers' lifecycles by gender

Germany **United States**

(a) Internal learning (b) External learning (c) Internal learning (d) External learning

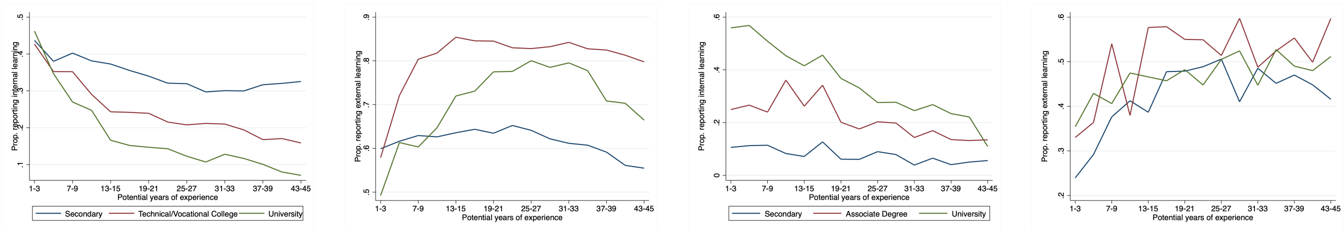


Notes: These figures plot the proportion of male and female workers reporting engaging in internal and external learning across different potential experience bins in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have between 1 and 45 years of potential experience in both settings. The results are weighted using the observational weights provided in the surveys. 95% confidence intervals are omitted for clarity.

Figure C.4: Sources of learning throughout workers' lifecycles by educational level

Germany **United States**

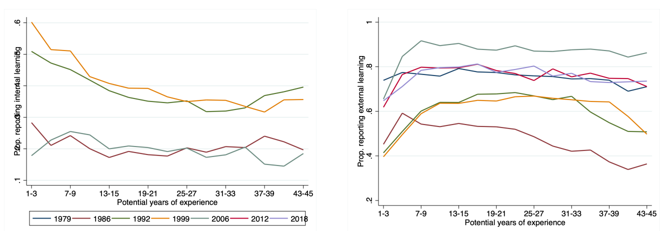
(a) Internal learning (b) External learning (c) Internal learning (d) External learning



Notes: These figures plot the proportion of workers with different educational attainment levels reporting engaging in internal and external learning across different potential experience bins in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have between 1 and 45 years of potential experience in both settings. The results are weighted using the observational weights provided in the surveys. 95% confidence intervals are omitted for clarity.

Figure C.5: Sources of learning throughout workers' lifecycles by survey wave in Germany

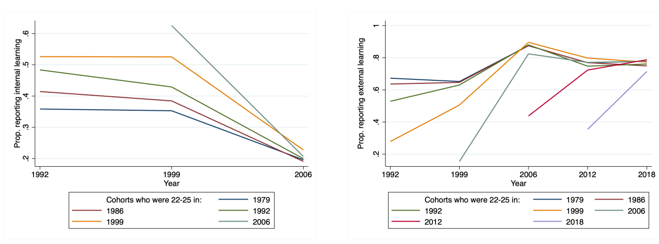
(a) Internal learning (b) External learning



Notes: These figures plot the proportion of workers reporting engaging in internal and external learning across different potential experience bins in different survey waves of the German BIBB/BAuA data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of potential experience. The results are weighted using the observational weights provided in the surveys. 95% confidence intervals are omitted for clarity.

Figure C.6: Sources of learning throughout workers' lifecycles by cohort in Germany

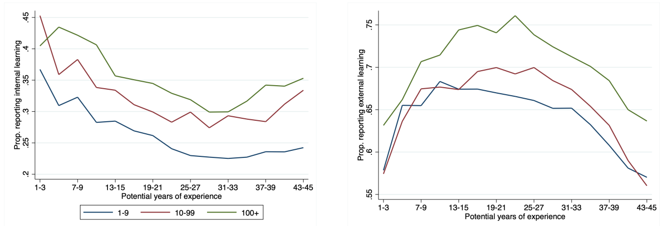
(a) Internal learning (b) External learning



Notes: These figures plot the proportion of workers of different cohorts reporting engaging in internal and external learning across different potential experience bins in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of potential experience. The results are weighted using the observational weights provided in the surveys. 95% confidence intervals are omitted for clarity.

Figure C.7: Sources of learning throughout workers' lifecycles by firm size in Germany

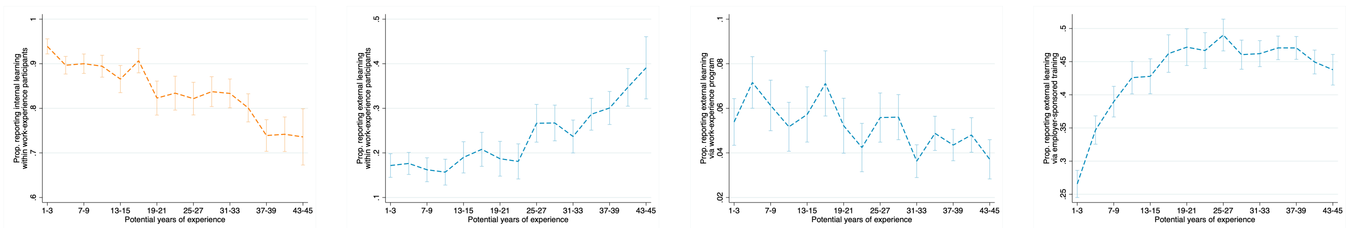
(a) Internal learning (b) External learning



Notes: These figures plot the proportion of workers employed in firms of different sizes reporting engaging in internal and external learning across different potential experience bins in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of potential experience. The results are weighted using the observational weights provided in the surveys. 95% confidence intervals are omitted for clarity.

Figure C.8: Decomposing different types of internal and external learning in the US

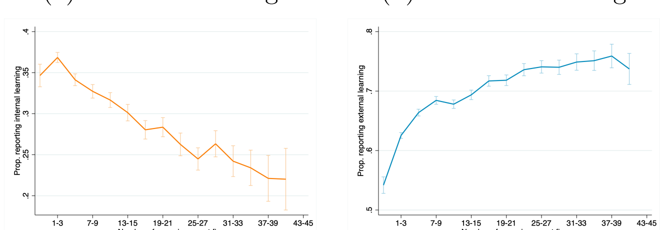
(a) Prevalence of internal learning within work-experience program participants (b) Prevalence of external learning within work-experience program participants (c) External learning via participation in work-experience program (d) External learning via certificate in employer-sponsored training



Notes: These figures plot the proportion of work-experience program participants reporting engaging in internal and external learning (panels (a) and (b), respectively), and the proportion of workers reporting engaging in external learning via participation in a work-experience program and certificate in an employer-sponsored program (panels (c) and (d), respectively) across different potential experience bins in the US NHES data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of potential experience. The results are weighted using the observational weights provided in the survey. 95% confidence intervals are plotted.

Figure C.9: Sources of learning by tenure in the current firm in Germany

(a) Internal learning (b) External learning



Notes: These figures plot the proportion of workers reporting engaging in internal and external learning across different tenure bins in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed, have between 1 and 45 years of potential experience, and have 42 years of tenure or less (we choose 42 specifically since it is one 3-year bin below our 45-year cut for potential experience). The results are weighted using the observational weights provided in the surveys. 95% confidence intervals are plotted.

Table C.2: Correlations between different sources of learning and tenure in Germany

	Internal learning				External learning			
Years w/ current emp.	-0.00591***	-0.00422***	-0.00311***	-0.00210**	0.00891***	0.00474***	0.00442***	0.00693***
	(0.000560)	(0.000669)	(0.000901)	(0.000968)	(0.000473)	(0.000514)	(0.000629)	(0.000668)
Years w/ current emp.^{2}	5.63e-05***	4.53e-05**	6.77e-05**	4.98e-05*	-0.000145***	-9.65e-05***	-0.000103***	-0.000118***
	(1.67e-05)	(1.99e-05)	(2.74e-05)	(2.91e-05)	(1.39e-05)	(1.48e-05)	(1.82e-05)	(1.90e-05)
Constant	0.372***	0.283***	0.342***	0.332***	0.609***	0.691***	0.671***	0.632***
	(0.00354)	(0.0125)	(0.0162)	(0.0167)	(0.00304)	(0.00963)	(0.0128)	(0.0131)
Observations	102,761	68,877	36,703	36,701	165,275	124,372	84,222	84,219
R-squared	0.007	0.127	0.077	0.079	0.009	0.193	0.206	0.211
Year FE		Y	Y	Y		Y	Y	Y
Demographic controls		Y	Y	Y		Y	Y	Y
Worker type FE		Y	Y	Y		Y	Y	Y
Industry FE		Y	Y	Y		Y	Y	Y
Occupation FE		Y	Y	Y		Y	Y	Y
Firm size FE		Y	Y	Y		Y	Y	Y
Wage controls			Y	Y			Y	Y
Age FE				Y				Y

Notes: This table shows the coefficients from regressing internal and external learning on years with current employer (tenure) and years with current employer squared in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed, and have between 1 and 45 years of potential experience. Regressions are weighted using the using the observational weights provided in the surveys. *Germany:* Year fixed effects correspond to the year of the survey. Demographic controls include educational attainment level and gender. Worker type categories include laborer, private employee, government employee, self-employed, freelancer, or family caregiver. Industry categories are at the 1-digit level. Occupation categories are at the 2-digit level (ISCO 88). Firm size is a categorical variable indicating whether the firm where the worker works has fewer than 4 workers, 5–9 workers, 10–49 workers, 50–99 workers, 100–499 workers, 500–999 workers, and 1000 or more workers. Wage controls include the current hourly wage. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C.1.2 Correlation between sources of learning

We study the correlation between internal and external learning by regressing the former on the latter and including a set of controls and fixed effects.[54] The results are summarized in Table C.3. We find a negative correlation between the two forms of learning in Germany, and a positive correlation in the US. This latter result likely reflects the fact that in the US both forms of learning occur in the context of a "work-experience program," namely a job with learning attributes. As such, several sources of learning are more likely to coexist.

[54]Please note that we include age fixed effects and education controls in these regressions, and thus do not include potential experience due to high collinearity between potential experience, education, and age.

Table C.3: Correlation between different sources of learning

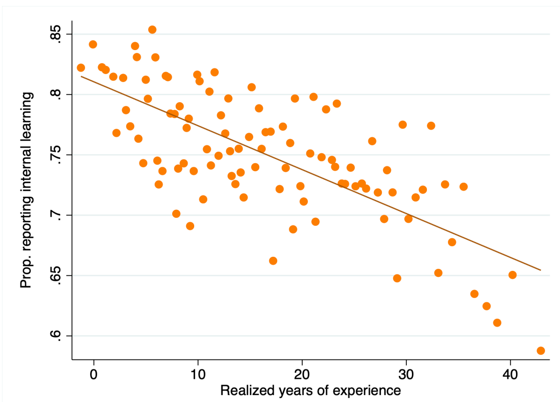
Germany			
		Internal learning	
External learning	-0.203***	-0.154***	-0.149***
	(0.00354)	(0.00456)	(0.00606)
Constant	0.438***	1.176***	1.227***
	(0.00301)	(0.0159)	(0.0200)
Observations	109,478	69,495	36,813
R-squared	0.045	0.150	0.101
Year FE		Y	Y
Demographic controls		Y	Y
Age FE		Y	Y
Worker type FE		Y	Y
Industry FE		Y	Y
Occupation FE		Y	Y
Firm size FE		Y	Y
Wage controls			Y
United States			
		Internal learning	
External learning	0.0949***	0.0903***	0.0936***
	(0.00773)	(0.00697)	(0.00710)
Constant	0.185***	0.0924***	0.0807***
	(0.00474)	(0.0105)	(0.0114)
Observations	29,399	29,398	27,766
R-squared	0.013	0.240	0.237
Demographic controls		Y	Y
Age FE		Y	Y
Worker type FE		Y	Y
Industry FE		Y	Y
Occupation FE		Y	Y
Wage controls			Y

Notes: This table shows the coefficients from regressing internal learning on external learning in the German BIBB/BAuA data and US NHES data. The samples encompass individuals who are currently employed, and have between 1 and 45 years of potential experience. Regressions are weighted using the using the observational weights provided in the surveys. *Germany:* Year fixed effects correspond to the year of the survey. Demographic controls include educational attainment level and gender. Worker type categories include laborer, private employee, government employee, self-employed, freelancer, or family caregiver. Industry categories are at the 1-digit level. Occupation categories are at the 2-digit level (ISCO 88). Firm size is a categorical variable indicating whether the firm where the worker works has fewer than 4 workers, 5–9 workers, 10–49 workers, 50–99 workers, 100–499 workers, 500–999 workers, and 1000 or more workers. Wage controls include the current hourly wage. *US:* Demographic controls include educational attainment level, race, census region, and gender. Worker type categories include private employee, government employee, self-employed, or working without pay. Industry and occupation categories are at the 2-digit level (ACS 2015). Wage controls include the current hourly wage. We do not not include potential experience in these regressions due to high collinearity between potential experience, education, and age. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

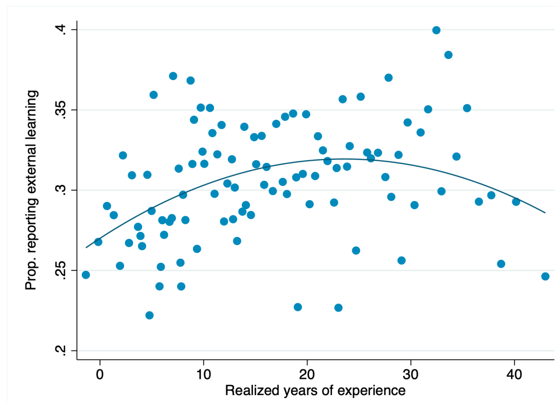
C.2 OECD PIAAC data

Figure C.10: Prevalence of different sources of learning throughout workers' lifecycles (with realized years of experience)

(a) Internal learning



(b) External learning



Notes: These binned scatterplots plot the proportion of workers reporting engaging in internal and external learning across different realized years of experience bins in the PIAAC data using pooled data from all countries considered and controlling for country fixed effects. The sample encompasses individuals who are currently employed, have between 1 and 45 years of both potential and realized experience. The results are weighted using the observational weights provided in the survey.

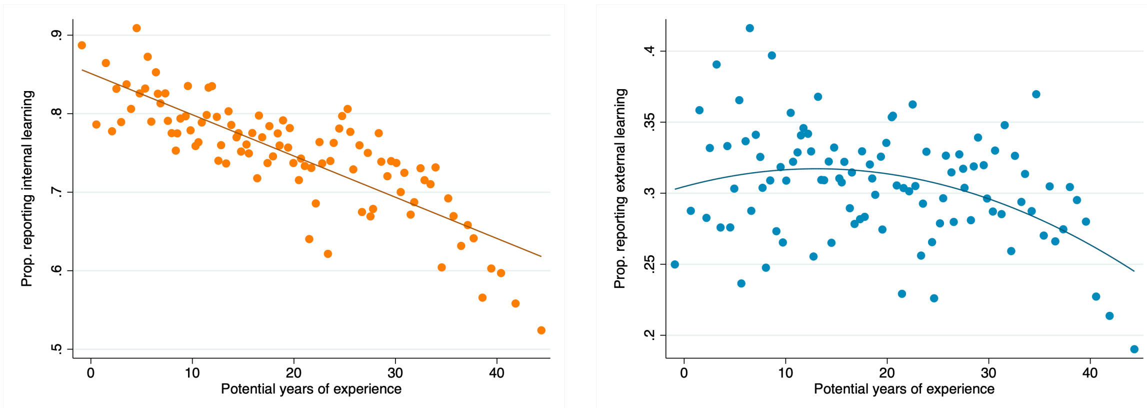
Table C.4: Correlations between different sources of learning and realized experience

	Internal learning			External learning		
Realized years of experience	-0.000834	-0.00432***	-0.00512***	0.00477***	0.00177	0.00148
	(0.00123)	(0.00121)	(0.00130)	(0.00108)	(0.00115)	(0.00126)
Realized years of experience^2	-7.28e-05**	-4.07e-06	1.09e-05	-0.000103***	-3.97e-05	-3.06e-05
	(3.15e-05)	(3.06e-05)	(3.31e-05)	(2.67e-05)	(2.83e-05)	(3.10e-05)
Constant	0.793***	0.719***	0.690***	0.266***	0.111***	0.106
	(0.00978)	(0.0521)	(0.153)	(0.00907)	(0.0359)	(0.111)
Observations	55,112	47,430	40,268	60,244	47,448	40,282
R-squared	0.043	0.121	0.123	0.051	0.159	0.158
Country FE	Y	Y	Y	Y	Y	Y
Demographic controls		Y	Y		Y	Y
Worker type FE		Y	Y		Y	Y
Industry FE		Y	Y		Y	Y
Occupation FE		Y	Y		Y	Y
Firm size FE		Y	Y		Y	Y
Wage controls			Y			Y

Notes: This table shows the coefficients from regressing internal and external learning on realized years of experience and realized years of experience squared in the PIAAC data. The sample encompasses individuals who are currently employed, and have between 1 and 45 years of potential experience. Regressions are weighted using the using the observational weights provided in the survey. Demographic controls include educational attainment level and gender. Worker type categories include private employee, government employee, non-profit employee and self-employed. Industry and occupation categories are at the 2-digit level (ISIC rev. 4 and ISCO 2008, respectively). Wage controls include the current hourly wage. We do not include age fixed effects due to the high collinearity between realized experience, education, and age. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure C.11: Prevalence of different sources of learning throughout workers' lifecycles (with potential years of experience)

(a) Internal learning (b) External learning



Notes: These binned scatterplots plot the proportion of workers reporting engaging in internal and external learning across different potential years of experience bins in the PIAAC data using pooled data from all countries considered and controlling for country fixed effects. The sample encompasses individuals who are currently employed, have between 1 and 45 years of both potential and realized experience. The results are weighted using the observational weights provided in the survey.

D Appendix: Additional results of quantitative model

D.1 Derivation of Match Value

We now provide a step-by-step discussion of how we derive the match value Equation (3). First, note that without accounting for spillover effects, the value function of a match between a worker i of age a_i and human capital h_{m_i} and a firm with productivity z, employing the workforce $\mathbb{I}(z)$ with human capital distribution $\mathbb{H}(z)$ encompasses the current output value net of training costs stemming from foregone production and payments to external trainers and the discounted future values contingent on whether worker i climbs the human capital ladder in the current period or not:

$$
\begin{aligned}
M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) = \max_{l,g} \Bigg[& \underbrace{z h_{m_i}(1-l)}_{\text{output value}} - \underbrace{(1-g)l \times q p_e(h_{m_i})}_{\text{pay for external trainers}} \\
& + \underbrace{s(h_{m_i}, \mathbb{H}(z)) \times \rho \mathbb{E}\left[(1-\kappa) M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\}\right]}_{\text{future value if successfully climb human capital ladder}} \\
& + \underbrace{[1 - s(h_{m_i}, \mathbb{H}(z))] \times \rho \mathbb{E}\left[(1-\kappa) M^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i}), V_{TR}^{a_i+1}(h_{m_i})\}\right]}_{\text{future value if don't successfully climb human capital ladder}} \Bigg].
\end{aligned}
\tag{D.1}
$$

This match value does not account for spillover effects across workers in the same firm. These spillovers arise because coworkers affect the human capital distribution of the firm, $\mathbb{H}(z)$, and thus influence its internal learning environment by impacting the probability each worker has of climbing the human capital ladder.

In particular, recall that the probability of climbing the human capital ladder, $s(\cdot)$, follows $s(h_{m_i}, \mathbb{H}(z)) = \min\left\{ \left[[A(z)p(h_{m_i}, \mathbb{H}(z))g]^{\frac{\sigma-1}{\sigma}} + [A_e(z)p_e(h_{m_i})(1-g)]^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} l^{\gamma} + \epsilon, 1 \right\}$, and thus encompasses the role of these spillovers by considering the probability of matching with a coworker with higher human capital than the own within the same firm, $p(h_{m_i}, \mathbb{H}(z))$, when learning internally.

To incorporate these spillover effects, we expand Equation (D.1) as:

$$
\begin{aligned}
M^{a_i}(h_{m_i}, \mathbb{H}(z)) = \max_{l,g} \Bigg[& \underbrace{z h_{m_i}(1-l)}_{\text{output value}} - \underbrace{(1-g)l \times q p_e(h_{m_i})}_{\text{pay for external trainers}} \\
& + \underbrace{\int_{j \in \mathbb{I}(z), j \neq i} M^{a_j}(h_{m_j}, z, \mathbb{H}(z)) - M^{a_j}(h_{m_j}, z, \mathbb{H}(z) \backslash \{i\}) dj}_{\text{spillover effects to other workers}} - \underbrace{\int_{j \in \mathbb{I}(z), j \neq i} M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) - M^{a_i}(h_{m_i}, z, \mathbb{H}(z) \backslash \{j\}) dj}_{\text{spillover effects to worker } i} \\
& + \underbrace{s(h_{m_i}, \mathbb{H}(z)) \times \rho \mathbb{E}\left[(1-\kappa) M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\} \right]}_{\text{future value if successfully climb human capital ladder}} \\
& + \underbrace{[1 - s(h_{m_i}, \mathbb{H}(z))] \times \rho \mathbb{E}\left[(1-\kappa) M^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i}), V_{TR}^{a_i+1}(h_{m_i})\} \right]}_{\text{future value if don't successfully climb human capital ladder}} \Bigg].
\end{aligned}
\tag{D.2}
$$

In the second line, we first incorporate the spillover effects to other workers in this firm stemming from the presence of worker i. Then, to avoid double accounting, and considering that workers may take a wage cut to fund human capital acquisition and compensate mentors inside the firm for the spillover effects they generate in their learning environment (Jarosch et al., 2021), we subtract the spillover effects to worker i stemming from the presence of other workers in this firm.

We isolate a worker j's effect on another worker i's match value by computing the increment in the match value worker i obtains from the fact that worker j is part of their same firm. In particular, we compare the match value worker i obtains from the match with firm z when worker j is in this firm's workforce in the current period, to the match value worker i would obtain if worker j were not in the firm's workforce in the current period, taking all else as given.[55] Thus, using the match value above, worker j's effect on another worker i's match

[55] As before, we write $\mathbb{H}(z)\backslash\{j\}$ to denote the human capital distribution of the firm's workforce absent

value can be written as:

$$
\begin{aligned}
& M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) - M^{a_i}(h_{m_i}, z, \mathbb{H}(z)\backslash\{j\}) \\
=& [zh_{m_i}(1 - l^{a_i}(h_{m_i}, z, \mathbb{H}(z))) - (1 - g^{a_i}(h_{m_i}, z, \mathbb{H}(z)))l^{a_i}(h_{m_i}, z, \mathbb{H}(z)) \times qp_e(h_{m_i})] \\
& - [zh_{m_i}(1 - l^{a_i}(h_{m_i}, z, \mathbb{H}(z)\backslash\{j\})) - (1 - g^{a_i}(h_{m_i}, z, \mathbb{H}(z)\backslash\{j\}))l^{a_i}(h_{m_i}, z, \mathbb{H}(z)\backslash\{j\}) \times qp_e(h_{m_i})] \\
& + [s(h_{m_i}, \mathbb{H}(z)) - s(h_{m_i}, \mathbb{H}(z)\backslash\{j\})] \times \rho\mathbb{E}\left[(1-\kappa)M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\}\right] \\
& - [s(h_{m_i}, \mathbb{H}(z)) - s(h_{m_i}, \mathbb{H}(z)\backslash\{j\})] \times \rho\mathbb{E}\left[(1-\kappa)M^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i}), V_{TR}^{a_i+1}(h_{m_i})\}\right] \\
\approx & \frac{\partial s(h_{m_i}, \mathbb{H}(z))}{\partial p(h_{m_i}, \mathbb{H}(z))}[p(h_{m_i}, \mathbb{H}(z)) - p(h_{m_i}, \mathbb{H}(z)\backslash\{j\})] \times \rho\mathbb{E}\left[(1-\kappa)\Delta M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}'(z)) + \kappa\Delta \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\}\right]
\end{aligned} \tag{D.3}
$$

The second equality is derived due to the envelope theorem, as the optimal training time $l^{a_i}(h_{m_i}, z, \mathbb{H}(z)))$ and time allocation $g^{a_i}(h_{m_i}, z, \mathbb{H}(z)))$ are derived to maximize the match value. In addition, we can ignore the integral terms capturing spillover effects in this expression since we consider a continuum of workers, and thus each worker's impact on other workers' spillover effects is negligible. Finally, since we only consider worker j's spillover effect in the current period, we take the future human capital distribution $\mathbb{H}'(z)$ as given.[56]

We can use this expression to show that the spillover effects to worker i stemming from the presence of other workers in this firm is approximately equal to zero. In particular, denote $O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) = \left[(1-\kappa)\Delta M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}'(z)) + \kappa\Delta \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\}\right]$ and notice that using the expression above, the value of these spillovers can be written as:

$$
\begin{aligned}
& \underbrace{\int_{j\in\mathbb{I}(z), j\neq i} M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) - M^{a_i}(h_{m_i}, z, \mathbb{H}(z)\backslash\{j\})dj}_{\text{spillover effects to worker } i} \\
&= \int_{j\in\mathbb{I}(z), j\neq i} \frac{\partial s}{\partial p}[p(h_{m_i}, \mathbb{H}(z)) - p(h_{m_i}, \mathbb{H}(z)\backslash\{j\})] \times O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z))dj \\
&= \frac{\partial s}{\partial p} O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) \int_{j\in\mathbb{I}(z), j\neq i} [p(h_{m_i}, \mathbb{H}(z)) - p(h_{m_i}, \mathbb{H}(z)\backslash\{j\})] \times dj
\end{aligned} \tag{D.4}
$$

Notice however that this is approximately equal to zero, since we consider a continuum of

worker j: $\mathbb{H}(z)\backslash\{j\} = \{h_{m_i}\}_{i\in\mathbb{I}(z), i\neq j}$.

[56] In principle, the hypothetical scenario of worker j not being part of the firm's workforce in the current period can also affect the future human capital distribution by affecting human capital evolution. However, given that we consider a continuum of workers, each worker's effect on overall human capital evolution is negligible.

workers in the workforce, and thus have

$$\int_{j\in\mathbb{I}(z),j\neq i}\left[p(h_{m_i},\mathbb{H}(z))-p(h_{m_i},\mathbb{H}(z)\backslash\{j\})\right]dj$$
$$\approx\int_{j\in\mathbb{I}(z),j\neq i}\left(\mathbf{1}_{\{h_{m_j}>h_{m_i}\}}\frac{1}{D(z)}-p(h_{m_i},\mathbb{H}(z))\frac{1}{D(z)}\right)dj=p(h_{m_i},\mathbb{H}(z))-p(h_{m_i},\mathbb{H}(z))=0,$$

where $D(z)$ captures the size of the workforce in firm z. In this equation, the first equality approximates the impact of worker j on worker i's learning probability which stems from two sources: (1) an increase in the mass of potential trainers if worker j's human capital is higher than worker i; and (2) an increase in the size of the workforce in firm z which reduces the chance of worker i to meet potential trainers. These two terms cancel out if we take an integral across the workforce, as we are considering the joint impact of all other workers, including those with both higher and lower human capital to worker i, in the firm.

D.2 Worker and firm values

Worker's value. The value for a worker of age a_i and human capital h_{m_i} matched with a firm with productivity z employing the workforce $\mathbb{I}(z)$ with human capital distribution $\mathbb{H}(z)$, and who perceives the share of revenue r is given by:

$$\begin{aligned}V^{a_i}(r,h_{m_i},z,\mathbb{H}(z))=&r\Bigg[\underbrace{zh_{m_i}(1-l)}_{\text{output value}}-\underbrace{(1-g)l\times qp_e(h_{m_i})}_{\text{pay for external trainers}}+\underbrace{\int_{j\in\mathbb{I}(z),j\neq i}\frac{\partial s}{\partial p}\left[p(h_{m_j},\mathbb{H}(z))-p(h_{m_j},\mathbb{H}(z)\backslash\{i\})\right]O^{a_j+1}(h_{m_j},z,\mathbb{H}'(z))dj}_{\text{spillover effects to other workers in firm}}\\&+\underbrace{s(h_{m_i},\mathbb{H}(z))\times\rho\mathbb{E}\left[(1-\kappa)V^{a_i+1}(r,h_{m_i+1},z,\mathbb{H}'(z))+\kappa\max\{V_U^{a_i+1}(h_{m_i+1}),V_{TR}^{a_i+1}(h_{m_i+1})\}\right]}_{\text{future value if successfully climb human capital ladder}}\\&+\underbrace{[1-s(h_{m_i},\mathbb{H}(z))]\times\rho\mathbb{E}\left[(1-\kappa)V^{a_i+1}(r,h_{m_i},z,\mathbb{H}'(z))+\kappa\max\{V_U^{a_i+1}(h_{m_i}),V_{TR}^{a_i+1}(h_{m_i})\}\right]}_{\text{future value if don't successfully climb human capital ladder}}\end{aligned}$$

Firm's value. The value for a firm with productivity z, employing the workforce $\mathbb{I}(z)$ with human capital distribution $\mathbb{H}(z)$, of matching with a worker of age a_i and human capital h_{m_i} who perceives the share of revenue r is given by:

$$\begin{aligned}J^{a_i}(r,h_{m_i},z,\mathbb{H}(z))=&(1-r)\Bigg[\underbrace{zh_{m_i}(1-l)}_{\text{output value}}-\underbrace{(1-g)l\times qp_e(h_{m_i})}_{\text{pay for external trainers}}\underbrace{\int_{j\in\mathbb{I}(z),j\neq i}\frac{\partial s}{\partial p}\left[p(h_{m_j},\mathbb{H}(z))-p(h_{m_j},\mathbb{H}(z)\backslash\{i\})\right]O^{a_j+1}(h_{m_j},z,\mathbb{H}'(z))dj}_{\text{spillover effects to other workers in firm}}\\&+\underbrace{s(h_{m_i},\mathbb{H}(z))\times\rho\mathbb{E}(1-\kappa)J^{a_i+1}(r,h_{m_i+1},z,\mathbb{H}'(z))}_{\text{future value if successfully climb human capital ladder}}+\underbrace{[1-s(h_{m_i},\mathbb{H}(z))]\times\rho\mathbb{E}(1-\kappa)J^{a_i+1}(r,h_{m_i},z,\mathbb{H}'(z))}_{\text{future value if don't succesfully climb human capital ladder}}.\end{aligned}$$

D.3 Proof of Proposition 1

Firms and workers jointly maximize the match value in Equation (3). Given the definitions of $s(h_{m_i}, \mathbb{H}(z))$ and $s^E(h_{m_i}, \mathbb{H}(z))$ in Equations (1) and (2), if $s(h_{m_i}, \mathbb{H}(z)) < 1$, which ensures internal solutions, the first-order condition with respect to l implies:

$$zh_{m_i} + (1-g)qp_e(h_{m_i}) = \gamma s^E(h_{m_i}, \mathbb{H}(z)) l^{\gamma-1} O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)).$$

Thus, we can solve l as:

$$l = \left(\frac{\gamma s^E(h_{m_i}, \mathbb{H}(z)) O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z))}{zh_{m_i} + (1-g)qp_e(h_{m_i})} \right)^{1/(1-\gamma)}.$$

The first-order condition with respect to g implies

$$\begin{aligned} & lqp_e(h_{m_i}) + \left[s^E(h_{m_i}, \mathbb{H}(z))^{\frac{1}{\sigma}} \left(\frac{(A(z)p(h_{m_i}, \mathbb{H}(z)))^{\frac{\sigma-1}{\sigma}}}{g^{\frac{1}{\sigma}}} - \frac{(A_e(z)p_e(h_{m_i}))^{\frac{\sigma-1}{\sigma}}}{(1-g)^{\frac{1}{\sigma}}} \right) \right] \\ & \times O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) l^{\gamma} = 0. \end{aligned}$$

Combining the above two equations we obtain:

$$1 = \frac{zh_{m_i} + (1-g)qp_e(h_{m_i})}{\gamma qp_e(h_{m_i})} \times \frac{(A_e(z)p_e(h_{m_i}))^{\frac{\sigma-1}{\sigma}} (1-g)^{-\frac{1}{\sigma}} - (A(z)p(h_{m_i}, \mathbb{H}(z)))^{\frac{\sigma-1}{\sigma}} g^{-\frac{1}{\sigma}}}{(A_e(z)p_e(h_{m_i}))^{\frac{\sigma-1}{\sigma}} (1-g)^{\frac{\sigma-1}{\sigma}} + (A(z)p(h_{m_i}, \mathbb{H}(z)))^{\frac{\sigma-1}{\sigma}} g^{\frac{\sigma-1}{\sigma}}} \tag{D.5}$$

All else being equal, it can be shown that g increases with $p(h_{m_i}, \mathbb{H}(z))$. This is because the right-hand side of Equation (D.5) increases with both $p(h_{m_i}, \mathbb{H}(z))$ and g, suggesting a positive partial derivative of g with regard to $p(h_{m_i}, \mathbb{H}(z))$.[57] Similarly, we can show that g decreases with q.

D.4 General equilibrium

Define $D^a(h_m, z)$ and $D_t^a(h_m)$ to be the measures of employed workers of age a and human capital level h_m ($m = 1, 2, ...$) at firm z and in the training sector, respectively. We define the general equilibrium of our model as follows:

[57] It is easy to show that the right-hand side declines with $p(h_{m_i}, \mathbb{H}(z))$ by taking the derivative. For the right-hand side to increase with g, $-\frac{(1-\frac{\sigma-1}{\sigma}\gamma)qp_e(h_{m_i})}{zh_{m_i}+(1-g)qp_e(h_{m_i})} + \frac{1}{\sigma} \frac{\left(A_e(z)p_e(h_{m_i})\right)^{\frac{\sigma-1}{\sigma}} (1-g)^{-\frac{1}{\sigma}-1} + \left(A(z)p(h_{m_i}, \mathbb{H}(z))\right)^{\frac{\sigma-1}{\sigma}} g^{-\frac{1}{\sigma}}}{\left(A_e(z)p_e(h_{m_i})\right)^{\frac{\sigma-1}{\sigma}} (1-g)^{-\frac{1}{\sigma}-1} - \left(A(z)p(h_{m_i}, \mathbb{H}(z))\right)^{\frac{\sigma-1}{\sigma}} g^{-\frac{1}{\sigma}}} > 0$ needs to hold, which is always true for all firms and human capital levels in our baseline calibration.

Definition 1. *The general equilibrium consists of meeting rates* $\{\lambda_U\}$*, employment distributions* $\{D^a_U(h_m), D^a(h_m, z), D^a_t(h_m)\}$*, firms' vacancy postings* $v(z)$*, the worker-firm joint decision of human capital accumulation* $\{g^a(h_m, z, \mathbb{H}(z)), l^a(h_m, z, \mathbb{H}(z))\}$*, workers' wage rates* $w^a(h_m, z, \mathbb{H}(z))$*, and the price of external training services* q*. These variables satisfy:*
(i) the wage rate $w^a(h_m, z, \mathbb{H}(z))$ *is determined by the bargaining processes in Section 4.2.2;*
(ii) the worker-firm joint decision of human capital accumulation $\{g^a(h_m, z, \mathbb{H}(z)), l^a(h_m, z, \mathbb{H}(z))\}$ *satisfies Equations (4) and (5) in Proposition 1, where the learning probabilities of internal and external learning sources are determined by the employment distributions (*$D^a(h_m, z)$ *within each firm and* $D^a_t(h_m)$ *in the training sector);*
(iii) firms' optimal vacancy postings $v(z)$ *are given by Equation* (9)*;*
(iv) the meeting rate λ_U *is determined by the number of unemployed* $U = \sum_a \sum_m D^a_U(h_m)$ *and the total number of vacancies* $V = \int v(z)ds(z)$*;*
(v) the employment distributions $\{D^a_U(h_m), D^a(h_m, z), D^a_t(h_m)\}$ *are consistent with vacancy postings* $v(z)$*, the worker-firm decision of human capital accumulation* $\{g^a(h_m, z), l^a(h_m, z)\}$*, exogenous job separations, and unemployed workers' sectoral choices; and*
(vi) the training price q *clears the market for external training services such that the total demand for external training services (external training time aggregated across all production workers) equals the total supply (training units provided by all external trainers).*

E Appendix: Additional quantitative results

E.1 Identification of parameters

Table E.1: Elasticities of targeted moments to parameters

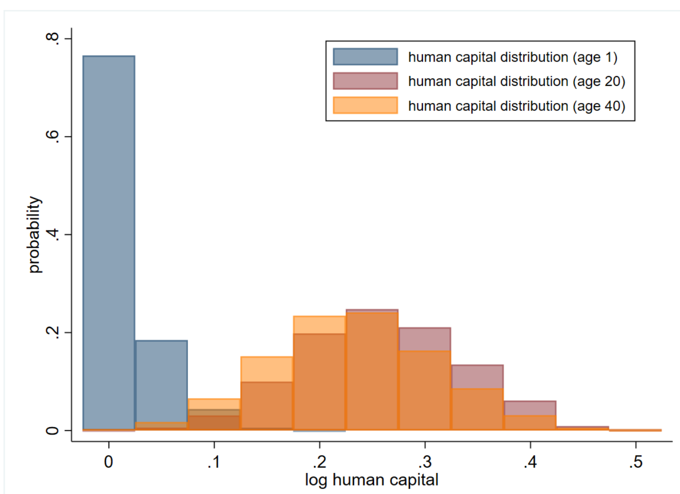
Moment	$\bar{A}$	$\bar{A}_e$	α	σ	ϵ	δ_h	c_m	c_v	ζ	κ
	Parameter									
Unemployment rate	-0.07	-0.05	-0.03	0.03	0.00	0.11	**-1.98**	0.31	0.41	**1.28**
Labor market tightness (#vacancies/#unemployed)	0.17	0.09	0.09	-0.05	0.03	-0.19	**1.39**	**-0.65**	**-0.82**	**-0.94**
Shape parameter of firm employment distribution	-0.01	-0.01	-0.02	0.00	0.00	0.00	0.00	0.00	**0.10**	0.01
Share of workers that remain employed in next quarter	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	**-0.03**
Ratio of new to all workers' avg external learning time	**1.07**	**-0.79**	-0.07	**3.08**	-0.01	0.42	0.00	0.00	0.22	0.04
Average wage growth (per quarter), 0–40 yrs of exp	**1.61**	**0.66**	**0.53**	-0.22	0.43	**-1.47**	0.14	0.02	-0.32	-0.12
Average wage growth (per quarter), 0–25 yrs of exp	**1.68**	**0.70**	**0.58**	-0.12	0.41	**-1.66**	0.20	0.07	-0.30	-0.12
Share of total time spent on external learning	0.28	**1.87**	0.47	**-0.73**	0.06	**-0.81**	-0.01	0.00	-0.38	-0.17
Share of total time spent on internal learning	**2.30**	-0.04	**0.57**	-0.09	0.20	**-1.18**	0.00	0.00	-0.22	-0.19
Ratio of learning time in 100+ to 50–99 worker firms	0.00	-0.02	**0.23**	-0.01	0.02	0.17	0.00	0.00	-0.13	-0.07

Notes: This table presents the elasticity of each moment targeted in the model calibration to each internally calibrated parameter. These elasticities are determined by calculating the percentage increase in each moment after a 10% change in the calibrated parameter value while holding the remaining parameters fixed. For each moment, we highlight in bold the elasticities surpassing an absolute value of 0.5, or the largest elasticity if no elasticities pertaining to that moment surpass this level.

E.2 Properties of equilibrium

E.2.1 Distribution of human capital levels at different ages

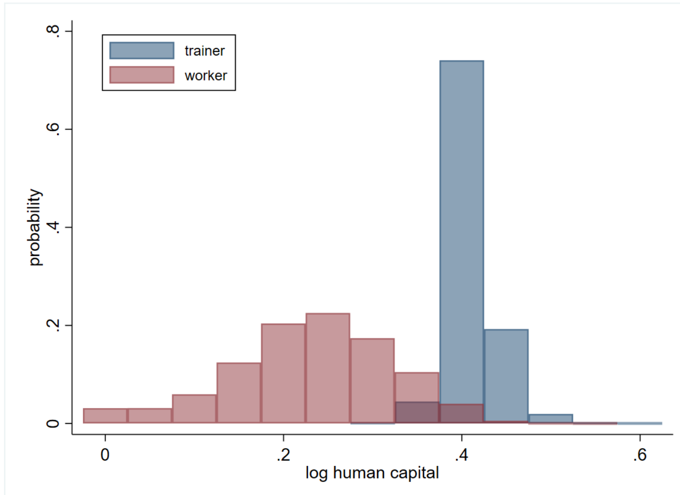
Figure E.1: Distribution of human capital at different ages



Notes: This figure illustrates how the distribution of human capital changes as workers age in the model by plotting the share of workers of each human capital level within a given cohort of workers observed at different ages.

E.2.2 Distribution of Trainers' and Workers' Human Capital

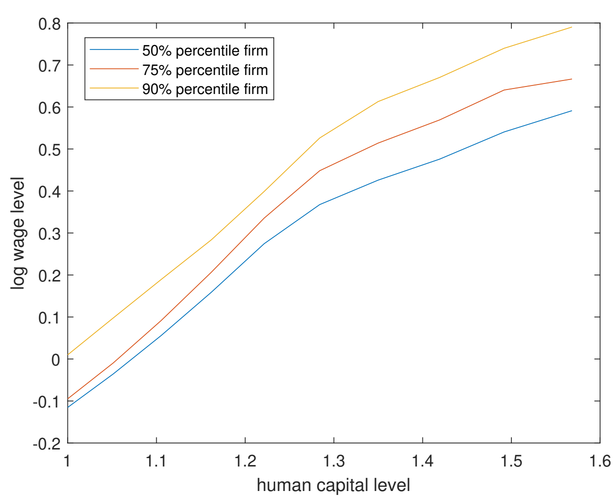
Figure E.2: Distribution of trainers' and workers' human capital



Notes: This figure illustrates the human capital distribution of production workers and external trainers in the model by plotting the share of workers of each human capital level within these two groups.

E.2.3 Wages by firm productivity and human capital

Figure E.3: Wages by firm productivity and human capital



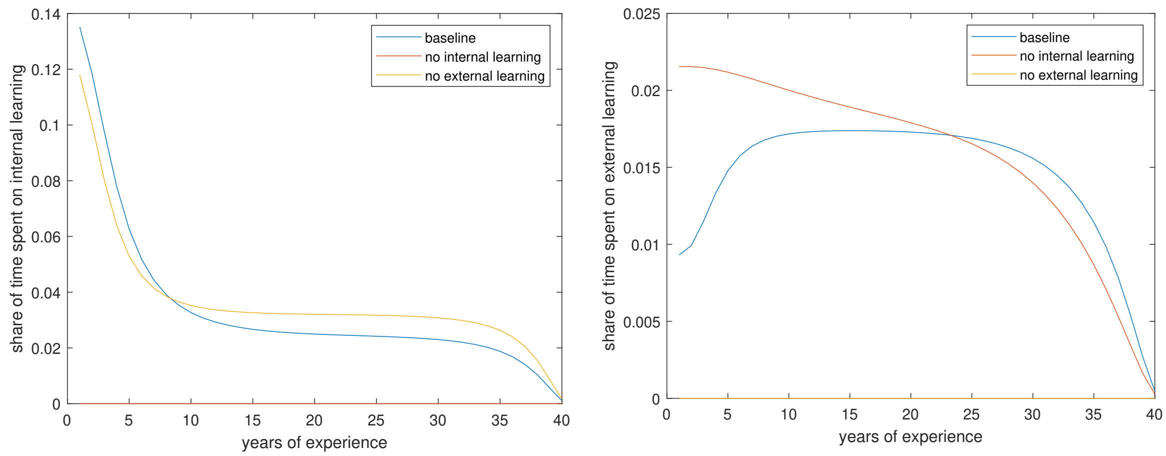
Notes: This figure plots the wages perceived by workers of each human capital level when working in firms with different productivity levels in the model.

E.3 Counterfactual exercises

E.3.1 Lifetime patterns of internal and external learning

Figure E.4: Lifecycle patterns of internal and external learning in baseline and counterfactual scenarios

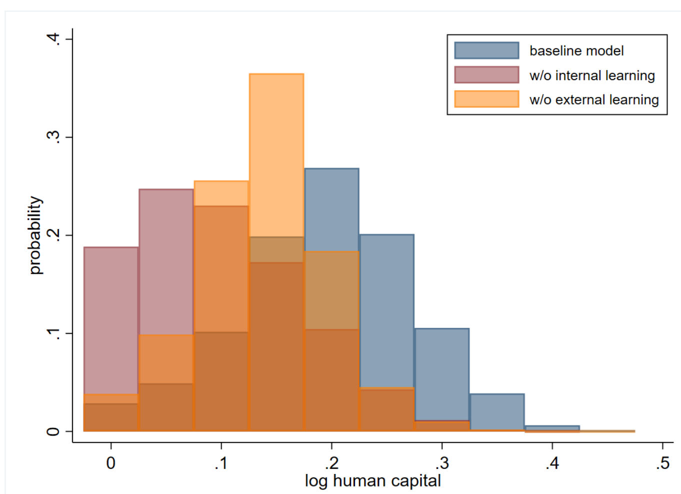
(a) Internal learning (b) External learning



Notes: These figures plot the share of time spent on internal and external learning, namely $l \times g$, and $l \times (1-g)$, respectively, for workers with different years of experience in the baseline model and counterfactual scenarios that shut down internal and external learning, respectively.

E.3.2 Human capital distribution

Figure E.5: Human capital distribution in baseline and counterfactual scenarios



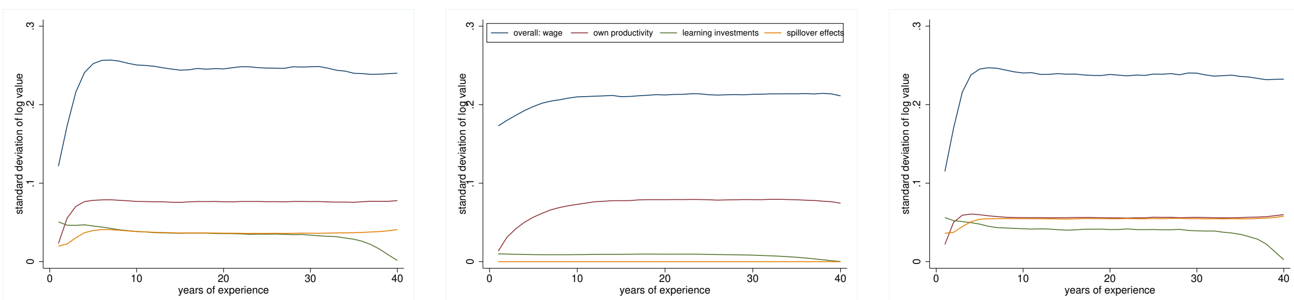
Notes: This figure plots the share of workers with different human capital levels in the baseline model and counterfactual scenarios that shut down internal and external learning, respectively.

E.3.3 Role of each learning source in wage dispersion

In Figure E.6 we plot the standard deviation of log wages among workers in the baseline and counterfactual scenarios, along with the various components driving it.[58] We highlight three notable findings. First, wage dispersion increases over the lifecycle in the baseline model, consistent with the literature.[59] This increase partially reflects the growing dispersion in workers' own productivity levels: all new workers begin with the same level of productivity, but some individuals have better luck and climb the human capital ladder more rapidly. Furthermore, we find that spillover effects to colleagues also become more dispersed as workers age, collectively contributing to the rise in wage dispersion. As workers refrain from investing in human capital when nearing retirement, the dispersion of learning costs declines at the oldest ages.

Figure E.6: Lifecycle wage dispersion and its components in the baseline and counterfactual scenarios

(a) Baseline model (b) Without internal learning (c) Without external learning



Notes: These figures illustrate the lifetime progression of wage dispersion and its different components (own productivity, spillover effects to coworkers, and learning investments and associated costs) by plotting the standard deviation of the log value of each of these for workers with different years of experience in the baseline model and counterfactual scenarios that shut down internal and external learning, respectively.

Second, the dispersion in workers' own productivity levels is more driven by external learning than by internal learning. Without external learning, the dispersion in workers' productivities remains low throughout the lifecycle as workers learn from and catch up fast to colleagues. Without internal learning, since skill acquisition is more expensive, only a small number of

[58] Similar to before, the standard deviations of the share of revenue attributed to workers and the productivity of the firms workers are matched with do not change significantly over workers' lifetimes. Thus, we omit these two components for the sake of clarity.

[59] The wage dispersion predicted by our model is relatively small compared to the data as we abstract from individual heterogeneity in innate abilities and other factors such as job ladders. The slight hump shape observed is consistent with the literature (Lagakos et al. (2018)).

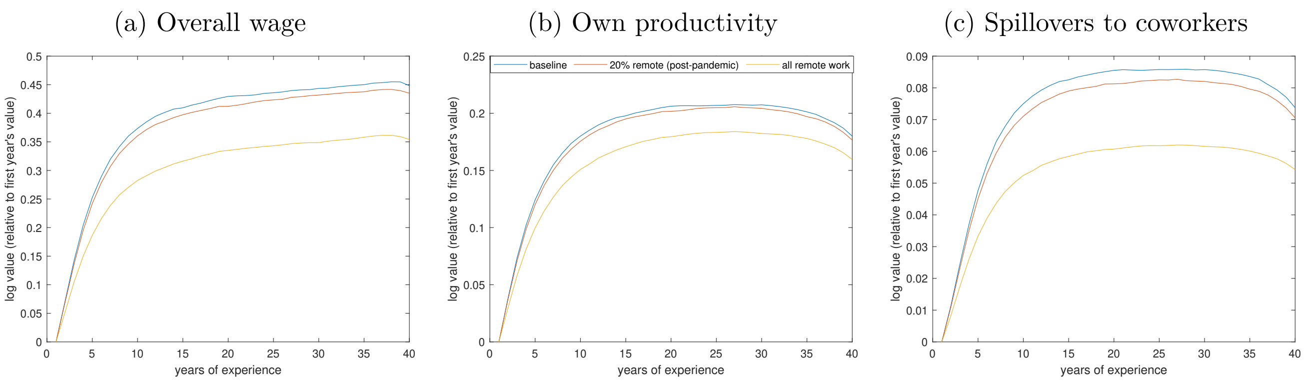
workers climb the human capital ladder, causing the dispersion in workers' own productivity levels to rise throughout the lifecycle.[60]

Third, we observe that even with the broader own-productivity dispersion, wage dispersion is actually lower when we shut down internal learning compared to when we shut down external learning. This can be primarily attributed to the spillover effects linked to internal learning, which contribute significantly to the wage gains for more-experienced workers.

E.4 Remote work

E.4.1 Lifetime wage growth and its components

Figure E.7: Impact of remote work on lifecycle wage growth and its components



Notes: These figures illustrate the lifetime progression of wage growth and two of its components (own productivity, and spillover effects to coworkers)[61] by plotting the average log value of each of these (normalized to the first year's value) for workers with different years of experience in the baseline model and 20% and 100% remote work scenarios.

E.5 Subsidies to learning

We now evaluate the role of subsidies in addressing learning inefficiencies in our framework. These inefficiencies arise because firms and workers do not take into account future employers' and coworkers' gains when choosing learning investments. To this end, we examine the impacts of government-sponsored subsidies that cover a portion of all learning expenses, and expenses pertaining solely to internal or external learning, respectively. These subsidies are

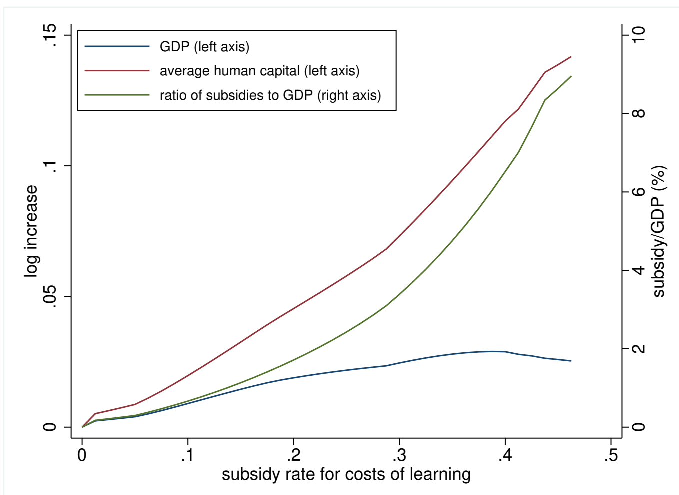
[60] In Figure E.5 we plot the distribution of human capital among workers in our baseline equilibrium and the counterfactual scenarios, showing that the distribution of human capital is more dispersed when we shut down internal learning compared to when we shut down external learning.

[61] To save on space, we do not plot changes in learning investments, which also change with remote work, and account for part of the decline in lifetime wage growth.

funded through lump-sum taxes, and encompass both trainers' fees (in the case of subsidies covering all or external learning expenses) and the value of lost production time.

Figure E.8 shows the gains in GDP and workers' average human capital in relation to the government's subsidy rates when subsidies cover a portion of all learning expenses.[62] GDP has a hump-shaped relationship with subsidy levels, since very high subsidy rates lead to a large portion of workers' time being allocated to learning, which subsequently lowers GDP. We find that the impact of the subsidy rate on GDP is maximized when the subsidy rate is around 38%. At this subsidy rate, equivalent to 6% GDP allocated for learning subsidies, average human capital and GDP increase by 11.1% and 2.9% in the stationary equilibrium, respectively, indicating potential advantages of government-sponsored learning policies.

Figure E.8: Subsidies to all learning costs



Notes: This figure illustrates the gains from learning subsidies covering a portion of all learning expenses by plotting the percent increase (measured through the log of the gross increase) in GDP and average human capital relative to the baseline model (left vertical axis), along with the ratio of subsidies to GDP (right vertical axis) from different subsidy rates. We compute GDP as the sum of the values produced by the production sector (which takes into account the loss of production time due to learning) and training sector.

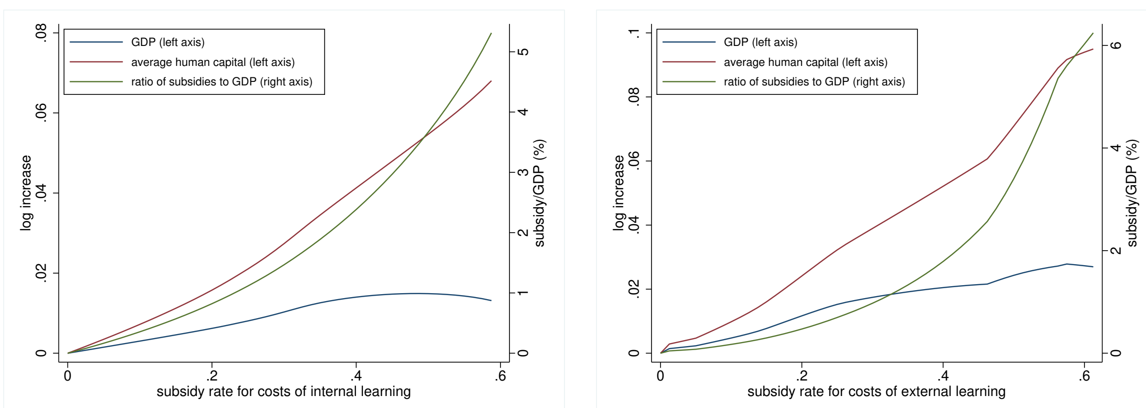
We now examine the effects of subsidizing expenses pertaining solely to internal or external learning, respectively. The results of these analyses are presented in Figure E.9. We have two main findings. First, we find that the optimal subsidy rate, which produces the largest impact on GDP, is higher when either internal or external learning is individually subsidized, compared to when both types of learning are subsidized simultaneously. Sec-

[62]Since our model incorporates two sectors (production and training), we compute GDP as the sum of the realized production value in the production sector (which takes into account the loss of production time due to learning) and the value of training services provided by trainers.

ond, and most importantly, we find that subsidizing external learning yields greater benefits than subsidizing internal learning. When the government subsidizes internal learning, the maximum gain in GDP is 1.5%, achieved at a subsidy rate of 49%. When the government subsidizes external learning, on the other hand, the maximum gain in GDP is 2.7%, achieved at a subsidy rate of 58%. This suggests that external learning exhibits larger externalities compared to internal learning. One possible explanation for this is that trainers, who deliver external learning, typically possess higher levels of human capital compared to production workers. In addition, promoting external learning contributes to an enhanced distribution of human capital within firms. Overall, our results highlight the potential for targeted learning subsidies to drive economic growth and improve human capital outcomes.

Figure E.9: Subsidies to each learning source

(a) Subsidies to internal learning (b) Subsidies to external learning



Notes: These figures illustrate the gains from learning subsidies covering a portion of internal and external learning expenses, respectively, by plotting the percent increase (measured through the log of the gross increase) in GDP and average human capital relative to the baseline model (left vertical axis), along with the ratio of subsidies to GDP (right vertical axis) from different subsidy rates. We compute GDP as the sum of the values produced by the production sector (which takes into account the loss of production time due to learning) and training sector.

F Appendix: Model extensions and robustness

F.1 Incorporating on-the-job search

F.1.1 Model extension

We now consider a model extension that incorporates on-the-job search following Cahuc et al. (2006), and thus assume that employed people can search for new jobs together with the unemployed. We denote $0 \leq \eta < 1$ as the search intensity of employed workers relative to unemployed workers whose search intensity is normalized to 1.

Once a worker receives an outside offer from a poaching firm with productivity z', there are three possible scenarios which determine whether the worker moves or not, and how her perceived fraction of current revenue r, which governs her wage, evolves.

1. The poaching firm cannot offer the worker a value higher than the worker's current value, $V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z)) > M^{a_i}(h_{m_i}, z', \mathbb{H}(z'))$. In this case, the worker will not move, and her perceived fraction of current revenue r does not change.

2. The match in the poaching firm z' is more valuable than the worker's current match with firm z, $M^{a_i}(h_{m_i}, z', \mathbb{H}(z')) > M^{a_i}(h_{m_i}, z, \mathbb{H}(z))$. In this case, the worker will move to the poaching firm and will use the match value in the current firm $M^{a_i}(h_{m_i}, z, \mathbb{H}(z))$ as the outside value in negotiation. The share of the revenue, and thus the value earned by the worker in the poaching firm will again be determined by Nash Bargaining:

$$\max_{r} \ [V^a(r, h_{m_i}, z', \mathbb{H}(z')) - M^{a_i}(h_{m_i}, z, \mathbb{H}(z))]^{\beta} \, J^{a_i}(r, h_{m_i}, z', \mathbb{H}(z'))^{1-\beta}.$$

   We can solve this problem and recover the worker's value obtained upon hiring in firm z' as $V^a(r', h_{m_i}, z', \mathbb{H}(z')) = M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) + \beta(M^{a_i}(h_{m_i}, z', \mathbb{H}(z')) - M^{a_i}(h_{m_i}, z, \mathbb{H}(z)))$.

3. The match in the current firm z is more valuable than the match in the poaching firm z', but the worker's value is low enough that the poacher's offer is attractive to the worker, $M^{a_i}(h_{m_i}, z', \mathbb{H}(z')) < M^{a_i}(h_{m_i}, z, \mathbb{H}(z))$ and $M^{a_i}(h_{m_i}, z', \mathbb{H}(z')) > V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))$. In this scenario, to prevent the worker from leaving, the current firm negotiates a new division of the match surplus with the worker. The worker now uses the poacher's match value $M^{a_i}(h_{m_i}, z', \mathbb{H}(z'))$ as the outside option in negotiation. Thus, the worker stays in the current firm, and the value of the worker in the current firm is determined by Nash Bargaining:

$$\max_{r} \ [V^a(r, h_{m_i}, z, \mathbb{H}(z)) - M^{a_i}(h_{m_i}, z', \mathbb{H}(z'))]^{\beta} \, J^{a_i}(r', h_{m_i}, z, \mathbb{H}(z))^{1-\beta}.$$

   We can solve this problem and recover the worker's value obtained upon hiring in firm z' as $V^a(r, h_{m_i}, z, \mathbb{H}(z)) = M^{a_i}(h_{m_i}, z', \mathbb{H}(z')) + \beta(M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) - M^{a_i}(h_{m_i}, z', \mathbb{H}(z')))$.

With the model extension of on-the-job search, the match value now incorporates the benefits from job-to-job transitions:

$$
\begin{aligned}
M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) = \max_{l,g} & \Bigg[\underbrace{z h_{m_i}(1-l)}_{\text{output value}} - \underbrace{(1-g)l \times q p_e(h_{m_i})}_{\text{pay for external trainers}} \Bigg] \\
& + \underbrace{\int_{j \in \mathbb{I}(z), j \neq i} \frac{\partial s}{\partial p} \left[p(h_{m_j}, \mathbb{H}(z)) - p(h_{m_j}, \mathbb{H}(z) \backslash \{i\}) \right] \times O^{a_j+1}(h_{m_j}, z, \mathbb{H}'(z)) dj}_{\text{spillover effects to other workers in firm}} \\
& + \underbrace{\rho(1-\kappa) \left(1 - \eta\theta q(\theta) \int \mathbf{1}_{move} \; dG(z') \right) M^{a_i+1,f}(h_{m_i}, z, \mathbb{H}(z))}_{\text{future match value if worker stays at current firm}} \\
& + \underbrace{\rho(1-\kappa)\eta\theta q(\theta) \int \mathbf{1}_{move} \left[M^{a_i+1,f}(h_{m_i}, z, \mathbb{H}(z)) + \beta(M^{a_i+1,f}(h_{m_i}, z', \mathbb{H}(z')) - M^{a_i+1,f}(h_{m_i}, z, \mathbb{H}(z))) \right] dG(z')}_{\text{worker's future value if move to poaching firm}} \\
& + \underbrace{\rho\kappa\mathbb{E} \left[s(h_{m_i}, \mathbb{H}(z)) \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\} + (1 - s(h_{m_i}, \mathbb{H}(z))) \max\{V_U^{a_i+1}(h_{m_i}), V_{TR}^{a_i+1}(h_{m_i})\} \right]}_{\text{worker's future value if separated from current firm}}
\end{aligned} \tag{F.1}
$$

where 1_{move} is an indicator for moving to the poaching firm, which happens if $M^{a_i+1,f}(h_{m_i}, z', \mathbb{H}(z')) > M^{a+1,f}(h_{m_i}, z, \mathbb{H}(z))$.[63] $G(z) = \int_0^z v(y)ds(y)/V$ is the offer distribution. We can compute the optimal learning levels according to Equation (F.1).

The worker's value is now given by:

$$
\begin{aligned}
V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z)) = r & \underbrace{\Big[z h_{m_i}(1-l) - (1-g)l \times q p_e(h_{m_i}) + X^{a_i}(h_{m_i}, z, \mathbb{H}(z)) \Big]}_{\text{current wage}} \\
& + \underbrace{\rho(1-\kappa) \left(1 - \eta\theta q(\theta) \int 1_{neg} G(z') \right) \mathbb{E} \left[s(h_{m_i}, \mathbb{H}(z)) V^{a_i+1}(r, h_{m_i}, z, \mathbb{H}(z)) + (1 - s(h_{m_i}, \mathbb{H}(z))) V^{a+1}(r, h_{m_i} + 1, z, \mathbb{H}(z)) \right]}_{\text{worker's future value if stay at current firm with no wage renegotiation}} \\
& + \underbrace{\rho(1-\kappa)\eta\theta q(\theta) \int \mathbf{1}_{neg} \left[(1-\beta)\{M^{a_i+1,f}(h_{m_i}, z, \mathbb{H}(z)), M^{a_i+1,f}(h_{m_i}, z', \mathbb{H}(z'))\}^- + \beta\{M^{a_i+1,f}(h_{m_i}, z, \mathbb{H}(z)), M^{a_i+1,f}(h_{m_i}, z', \mathbb{H}(z'))\}^+ \right] dG(z')}_{\text{worker's future value if get an attractive offer and negotiate a new wage (either in current or poaching firm)}} \\
& + \underbrace{\rho\kappa\mathbb{E} \left[s(h_{m_i}, \mathbb{H}(z)) \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\} + (1 - s(h_{m_i}, \mathbb{H}(z))) \max\{V_U^{a_i+1}(h_{m_i}), V_{TR}^{a_i+1}(h_{m_i})\} \right]}_{\text{worker's future value if separated from current firm}}
\end{aligned} \tag{F.2}
$$

where 1_{neg} is an indicator for wage renegotiation, which happens if the worker receives an offer that is more attractive than the current match, and thus either renegotiates in the current firm or switches firms.[64] As before, for each bargaining outcome, we can search for

[63] We define $M^{a_i+1,f}(h_{m_i}, z, \mathbb{H}(z)) = \mathbb{E} \left[s(h_{m_i}, \mathbb{H}(z)) M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}(z)) + (1 - s(h_{m_i}, \mathbb{H}(z))) M^{a_i+1}(h_{m_i}, z, \mathbb{H}(z)) \right]$ to save on notation.

[64] We also denote $\{M^{a_i+1,f}(h_{m_i+1}, z, \mathbb{H}(z)), M^{a_i+1,f}(h_{m_i+1}, z', \mathbb{H}(z'))\}^-$ and $\{M^{a_i+1,f}(h_{m_i+1}, z', \mathbb{H}(z')), M^{a_i+1,f}(h_{m_i+1}, z', \mathbb{H}(z'))\}^+$ as the minimum and maximum of $M^{a_i+1,f}(h_{m_i+1}, z, \mathbb{H}(z))$ and $M^{a_i+1,f}(h_{m_i+1}, z', \mathbb{H}(z'))$, respectively.

fraction r in Equation (F.2) that delivers the same value as the worker's bargaining outcome and thus renders the wage.

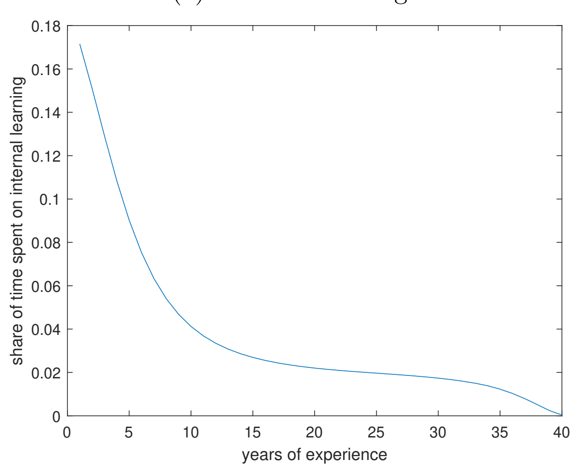
F.1.2 Calibration and quantitative findings

We set the on-the-job search intensity η to 0.3 following Faberman et al. (2017), who find that the average number of offers per month received by employed workers is around 30–40% of that for unemployed people in the US. All other externally calibrated parameters are identical to the baseline values. We still choose the internally calibrated parameters to match the targeted moments in Table 4.2, except that now we target the full average wage growth per quarter at both 0–40 and 0–25 years of experience, not just that coming from human capital. With on-the-job search, our model predicts 0.41% for the average quarterly wage growth at 0–40 years of experience, which is close to the data (0.45%). All other model moments still match the targeted data moments well.[65]

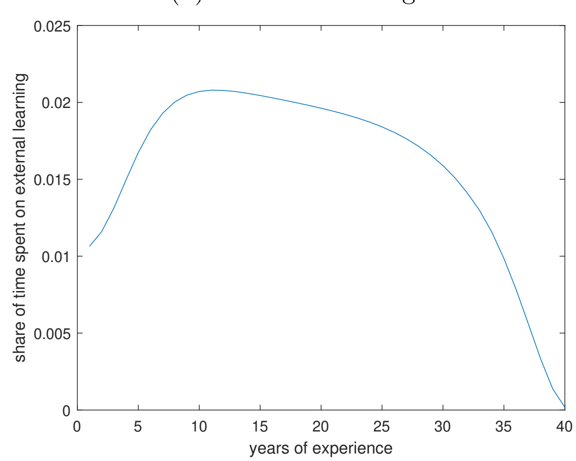
Figures F.1 and F.2 show that the model with on-the-job search matches our empirical findings, namely the lifecycle patterns of learning, and the fact that larger firms provide more variety in learning options.

Figure F.1: Lifecycle patterns of internal and external learning (with on-the-job search)

(a) Internal learning



(b) External learning



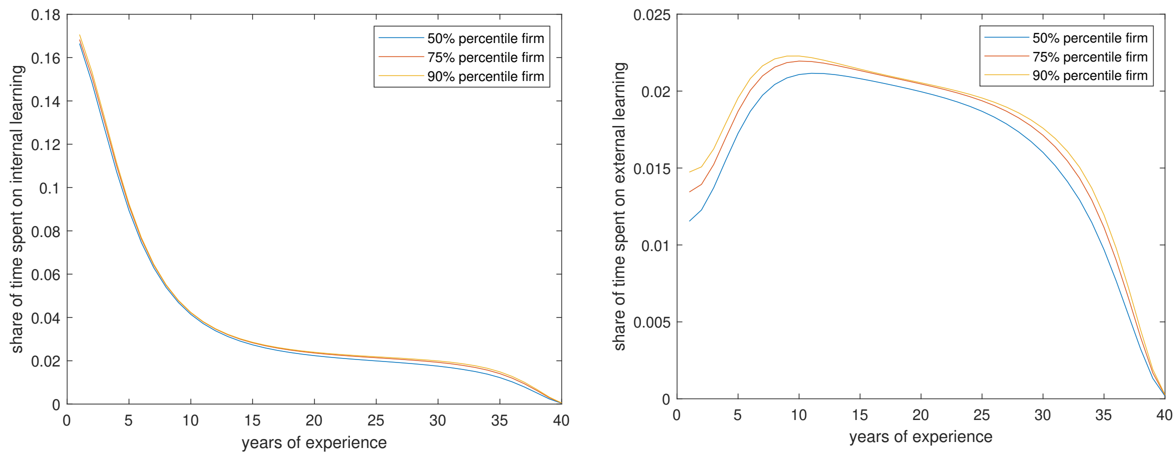
Notes: These figures plot the share of time spent on internal and external learning, namely $l \times g$, and $l \times (1 - g)$, respectively, for workers with different years of experience in the model with on-the-job search.

Panel A in Table F.1 summarizes the average share of time spent on each source of learning along with the average level of human capital in the stationary equilibrium of our extended

[65]Tables with details on the parameter values and comparison between the data and model moments are available upon request.

Figure F.2: Lifecycle patterns of internal and external learning by firm productivity (with on-the-job search)

(a) Internal learning (b) External learning

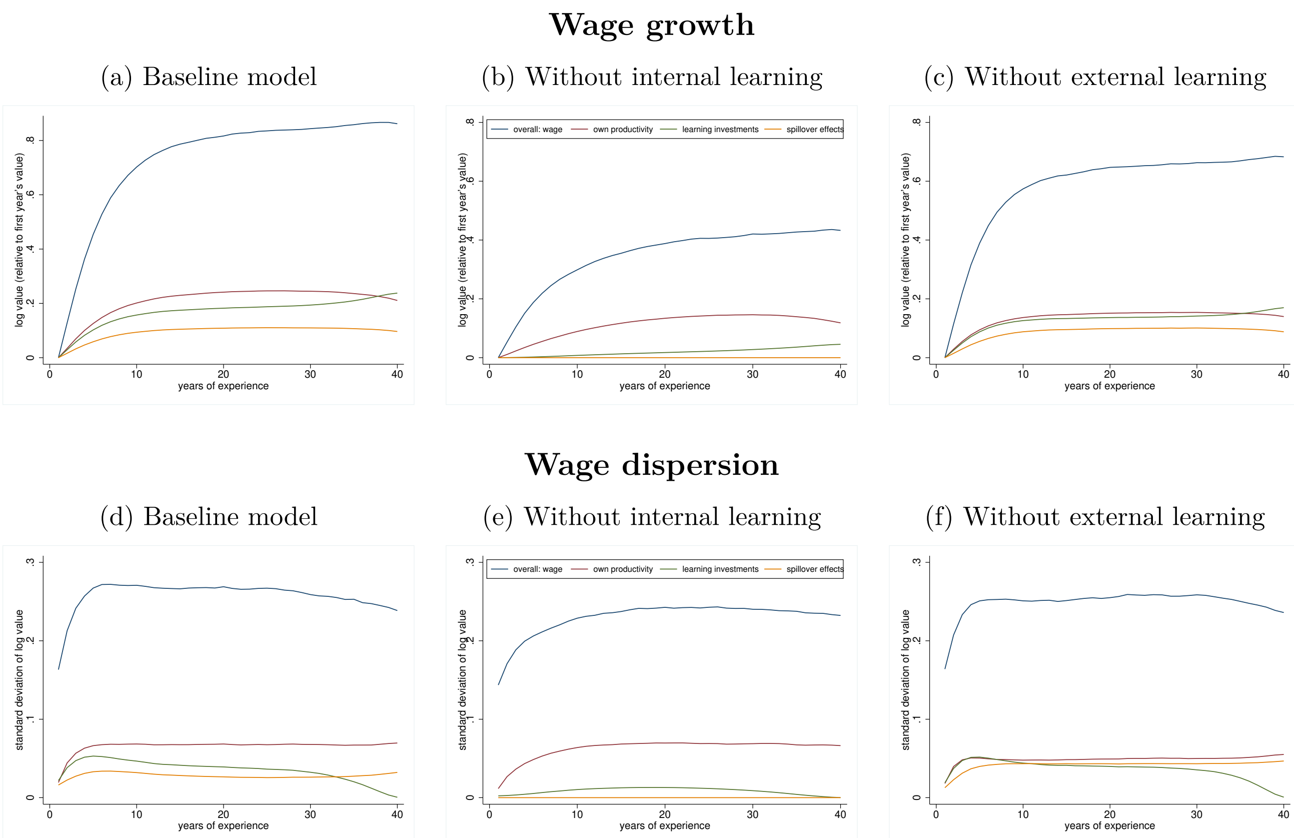


Notes: These figures plot the share of time spent on internal and external learning, namely $l \times g$, and $l \times (1-g)$, respectively, for workers with different years of experience and working in firms with different productivity levels in the model with on-the-job search.

model, and the stationary equilibria of the extended model without different forms of learning. We find that the impacts of each type of learning on human capital are similar to the baseline results. Figure F.3 shows that the predictions about how each type of learning affects wage growth and dispersion in this extended model are similar to those in the baseline model, though the lifetime growth in wages and wage dispersion becomes larger due to the inclusion of job-to-job transitions.

We report the impact of remote work on average human capital and lifecycle wage growth in panel A of Table F.2. In both scenarios of remote work, the decline in average human capital (0.97% and 3.51%) in the extended model is more pronounced than in the baseline model (0.44% and 2.71%). One reason for this disparity is that by hindering coworker learning, remote work disproportionately reduces the incentives of highly productive firms to post job vacancies since these firms have a larger share of skilled workers and thus higher intensity of coworker learning than unproductive firms. This disincentive is more severe in the scenario that includes on-the-job search since through poaching, skilled workers tend to be more concentrated in highly productive firms than in the baseline model. Relatedly, due to the same mechanisms, we also find that the remote-work-driven decline in wage growth from spillover effects is also more severe in the extended model than in the baseline model.

Figure F.3: Lifecycle wage growth and dispersion and their components in the baseline and counterfactual scenarios (with on-the-job search)



Notes: These figures illustrate the lifetime progression of wage growth and wage dispersion and their different components (own productivity, spillover effects to coworkers, and learning investments and associated costs) by plotting the average log value (normalized to the first year's value) and standard deviation of the log value of each of these for workers with different years of experience in the baseline model with on-the-job search and corresponding counterfactual scenarios that shut down internal and external learning, respectively.

F.2 Incorporating time costs for internal trainers

F.2.1 Model extension

In our baseline model, we considered internal learning to be costless for the colleagues workers are learning from (internal trainers). This assumption is based on the qualitative observation that workers who are learning are often included in projects with senior colleagues to observe and learn. However, mentoring might still imply a time cost for internal trainers. To incorporate this, we modify the per-period output value in Equation (3) for worker i in firm z to be

$$zh_{m_i}(1 - l - \zeta \bar{l}_{int}(z, \mathbb{H}(z))F_{int}(h_{m_i}, z, \mathbb{H}(z))).$$

$\bar{l}_{int}(z, \mathbb{H}(z))$ is the average time spent on internal learning by workers in firm z, capturing the average time worker i spends mentoring internal learners given the assumption of random matching. $F_{int}(h_{m_i}, z, \mathbb{H}(z))$ is the share of workers (weighted by each worker's internal learning time) with human capital levels lower than h_{m_i} in firm z, capturing the share of workers in the firm that worker i needs to actively mentor. ζ is a parameter governing the time loss from mentoring for internal trainers.

F.2.2 Calibration and quantitative findings

We calibrate the newly introduced ζ together with other parameters to match targeted moments. To calibrate ζ we introduce a new targeted moment—the share of internal trainers' wage costs relative to learners' wage costs, which is 10.7% according to the US-SEPT data.[66] The other targeted moments are the same as in Table 4.2. The results of our calibration show that $\zeta = 0.18$, indicating that internal trainers incur some time costs, though these are not very large.

Figure F.4 shows that the predictions about how each type of learning affects wage growth and dispersion in this extended model are similar to those in the baseline model. In particular, we find that 12.5% of wage growth over 25 years of experience can be attributed to spillover effects resulting from teaching colleagues, which is slightly lower than the baseline results (14%) but still considerable.

Furthermore, panel B of Table F.1 shows that the contributions of external and internal learning to aggregate human capital are 13.5% in each case, which is very similar to the baseline result.

Finally, panel B of Table F.2 shows that the results of the remote work exercises in this alternative scenario match the baseline results closely. For example, in this scenario, the declines in wage growth in the 20% and 100% remote work exercises are 4.60% and 20.69% respectively, which track closely with the results of 2.86% and 21.29% found in the baseline model. This further indicates that the addition of mentoring time costs does not alter the importance of human capital spillovers from internal learning in the model.[67]

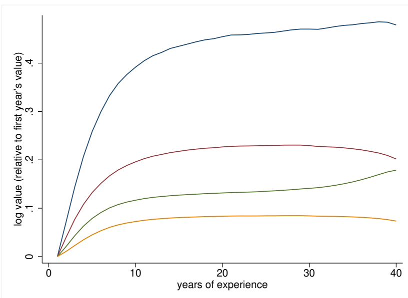
[66] We use the US-SEPT data to obtain information on training expenditures related to the wages and salaries of in-house trainers in 1995. We combine this information with the wage and salary costs of trainees to calculate the proportion of internal trainers' wage costs in relation to learners' wage costs.

[67] The rest of the results in this exercise are very similar to the baseline results and available upon request.

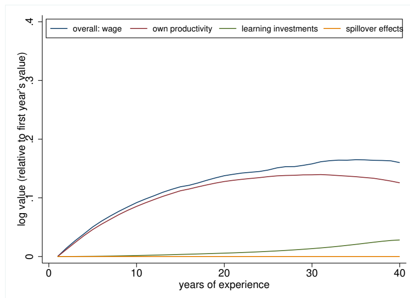
Figure F.4: Components driving wage growth and dispersion over the lifecycle (with time costs for internal trainers)

Wage growth

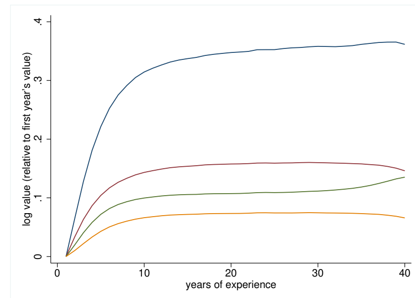
(a) Baseline model



(b) Without internal learning

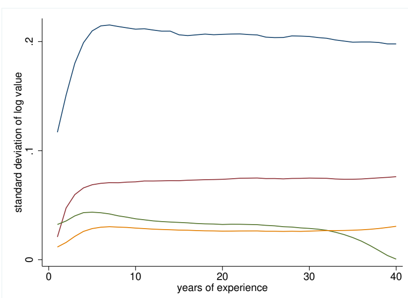


(c) Without external learning

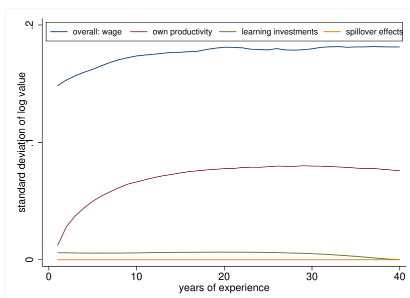


Wage dispersion

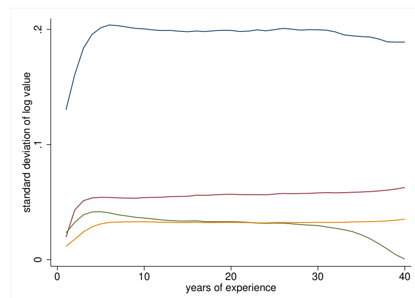
(d) Baseline model



(e) Without internal learning



(f) Without external learning



Notes: These figures illustrate the lifetime progression of wage growth and wage dispersion and their different components (own productivity, spillover effects to coworkers, and learning investments and associated costs) by plotting the average log value (normalized to the first year's value) and standard deviation of the log value of each of these for workers with different years of experience in the baseline model with time costs for internal trainers and corresponding counterfactual scenarios that shut down internal and external learning, respectively.

F.3 Alternative parameterizations

We consider the robustness of our results to considering different values for two key parameters driving wage bargaining and human capital formation in the model.[68]

First, we change workers' bargaining power β to be either 0.3 or 0.7 (the baseline value is 0.5), matching the estimates in Bagger et al. (2014) and Gregory (2019), respectively. In each case, we recalibrate all internally calibrated parameters to jointly match the targeted

[68] For these exercises, we focus on the role of each source of learning in human capital, and the role of remote work in lifecycle wage growth as summarized in Tables F.1 and F.2. The rest of the results in these exercises are very similar to the baseline results and are available upon request.

moments in Table 4.2. As reported in panels C and D of Table F.1, we find that that the contributions of external and internal learning to aggregate human capital are 13–14% in each case. These results are very similar to the baseline results reported in Table 5.1. In addition, panels C and D of Table F.2 also show that the results from the remote work exercises match the baseline results closely. For example, the decline in wage growth in the 20% remote work exercise is 3.22–4.68% when changing β, which tracks closely with the result of 2.86% found in the baseline model.

In the second exercise, we allow our two learning modes to be more highly substitutable by setting the elasticity of substitution to be $\sigma = 5$, instead of estimating it internally (the baseline value is 3.6). Following this, we recalibrate all other internally calibrated parameters to match the targeted moments presented in Table 4.2. The results are very similar to the baseline, with two noteworthy differences. First, as reported in panel E of Table F.1, the importance of internal learning slightly declines in this case to 12%. Second, due to the decreased importance of internal learning, panel E of Table F.2 indicates that the 20% remote work exercise yields a slightly lower decrease (2.32%) in lifecycle wage growth relative to the baseline result (2.86%).

F.4 Role of each learning source in alternative model setups

Table F.1: Learning and human capital in the baseline and counterfactual scenarios in alternative model setups

	Workers' share of time spent on learning		
	External learning	Internal learning	Avg Human Capital
	Panel A: With on-the-job search		
Calibrated economy	1.85%	3.57%	1.42
W/o external learning	0	2.71%	1.21
W/o internal learning	3.29%	0	1.23
	Panel B: Incorporating time costs for internal trainers		
Calibrated economy	1.35%	3.19%	1.41
W/o external learning	0	2.25%	1.22
W/o internal learning	2.51%	0	1.22
	Panel C: Workers' bargaining power $\beta = 0.3$		
Calibrated economy	1.47%	3.32%	1.41
W/o external learning	0	2.29%	1.21
W/o internal learning	2.58%	0	1.21
	Panel D: Workers' bargaining power $\beta = 0.7$		
Calibrated economy	1.52%	3.10%	1.41
W/o external learning	0	2.21%	1.21
W/o internal learning	2.93%	0	1.23
	Panel E: Elasticity of substitution between learning modes $\sigma = 5$		
Calibrated economy	2.05%	2.59%	1.44
W/o external learning	0	1.96%	1.23
W/o internal learning	4.42%	0	1.27

Notes: This table presents the share of time workers spend on internal and external learning along with the average level of human capital in the baseline calibrated model economy of different alternative model setups and the corresponding counterfactual scenarios that shut down internal and external learning, respectively. The alternative model setups considered encompass incorporating on-the-job search (panel A), incorporating time costs for internal trainers (panel B), setting workers' bargaining power to $\beta = 0.3$ (panel C), setting workers' bargaining power to $\beta = 0.7$ (panel D), and setting the elasticity of substitution between learning modes to $\sigma = 0.3$ (panel E).

F.5 Impact of remote work in alternative model setups

Table F.2: Impact of remote work on human capital and wage growth in alternative model setups

		Impact of remote work (rel. to baseline of each scenario)	
	Baseline of scenario	20% remote work	100% remote work
Panel A: With on-the-job search			
Avg human capital	1.42	-0.97%	-3.51%
Human capital gains from internal learning	0.19	-9.53%	-34.47%
Lifecycle wage growth (25 years of experience)	0.94	-4.90%	-18.19%
Wage growth from spillover effects	0.09	-8.57%	-32.38%
Panel B: Incorporating time costs for internal trainers			
Avg human capital	1.41	-0.48%	-2.97%
Human capital gains from internal learning	0.19	-4.96%	-30.69%
Lifecycle wage growth (25 years of experience)	0.51	-4.60%	-20.69%
Wage growth from spillover effects	0.06	-5.62%	-29.21%
Panel C: Workers' bargaining power $\beta = 0.3$			
Avg human capital	1.41	-0.56%	-3.12%
Human capital gains from internal learning	0.2	-5.54%	-30.84%
Lifecycle wage growth (25 years of experience)	0.55	-4.68%	-22.03%
Wage growth from spillover effects	0.07	-5.73%	-29.73%
Panel D: Workers' bargaining power $\beta = 0.7$			
Avg human capital	1.41	-0.47%	-2.98%
Human capital gains from internal learning	0.20	-5.17%	-32.75%
Lifecycle wage growth (25 years of experience)	0.51	-3.22%	-23.09%
Wage growth from spillover effects	0.07	-4.36%	-30.37%
Panel E: Elasticity of substitution between learning modes $\sigma = 5$			
Avg human capital	1.44	-0.42%	-2.27%
Human capital gains from internal learning	0.17	-6.56%	-35.31%
Lifecycle wage growth (25 years of experience)	0.50	-2.32%	-17.57%
Wage growth from spillover effects	0.05	-5.63%	-33.31%

Notes: This table presents the following in both the baseline model of different alternative model setups and the corresponding scenarios with 20% and 100% remote work: the average human capital level, the average human capital gains from internal learning (built by subtracting the average human capital in the scenario where internal learning is shut down from the average human capital in the baseline scenario), the average wage growth after 25 years of experience, and the average wage growth stemming from coworker spillover effects. The remote work results report the changes in the absolute values relative to the corresponding baseline results. The alternative model setups considered encompass incorporating on-the-job search (panel A), incorporating time costs for internal trainers (panel B), setting workers' bargaining power to $\beta = 0.3$ (panel C), setting workers' bargaining power to $\beta = 0.7$ (panel D), and setting the elasticity of substitution between learning modes to $\sigma = 0.3$ (panel E).

G Appendix: Evidence on testable predictions of theory and validation

In this appendix we present evidence on the testable predictions of our theory. We start by presenting evidence matching key predictions on how the incentives to engage in each source of skill acquisition evolve through workers' lifecycles as they accumulate human capital and age. Then, we show evidence on the importance of coworker instruction across occupations in the US.

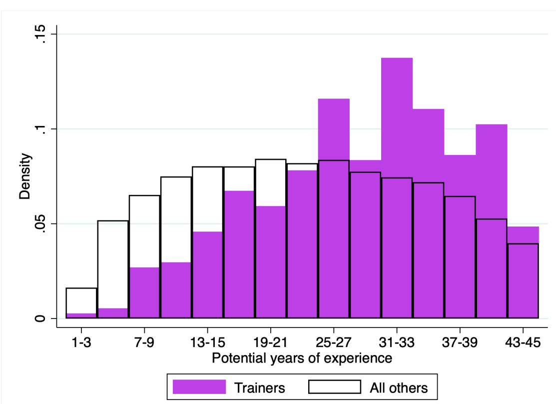
G.1 Evidence on testable predictions of lifecycle theory

G.1.1 Human capital distribution of trainers is left-skewed

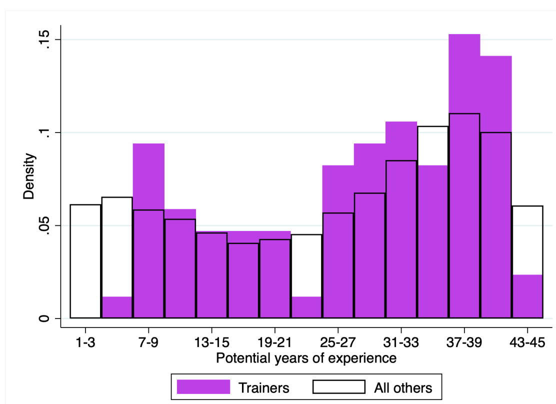
Our theory suggests that the human capital of trainers is higher than that of production workers, motivating mid-career workers to switch to the former in order to continue learning once they have exhausted the learning opportunities within their firm. We provide support for this in Figure G.1, where we plot the histograms of potential experience for trainers and production workers in both Germany and the US. We define trainers as workers engaged in an occupation that involves training, teaching, or instruction activities outside of school and university education. Production workers, on the other hand, capture all other workers outside of trainers. In Appendix H.1.1 we provide further details on the construction of the trainer and production worker variables in the German and US surveys.

Figure G.1: Histograms of potential experience for trainers and production workers

(a) Germany



(b) United States



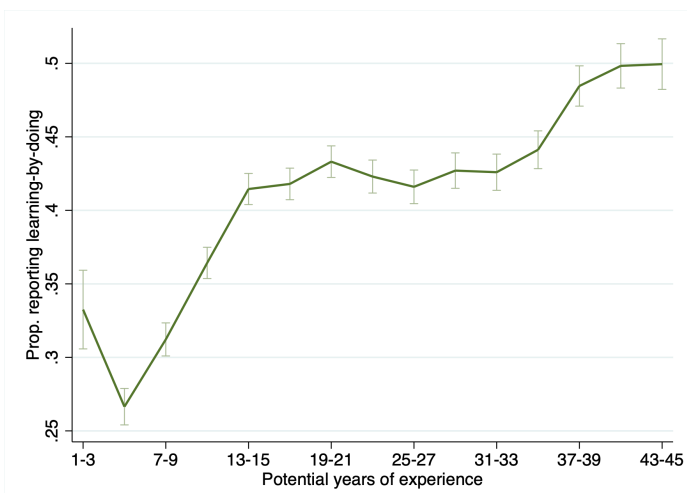
Notes: These figures plot the histograms of potential experience for trainers and all other workers in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have between 1 and 45 years of potential experience in both settings. These results are unweighted for ease of interpretation.

The plots show that the distribution of trainers in both Germany and the US heavily concentrates among higher levels of potential experience relative to that of other workers.[69] In Table H.1 we present the results of quantile regressions at the first, second, and third quartiles of potential years of experience on the trainer variable (where the omitted category is production workers). The results from these regressions indicate that the 25th, 50th and 75th percentiles of potential years of experience for trainers are generally higher (though not always statistically significant) than those of production workers in both settings, even after including controls.

G.1.2 Portion of workers who learn-by-doing rises with human capital

Another key prediction of our model is that the portion of workers who do not make internal and external learning investments, and therefore are more prone to learn-by-doing, rises with age. We provide evidence for this prediction using our German data. We construct a measure of "learning-by-doing" which captures individuals who did not invest in explicit forms of learning to acquire skills for their job, but rather acquired the necessary professional skills by doing the job itself.[70] In Figure G.2 we find that consistent with our model's

Figure G.2: Prevalence of learning-by-doing throughout workers' lifecycles in Germany



Notes: This figure plots the proportion of workers reporting engaging in learning-by-doing across different potential experience bins in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of potential experience. The results are weighted using the observational weights provided in the surveys. 95% confidence intervals are plotted.

[69] In Figure H.1 and Figure H.2 we show that these results hold when we compare the distribution of trainers to the distribution of workers in professional and technical occupations, and to the distributions of internal and external learners as defined in Section 3.

[70] Please see Appendix H.2.1 for details on the contruction of this variable.

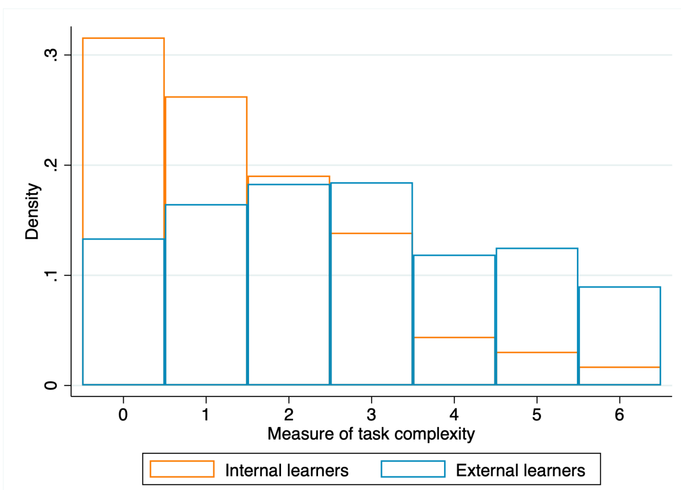
prediction, learning-by-doing increases with potential experience. In Table H.2 we show that the positive correlation between experience and learning-by-doing is statistically significant even after controlling for several demographic and firm-level variables.

G.1.3 Task complexity ranking across different types of workers

Our theory predicts that internal learners have lower human capital levels than external learners, and are thus more easily trained by coworkers who have lower human capital levels than external trainers. We provide a test of these human capital differences between internal and external learners by exploring how the skill content of the tasks performed and the tools used for these tasks vary across each of these two groups of workers.

We rely on information from our German data about the skills workers report using in their jobs. We construct a measure of task complexity by counting how many of the following skills workers use on their jobs: Math and Stats; Foreign Language; Computing; Accounting, Purchasing, Financing and Taxes; Marketing; and Management and Organization.[71] Larger values of this measure imply a higher number of skills used in the job, and thus higher task complexity. In Figure G.3 we show that the distribution of task complexity concentrates more heavily among lower levels for individuals learning internally than those learning externally. We then formally test these distribution differences through quantile regressions of

Figure G.3: Task complexity for internal and external learners in Germany



Notes: These figures plot the histograms of task complexity for internal and external learners in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of potential experience. These results are unweighted for ease of interpretation.

[71]Please see Appendix H.3.1 for details on the construction of these skill variables.

the median of task complexity on the external learning variable (where the omitted category is internal learning) in Table H.3.[72] The results from these regressions indicate that the median of task complexity for external learners is larger than that of internal learners.

In Appendix H.3.2 we provide an additional test of the human capital differences between internal and external learners by exploring the differences in the tools used by each of these two groups of workers. We find that external learners are more likely to use "white-collar" tools than internal learners, while the opposite is true for "blue-collar" tools.

G.2 Evidence on importance of coworker instruction

Our theory and quantitative results predict that wage compensation stemming from coworker spillovers for workers who can act as "teachers" in the firm constitutes a very important component of lifecycle wage growth, and is particularly key for high-skill workers who can effectively teach their colleagues. We now show evidence consistent with this by documenting that tutoring coworkers is an important task in many occupations, and that these tutoring tasks tend to be performed by high-skill workers.

To do this, we use data from the United States Department of Labor's O*NET project, which aims to characterize the tasks pertaining to each occupation, along with the mix of knowledge, skills, and abilities required to perform these tasks. The data includes information about 900 occupations. For each of these occupations, analysts at O*NET give a score characterizing the importance of different work activities, skills, and other descriptors.[73] Please see Appendix H.4 for details about this data.

G.2.1 Tutoring tasks are important

First, we show that coworker tutoring is an important part of the job in many non-teaching occupations.[74] We focus on three work activities and one skill that capture coworker tutoring

[72] We do not perform a quantile regression for other quantiles here, since the measure of task complexity contains only 7 values, and does not have enough variation across groups at the lower and upper ends of the distribution. In addition, we include age fixed effects and education controls in these regressions, and thus do not include potential experience due to high collinearity between potential experience, education, and age.

[73] We use the analysts' database, version 4.0, which only includes data from analysts and should yield a more consistent picture across occupations.

[74] In order to focus on non-teaching occupations, we first drop occupations related to teaching in child-care, preschool, elementary, secondary, and post-secondary programs, along with occupations related to education administration at all of these levels. In addition, we drop occupations that refer to training within firms (such as training and development managers and specialists), and occupations engaging in

tasks within non-teaching occupations: (1) Guiding, Directing, and Motivating Subordinates Activity; (2) Coaching and Developing Others Activity; (3) Training and Teaching Others Activity; and (4) Instructing Skill. Please see Appendix H.4.2 for further details on these activities and skills.

In Columns 1–3 of Table G.1 we present the number and share of non-teaching occupations with a medium or higher score in each of these activities and skill, along with the corresponding average score among these occupations.[75] We find that a sizeable number of non-teaching occupations place high importance on coworker tutoring tasks and skills. For instance, 57 occupations, corresponding to 6.7% of all non-teaching occupations, report a medium or higher score in the importance of activities related to guiding, directing, and motivating subordinates. In Table H.6 we tabulate these occupations, and find that most of them refer to managerial and supervising occupations.[76] This matches the literature suggesting managerial inputs are an important input of on-the-job human capital formation for workers (Burstein and Monge-Naranjo (2007), Luttmer (2014)).

Table G.1: Importance of tutoring activities/skills in non-teaching occupations in the US (information for non-teaching occupations with 3.5+ score in each activity/skill)

	# Occs	Share occs	Avg score	Avg job zone
Importance of guiding, directing, and motivating subordinates activity	57	0.067	4.24	3.97
Importance of coaching and developing others activity	27	0.032	4.22	4.15
Importance of training and teaching others activity	46	0.054	4.12	4.22
Importance of instructing skill	99	0.12	4.13	3.88

Notes: This table illustrates the importance of tutoring activities/skills in the US O*NET data. Columns 1 and 2 report the number of non-teaching occupations with a medium or higher score in each of the considered activities or skills, and the share of these out of the total number of non-teaching occupations, respectively. Column 3 reports the average score for these occupations in each of the considered activities or skills. Column 4 reports the average job zone score for these occupations.

G.2.2 Occupations with tutoring tasks are performed by skilled workers

We now show that these tutoring-heavy occupations are mainly performed by high-skill workers. To do this, we rely on information on the experience and education requirements of each of these occupations. In particular, we use the "Job Zone" information provided in the

community-wide education (such as health educators or sports coaches).

[75]The scores in each activity and skill range from 0 to 7, and as such medium or higher scores correspond to values of 3.5 or higher.

[76]Lists of the occupations reporting a medium or higher score in the importance of the other tutoring activities and skill are available upon request.

O*NET data, which sorts occupations into five bins based on the education, experience, and training required. A higher bin denotes a higher level of skill required for the occupation. Please see Appendix H.4.3 for further details on these job zones and descriptions.

In Column 4 of Table G.1 we present the average job zone for occupations with a medium or higher score in each of the activities and skill of interest. We find that tutoring-heavy occupations tend to be high-skill on average. For instance, the 57 occupations reporting a medium or higher score in the importance of activities related to guiding, directing, and motivating subordinates, have an average job zone score of 3.97, indicating considerable preparation needed in education and/or experience.

H Appendix: Additional information for testable predictions

H.1 Additional information for Appendix G.1.1

H.1.1 Definition of trainer and production worker

- Trainer is constructed as a binary variable that takes a value of one for workers who report an occupation that involves training, teaching, or instruction activities outside of school and university education.
  - German BIBB/BAuA data: For the German data, we define trainers as those having occupations of "other teaching professionals" or "other teaching associate professionals," meaning workers who engage in teaching activities other than those connected with primary, pre-primary, and special education school levels. The specific 3-digit ISCO 1988 codes we use to define trainers are 2359 and 3340.
  - US NHES data: For the US data, we define trainers as those having occupations of "training and development managers," "training and development specialists," or "other education, training, and library workers," meaning training professionals or specialists, and teachers outside of post-secondary, preschool, and kindergarten, elementary and middle school, and secondary and special education. The specific ACS 2000 codes we use to define trainers are 0137, 0650, and 2550.
- Production worker is constructed as a binary variable that takes a value of one for workers who report any occupation outside of being a trainer.

- Professional and technical production worker is constructed as a binary variable that takes a value of one for workers who report any professional or technical occupation outside of being a trainer.[77]

H.1.2 Additional results

Table H.1: Quantile regression of potential experience for trainers vs. production workers

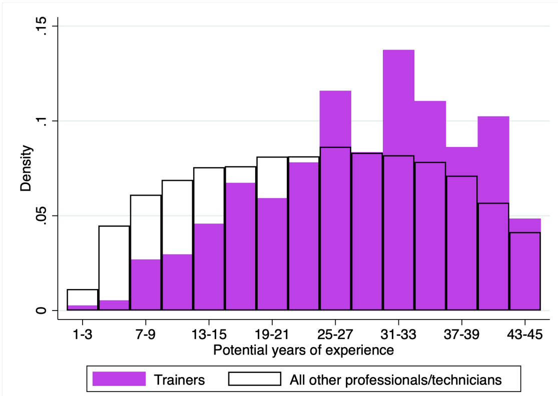
Germany						
	Potential years of experience					
	25th percentile		50th percentile		75th percentile	
Trainer	4***	4***	3	3***	3**	2.5*
	(1.254)	(1.463)	(2.005)	(1.015)	(1.254)	(1.467)
Constant	14***	3***	24***	14***	33***	25***
	(0.0529)	(0.354)	(0.0423)	(0.266)	(0.0529)	(0.422)
Observations	173,639	165,265	173,639	165,265	173,639	165,265
Year FE		Y		Y		Y
Demographic controls		Y		Y		Y
Worker type FE		Y		Y		Y
Firm size FE		Y		Y		Y
United States						
	Potential years of experience					
	25th percentile		50th percentile		75th percentile	
Trainer	4	4	-1	4**	1	3
	(7.557)	(6.560)	(6.044)	(1.593)	(3.782)	(6.761)
Constant	10***	11***	20***	24***	31***	35***
	(0.186)	(0.375)	(0.149)	(0.496)	(0.186)	(0.410)
Observations	29,217	29,217	29,217	29,217	29,217	29,217
Demographic controls		Y		Y		Y
Worker type FE		Y		Y		Y

Notes: This table shows the coefficients from quantile regressions at the first, second, and third quartiles of potential experience on the trainer indicator variable in the German BIBB/BAuA data and US NHES data. The samples encompass individuals who are currently employed, and have between 1 and 45 years of potential experience. Regressions are weighted using the using the observational weights provided in the surveys. *Germany:* Year fixed effects correspond to the year of the survey. Demographic controls include educational attainment level and gender. Worker type categories include laborer, private employee, government employee, self-employed, freelancer, or family caregiver. Firm size is a categorical variable indicating whether the firm where the worker works has fewer than 4 workers, 5–9 workers, 10–49 workers, 50–99 workers, 100–499 workers, 500–999 workers, and 1000 or more workers. Wage controls include the current hourly wage. *US:* Demographic controls include educational attainment level, race, census region, and gender. Worker type categories include private employee, government employee, self-employed, or working without pay. We do not include wage controls, occupation fixed effects, and industry fixed effects in these regressions since trainers and production workers have inherently different wage levels, occupations, and industries. We do not include age fixed effects due to high collinearity between potential experience, education, and age. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

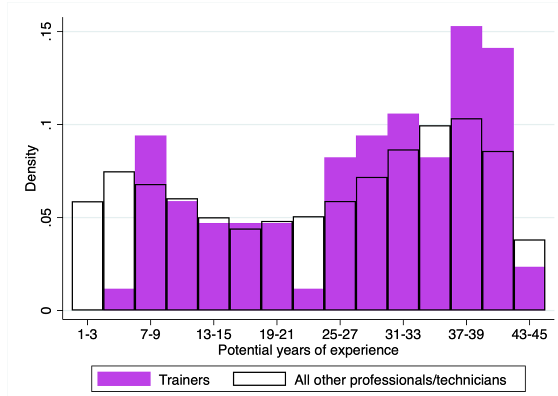
[77] In the German data, professional and technical occupations encompass 3-digit ISCO 1988 codes in the 100s, 200s and 300s. In the US data, professional and technical occupations encompass ACS 2000 codes below 3700.

Figure H.1: Histograms of potential experience for trainers and professional and technical production workers

(a) Germany



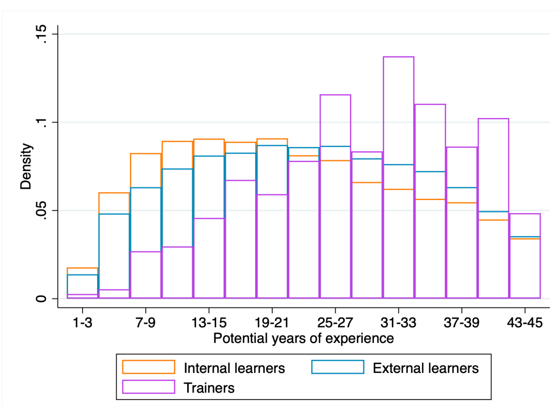
(b) United States



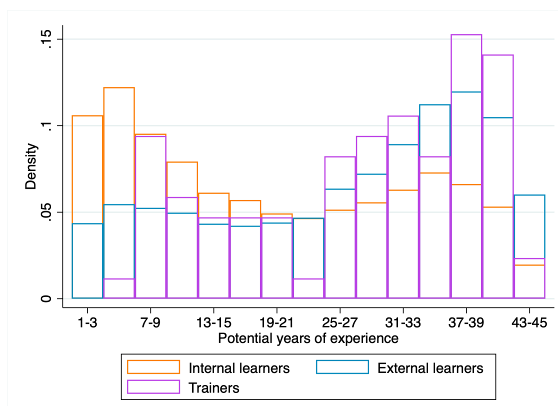
Notes: These figures plot the histograms of potential experience for trainers and professional and technical production workers in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have between 1 and 45 years of potential experience in both settings. These results are unweighted for ease of interpretation.

Figure H.2: Histograms of potential experience for internal and external learners and trainers

(a) Germany



(b) United States



Notes: These figures plot the histograms of potential experience for internal learners, external learners, and trainers in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have between 1 and 45 years of potential experience in both settings. These results are unweighted for ease of interpretation.

H.2 Additional Information for Appendix G.1.2

H.2.1 Definition of "learning-by-doing"

- "Learning-by-doing" is a binary variable that indicates whether the interviewee acquired professional skills by doing the job itself.
  - 1985/1986, 1991/1992, 1998/1999: All of the listed surveys contain questions that determine whether or not the interviewee claims to have acquired professional knowledge/skills by doing his or her job. The 1979 survey does not distinguish between learning-by-doing and internal learning, and thus is not used. The three most recent survey waves in 2005/2006, 2011/2012, and 2017/2018 do not include this information.

H.2.2 Additional results

Table H.2: Correlations between learning-by-doing and potential experience in Germany

	Learning-by-doing		
Potential years of experience	0.0091***	0.0084***	0.0091***
	(0.0007)	(0.0009)	(0.0014)
Potential years of experience^2	-9.87e-05***	-9.18e-05***	-9.82e-05***
	(1.54e-05)	(1.93e-05)	(2.99e-05)
Constant	0.267***	0.150***	0.178***
	(0.00772)	(0.0216)	(0.0303)
Observations	90,536	51,455	21,378
R-squared	0.011	0.123	0.081
Year FE		Y	Y
Demographic controls		Y	Y
Worker type FE		Y	Y
Industry FE		Y	Y
Occupation FE		Y	Y
Firm size FE		Y	Y
Wage controls			Y

Notes: This table shows the coefficients from regressing learning-by-doing on potential experience and potential experience squared in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed, and have between 1 and 45 years of potential experience. Regressions are weighted using the using the observational weights provided in the surveys. *Germany:* Year fixed effects correspond to the year of the survey. Demographic controls include educational attainment level and gender. Worker type categories include laborer, private employee, government employee, self-employed, freelancer, or family caregiver. Industry categories are at the 1-digit level. Occupation categories are at the 2-digit level (ISCO 88). Firm size is a categorical variable indicating whether the firm where the worker works has fewer than 4 workers, 5–9 workers, 10–49 workers, 50–99 workers, 100–499 workers, 500–999 workers, and 1000 or more workers. Wage controls include the current hourly wage. We do not include age fixed effects due to high collinearity between potential experience, education, and age. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

H.3 Additional information for Appendix G.1.3

H.3.1 Construction of task complexity

We construct a measure of task complexity by counting how many complex skills workers report using in their jobs. The waves used to construct this measure encompass 1992, 1999, 2006, 2012, and 2018. Earlier waves do not contain this information. There are six categories of skills, summarized in the following variables:

- Math and statistics is constructed as a binary variable that takes a value of one for workers who report needing knowledge about math and statistics for their current job.
- Foreign language is constructed as a binary variable that takes a value of one for workers who report needing to use a language other than German for their current job.
- Computing is constructed as a binary variable that takes a value of one for workers who report needing computing knowledge for their current job.
- Accounting, purchasing, financing, and taxes is constructed as a binary variable that takes a value of one for workers who report needing accounting, purchasing, financing, tax, or related knowledge for their current job.
- Marketing is constructed as a binary variable that takes a value of one for workers who report needing marketing or related knowledge for their current job.
- Management and organization is constructed as a binary variable that takes a value of one for workers who report needing management and organization knowledge for their current job.

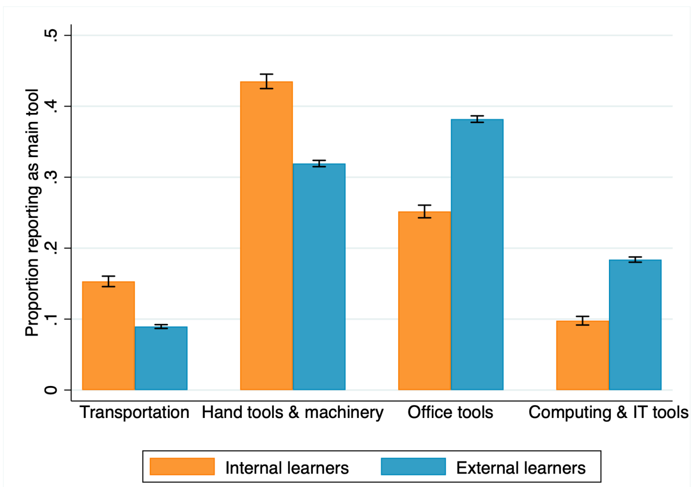
H.3.2 Tool use for internal and external learners

We provide an additional test of the human capital differences between internal and external learners by exploring differences in the tools used by each of these two groups of workers. We build binary variables capturing whether the main tool employed by the worker in her job corresponds to transportation equipment (such as trucks or forklifts), hand tools and machinery (such as hammers, drills or hair dryers), office equipment (such as writing materials, phones, or calculators), or computers and other IT equipment.[78] This tool information provides insights into the attributes of the worker's job, and particularly the skill

[78]Please see Appendix H.3.3 for details on the construction of these tool use variables.

level required, as suggested by DiNardo and Pischke (1997). Specifically, these tool categories separate blue-collar occupations (main tools used are transportation and hand tools) from white-collar occupations (main tools used are office equipment and computers). In Figure H.3 we plot the proportion of external and internal learners who report their main tool to be in each of the four categories above, along with 95% confidence intervals. The plot suggests that external learners are more likely to use "white-collar" tools than internal learners, while the opposite is true for "blue-collar" tools. We formally test the difference in "white-" versus "blue-collar" tool use in Table H.4. The results from these regressions indicate that external learners are more likely to employ "white-collar" tools than internal learners even after controls.

Figure H.3: Main tool use for internal and external learners in Germany



Notes: This figure plots the proportion of internal and external learners reporting that the main tool they employ in their jobs corresponds to transportation equipment, hand tools and machinery, office tools, and computing and IT tools, respectively, in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of potential experience. The results are weighted using the observational weights provided in the surveys.

H.3.3 Construction of job-related tool use

We construct binary variables capturing whether the main tool employed by the worker in her job corresponds to different categories. The waves used to construct these variables are 1979, 1986, 1992, and 1999. Latter waves appear to collect this information, but it is not available in the data files.

We consider four specific tool categories, summarized in the following variables:

- Transportation equipment is constructed as a binary variable that takes a value of one for workers who report that the main tool used in their current jobs corresponds to transportation equipment such as motor vehicles, tractors, snowplows, bulldozers, forklifts, cranes, hoists, rail vehicles, handcarts, etc.
- Hand tools is constructed as a binary variable that takes a value of one for workers who report that the main tool used in their current jobs corresponds to hand tools or machinery such as hammers, screwdrivers, gauges, welding machines, drills, hair dryers, ovens, sewing machines, elevators, etc.
- Office equipment is constructed as a binary variable that takes a value of one for workers who report that the main tool used in their current jobs corresponds to office equipment such as pencils, rulers, stamps, phones, calculators, files, books, copiers, cash registers, etc.
- Computer and other IT Equipment is constructed as a binary variable that takes a value of one for workers who report that the main tool used in their current jobs corresponds to a computer or other IT equipment such as network devices, digital graphics systems, terminals, etc.

H.3.4 Additional results

Table H.3: Quantile regression of task complexity for external learners vs. internal learners in Germany

	Task complexity	
	50th percentile	
External learner	2***	0.584***
	(0.0991)	(0.222)
Constant	1***	3.827***
	(0.0940)	(0.840)
Observations	84,315	29,322
Year FE		Y
Demographic controls		Y
Age FE		Y
Worker type FE		Y
Industry FE		Y
Occupation FE		Y
Firm size FE		Y
Wage controls		Y

Notes: This table shows the coefficients from a quantile regression at the second quartile of task complexity on an external learner indicator variable in the German BIBB/BAuA data. The baseline group is internal learners. The sample encompasses individuals who are currently employed, and have between 1 and 45 years of potential experience. Regressions are weighted using the using the observational weights provided in the surveys. *Germany:* Year fixed effects correspond to the year of the survey. Demographic controls include educational attainment level and gender. Worker type categories include laborer, private employee, government employee, self-employed, freelancer, or family caregiver. Firm size is a categorical variable indicating whether the firm where the worker works has fewer than 4 workers, 5–9 workers, 10–49 workers, 50–99 workers, 100–499 workers, 500–999 workers, and 1000 or more workers. Wage controls include the current hourly wage. We do not include potential experience due to high collinearity between potential experience, education, and age. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table H.4: Regression of white-collar tool use for external learners vs. internal learners in Germany

	White-collar tools		
External learner	0.209***	0.0419***	0.0433***
	(0.00508)	(0.00496)	(0.00845)
Constant	0.373***	0.363***	0.262***
	(0.00458)	(0.0183)	(0.0283)
Observations	87,047	53,163	29,162
R-squared	0.026	0.617	0.623
Year FE		Y	Y
Demographic controls		Y	Y
Age FE		Y	Y
Worker type FE		Y	Y
Industry FE		Y	Y
Occupation FE		Y	Y
Firm size FE		Y	Y
Wage controls			Y

Notes: This table shows the coefficients from regressing task complexity on an external learner indicator variable in the German BIBB/BAuA data. The baseline group is internal learners. The sample encompasses individuals who are currently employed, and have between 1 and 45 years of potential experience. Regressions are weighted using the using the observational weights provided in the surveys. *Germany:* Year fixed effects correspond to the year of the survey. Demographic controls include educational attainment level and gender. Worker type categories include laborer, private employee, government employee, self-employed, freelancer, or family caregiver. Firm size is a categorical variable indicating whether the firm where the worker works has fewer than 4 workers, 5–9 workers, 10–49 workers, 50–99 workers, 100–499 workers, 500–999 workers, and 1000 or more workers. Wage controls include the current hourly wage. We do not include potential experience due to high collinearity between potential experience, education, and age. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

H.4 Additional information for Appendix G.2

H.4.1 O*NET 4.0

The US Department of Labor's Occupational Information Network (O*NET) project aims to characterize the tasks pertaining to each occupation, along with the mix of knowledge, skills, and abilities required to perform these tasks. The data includes information about 900 occupations. The analysts' database, version 4.0, uses data from analysts to characterize these occupations. For each occupation, the analysts give a score characterizing the importance of different skills, knowledge, abilities, etc. Some summary statistics of interest for this data, including the ratings for work activities and skills related to on-the-job tutoring along with education and experience requirements and the prevalence of non-teaching occupations, are presented in Table H.5.

Table H.5: Summary statistics in O*NET data

	Mean	Std. dev.	Median	Min.	Max.	# Obs.
Importance of guiding, directing, and motivating subordinates activity	1.11	1.35	0.4	0	5.66	900
Importance of coaching and developing others activity	1.32	1.17	1	0	6.2	900
Importance of training and teaching others activity	1.46	1.31	1	0	6.16	900
Importance of instructing skill	1.87	1.39	1.33	0	6.33	900
Job zone rating	2.98	1.26	3	1	5	900
Non-teaching occupation	0.949	0.22	1	0	1	900

H.4.2 Work Activities and Skills Relating to On-the-Job Tutoring

In order to characterize the importance of tutoring in non-teaching occupations, we focus on the following work activities and skills, and their corresponding score in O*NET. All of these work activities and skills follow a scale of 0 to 7 ranking their importance to the job.

- Activities:
  - Guiding, directing, and motivating subordinates: Providing guidance and direction to subordinates, including setting performance standards and monitoring performance.
  - Coaching and developing others: Identifying the developmental needs of others and coaching, mentoring, or otherwise helping others to improve their knowledge or skills.
  - Training and teaching others: Identifying the educational needs of others, developing formal educational or training programs or classes, and teaching or instructing others.
- Skills:
  - Instructing: Teaching others how to do something

H.4.3 Job zones and education and experience requirements

We can categorize occupations by skill using the "Job zone" information provided in the O*NET data, which sorts occupations into five bins based on the education, experience and training levels required. A higher bin denotes a higher level of skill required for the occupation. We describe each of these job zones below.

- Job zone 1: Little or no preparation needed
  - Experience: No previous work-related skill, knowledge, or experience is needed for these occupations. For example, a person can become a general office clerk even if he/she has never worked in an office before.
  - Education: These occupations may require a high school diploma or GED certificate. Some may require a formal training course to obtain a license.
  - Training: Employees in these occupations need anywhere from a few days to a few months of training. Usually, an experienced worker could show you how to do the job.
  - Examples: These occupations involve following instructions and helping others. Examples include bus drivers, forest and conservation workers, general office clerks, home health aides, and waiters/waitresses.
- Job zone 2: Some preparation needed
  - Experience: Some previous work-related skill, knowledge, or experience may be helpful in these occupations, but usually is not needed. For example, a drywall installer might benefit from experience installing drywall, but an inexperienced person could still learn to be an installer with little difficulty.
  - Education: These occupations usually require a high school diploma and may require some vocational training or job-related course work. In some cases, an associate's or bachelor's degree could be needed.
  - Training: Employees in these occupations need anywhere from a few months to one year of working with experienced employees.
  - Examples: These occupations often involve using your knowledge and skills to help others. Examples include drywall installers, fire inspectors, flight attendants, pharmacy technicians, salespersons (retail), and tellers.

- Job zone 3: Medium preparation needed
  - Experience: Previous work-related skill, knowledge, or experience is required for these occupations. For example, an electrician must have completed three or four years of apprenticeship or several years of vocational training, and often must have passed a licensing exam, in order to perform the job.
  - Education: Most occupations in this zone require training in vocational schools, related on-the-job experience, or an associate's degree. Some may require a bachelor's degree.
  - Training: Employees in these occupations usually need one or two years of training involving both on-the-job experience and informal training with experienced workers.
  - Examples: These occupations usually involve using communication and organizational skills to coordinate, supervise, manage, or train others to accomplish goals. Examples include dental assistants, electricians, fish and game wardens, legal secretaries, personnel recruiters, and recreation workers.
- Job zone 4: Considerable preparation needed
  - Experience: A minimum of two to four years of work-related skill, knowledge, or experience is needed for these occupations. For example, an accountant must complete four years of college and work for several years in accounting to be considered qualified.
  - Education: Most of these occupations require a four-year bachelor's degree, but some do not.
  - Training: Employees in these occupations usually need several years of work-related experience, on-the-job training, and/or vocational training.
  - Examples: Many of these occupations involve coordinating, supervising, managing, or training others. Examples include accountants, chefs and head cooks, computer programmers, historians, pharmacists, and police detectives.
- Job zone 5: Extensive preparation needed
  - Experience: Extensive skill, knowledge, and experience are needed for these occupations. Many require more than five years of experience. For example, surgeons

must complete four years of college and an additional five to seven years of specialized medical training to be able to do their job.

– Education: A bachelor's degree is the minimum formal education required for these occupations. However, many also require graduate school. For example, they may require a master's degree, and some require a Ph.D., M.D., or J.D. (law degree).

– Training: Employees may need some on-the-job training, but most of these occupations assume that the person will already have the required skills, knowledge, work-related experience, and/or training.

– Examples: These occupations often involve coordinating, training, supervising, or managing the activities of others to accomplish goals. Very advanced communication and organizational skills are required. Examples include athletic trainers, lawyers, managing editors, physicists, social psychologists, and surgeons.

H.4.4 Additional results

Table H.6: Non-teaching occupations with a medium or higher rating for the importance of guiding, directing, and motivating subordinates activity

Administrative services managers
Advertising and promotions managers
Agricultural crop farm managers
Agricultural engineers
Chemical engineers
Clinical psychologists
Compensation and benefits managers
Computer programmers
Computer and information systems managers
Construction managers
Directors- stage, motion pictures, television, and radio
Electrical drafters
Engineering managers
Financial managers, branch or department

Continued on next page

First-line supervisors and manager/supervisors- construction trades workers
First-line supervisors, administrative support
First-line supervisors, customer service
First-line supervisors/managers of mechanics, installers, and repairers
First-line supervisors/managers of non-retail sales workers
First-line supervisors/managers of police and detectives
First-line supervisors/managers of production and operating workers
First-line supervisors/managers of retail sales workers
First-line supervisors/managers of transportation and material-moving machine and vehicle operators
Fish hatchery managers
Food service managers
Forest fire fighting and prevention supervisors
Gaming managers
Gaming supervisors
Government service executives
Housekeeping supervisors
Human resources managers
Industrial engineers
Industrial production managers
Lawn service managers
Lodging managers
Management analysts
Mapping technicians
Marketing managers
Mates- ship, boat, and barge
Medical and health services managers
Meeting and convention planners
Mining and geological engineers, including mining safety engineers
Natural sciences managers
Nursery and greenhouse managers
Physical therapists

Continued on next page

Table H.6 – *Continued from previous page*

Postmasters and mail superintendents
Private sector executives
Producers
Program directors
Recreation workers
Sales managers
Social and community service managers
Storage and distribution managers
Technical directors/managers
Transportation managers
Treasurers, controllers, and chief financial officers
Veterinarians