University of Connecticut

Department of Economics Working Paper Series

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

by

Xiao Ma
Peking University

Alejandro Nakab
Universidad Torcuato Di Tella

Daniela Vidart
University of Connecticut

Working Paper 2022-11R4
April 2022, Revised July 2022, Revised August 2023
Revised June 2024, Revised December 2024

365 Fairfield Way, Unit 1063
Storrs, CT 06269-1063
Phone: (860) 486-3022
Fax: (860) 486-4463
http://www.econ.uconn.edu/

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*

December 2024

Abstract

How do the sources of worker learning change over the lifecycle, and how do these changes impact human capital and wages? Using data from Germany and the US, we document that workers' reliance on internal learning (learning from coworkers) decreases with experience, while external learning (on-the-job training) follows an inverted U-shape. Based on this evidence, we develop a search model featuring multiple learning sources whose benefits evolve as workers age and accumulate human capital. Quantitative results indicate that the interaction between internal and external learning is key to understanding lifecycle wage dynamics and the effects of learning policies and shocks.

Keywords: On-the-job learning; Human capital accumulation; Lifecycle wage growth

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, Gregor Jarosch, 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, Universidad 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, University of Arizona, Harvard, UCSB, University of Houston, University of Toronto, Boston College, Clemson, UCSB-LAEF Labor Markets Conference, CMU, CICM, SITE: Frontiers of Macroeconomic Research, BU, UC Berkeley, St. Louis Fed – Washington Univ. Macro-Labor Conference, and Duke.

1 Introduction

Since Becker's seminal work in 1962 (Becker, 1962), the economics literature has increasingly emphasized the role of on-the-job learning in shaping lifecycle wage dynamics (Rubinstein and Weiss, 2006). Studies have identified several key drivers of on-the-job human capital acquisition, including on-the-job training (Acemoglu, 1997; Acemoglu and Pischke, 1998; Moen and Rosén, 2004; Ma, Nakab, and Vidart, 2024), learning-by-doing (Bagger, Fontaine, Postel-Vinay, and Robin, 2014; Gregory, 2019), and coworker interactions (Nix, 2017; Akcigit, Caicedo, Miguelez, Stantcheva, and Sterzi, 2018; Jarosch, Oberfield, and Rossi-Hansberg, 2021; Herkenhoff, Lise, Menzio, and Phillips, 2024).[1] However, most studies have examined these inputs in isolation, overlooking key interactions between them. This paper seeks to fill this gap by exploring how two primary sources of learning interact and collectively influence lifetime human capital and wage dynamics: internal learning, which captures learning from colleagues within the firm; and external learning, which captures on-the-job training from external sources.[2]

Distinguishing between internal and external learning and examining their combined impact is important for several reasons. First, since these learning sources draw from different knowledge pools separated by the boundary of the firm, their relevance varies with worker and firm characteristics. Internal learning, which largely comes from day-to-day workplace interactions, is common in many firms, but its benefits depend heavily on the quality of the firm's workforce. In contrast, external learning involves additional steps, such as hiring external trainers, but its benefits are tied to broader societal knowledge. This has important implications for the design of policies aimed at enriching worker learning, and for the understanding of shocks that affect these learning sources asymmetrically. A notable example of such a shock, which we explore in this paper, is the post-pandemic surge in remote work, which has disrupted face-to-face coworker interactions and affected internal learning. Second, these learning sources provide additional incentives for human capital acquisition beyond usual productivity gains. Internal learning allows knowledgeable workers to be compensated for generating valuable spillovers, while external learning equips workers with new knowledge enhancing their ability to create such spillovers. We show in this paper that these

[1]Other factors explored include formal schooling (Ben-Porath, 1967), knowledge hierarchies (Garicano, 2000; Garicano and Rossi-Hansberg, 2004, 2006; Caicedo, Lucas, and Rossi-Hansberg, 2019), materials (Manuelli and Seshadri, 2014), and managerial inputs (Burstein and Monge-Naranjo, 2009; Luttmer, 2014).

[2]Our model also incorporates learning-by-doing, another key component of on-the-job human capital accumulation. However, given that this type of learning is costless and often indistinguishable from work, we do not focus on it extensively in our data analysis.

non-own-productivity aspects of human capital are central to lifecycle wage growth, and that the interaction between internal and external learning is key to their impact.

In the empirical section of the paper, we document two novel facts that underscore the importance of internal and external learning for both firms and workers. First, using firm survey data from Europe, we show that both learning sources are widely provided by firms, and that larger firms offer more diverse learning opportunities by allowing workers to engage in both internal and external learning. Second, using detailed worker qualification data from Germany and the US, we show that both learning sources are significant for workers, but follow distinct lifecycle patterns: internal learning decreases with experience, while external learning exhibits an inverted U-shape over experience. These patterns are robust to alternate definitions of experience, as well as controlling for industry, occupation, and demographics, and decomposing the data across several dimensions such as education level, gender, and firm size.[3] We then review the literature examining the impacts of internal and external learning on human capital and wages, finding consistent and significant positive effects of both learning sources on worker productivity and wages across various settings and time periods.

To shed light on our empirical findings and assess the role of internal and external learning in lifecycle wage dynamics, we build a model with a novel two-source learning framework. The model follows an overlapping generations structure with a training sector and a final-good production sector, which workers can move between. The training sector is frictionless and employs trainers who provide external learning services, while the production sector features a frictional labor market with firm heterogeneity in productivity, where firms and workers match through random search. Upon matching, workers and firms in the production sector engage in Nash bargaining to determine compensation and jointly decide to allocate time for internal and external learning to maximize the match value. Internal learning arises from interactions with higher-skilled coworkers, while external learning arises from interactions with higher-skilled trainers. Both learning modes involve an opportunity cost due to forgone production, but external learning incurs an additional cost from the purchase of training services. Learning enhances human capital, which follows a ladder structure and determines labor productivity in both sectors.

We calibrate the model to the US economy and show that the stationary equilibrium repli-

[3]We provide further evidence supporting the first finding in the US using aggregate data from the Survey of Employer-Provided Training and the second finding in other OECD countries using data from the Program for the International Assessment of Adult Competencies (PIAAC).

cates the empirical lifecycle learning patterns we document. This arises because, as workers age and acquire human capital, their position within the firm's human capital distribution shifts, shaping the benefits derived from each learning source. In particular, and consistent with our empirical findings, young workers primarily rely on internal learning, as it is relatively cheap and the probability of encountering coworkers with higher human capital is high. However, as workers climb the human capital ladder, the availability of coworkers with higher human capital decreases, prompting a shift toward external learning, since trainers have higher average human capital than production workers. Eventually, as workers continue to age, the opportunity cost of learning increases, and the benefits decline due to a shorter remaining work horizon, resulting in a reduction in external learning.[4]

The calibrated model also highlights the importance of firms' learning environments for human capital formation. Across all levels of human capital, workers in more productive firms allocate more time to both internal and external learning. This matches our empirical findings and is consistent with the literature showing that workers in more productive firms acquire skills at faster rates (Engbom, 2017; Arellano-Bover, 2020; Arellano-Bover and Saltiel, 2023). More productive firms foster higher levels of learning due to greater returns to skill acquisition, driven by the supermodularity of the production function, improved learning efficiency, and a better pool of coworkers, which emerges from the more favorable learning environments in these firms.

To assess the role of internal and external learning in lifecycle human capital and wage dynamics, we conduct counterfactual analyses by successively disabling each source of learning in our model. We highlight two main findings. First, removing external learning decreases workers' average human capital by 7%, while removing internal learning reduces it by 11%, indicating that internal learning is 1.57 times more important than external learning for overall lifecycle human capital. However, the relative importance of the two sources shifts over the lifecycle: internal learning is more important to human capital formation early on, whereas external learning becomes more critical as workers age. Second, incorporating the non-own productivity dimension of human capital—reflecting the benefits of human capital acquisition beyond the increase in the worker's own output—is crucial for understanding wage dynamics. Specifically, we find that changes in learning costs and compensation from

[4]We also provide empirical evidence supporting key predictions from this lifecycle theory. For instance, the model predicts that trainers possess higher average human capital than production workers, since the returns to human capital are higher in the training sector. We confirm this empirically by documenting that trainers have higher average human capital levels than production workers in both the US and Germany.

coworker spillovers account for 48% of wage gains from human capital over 25 years of experience. Moreover, both internal and external learning are important to this process: both influence learning costs; and internal learning allows knowledgeable workers to be compensated for fostering valuable spillovers among colleagues, while external learning equips workers with new skills that enhance their ability to create these beneficial spillovers.

We further underscore the importance of considering both internal and external learning by using the model to evaluate the effects of remote work and training subsidies on human capital and wage dynamics. To assess the effects of remote work, we draw on Barrero, Bloom, and David (2021), who document an increase in remote work days from 5% pre-pandemic to 20% post-pandemic, and Emanuel, Harrington, and Pallais (2023), who find a 15% reduction in colleagues' feedback following the shift to remote work, indicating a decline in internal learning efficiency. The post-pandemic rise in remote work leads to a 0.45% drop in average human capital and a 5.05% decline in wage growth for workers with 0–5 years of experience. However, as young workers adapt to remote work by increasing external learning, these losses would be overstated if external learning were not accounted for in the model.

Motivated by the widespread practice of subsidizing formal training expenses, we then examine the effects of a 50% subsidy for external learning, the median policy rate across US states. This subsidy raises average human capital by 6.34% and increases wage growth by 32.93% for workers with 0–5 years of experience. These gains stem not only from the reduced cost of external learning, but also from an improved human capital distribution within firms, as more workers have access to highly skilled trainers. However, these effects would also be overstated if internal learning were omitted, as the exclusive reliance on external learning would amplify the subsidy's impact. Together, these exercises show that focusing exclusively on one form of learning skews the impact of shocks and policies that asymmetrically affect internal and external learning, as it overlooks workers' capacity to reallocate learning time across sources.

The paper is organized as follows. In Section 2, we conduct a literature review. In Section 3, we describe the data and present the empirical findings. We present the model, its calibration, and equilibrium properties in Section 4. In Section 5, we present our counterfactual results assessing the significance of internal and external learning for human capital and wage dynamics. We evaluate the impact of remote work and training subsidies 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 theory features a novel two-source learning framework that jointly considers internal and external learning, thus relating to various strands of the literature on on-the-job human capital formation. First, by incorporating internal learning, our paper connects to the literature on knowledge diffusion within coworker and production teams (Garicano, 2000; Garicano and Rossi-Hansberg, 2004, 2006; Azoulay, Wang, and Zivin, 2010; Luttmer, 2014; Nix, 2017; Akcigit et al., 2018; Caicedo et al., 2019; Jarosch et al., 2021; Emanuel et al., 2023; Wallskog, 2023; Herkenhoff et al., 2024).[5] Our work is particularly related to Akcigit et al. (2018), who develop a model where inventors learn through both interactions and an exogenous external source. Similarly, our model integrates learning from both coworkers and external sources, but we define the firm's boundary as the key factor distinguishing internal from external learning, and endogenize both choices. Additionally, we focus on workers in general rather than inventors, and emphasize how shifts in a worker's position within the firm's human capital distribution can explain the observed lifecycle patterns in human capital acquisition. Second, by incorporating external learning, our paper builds on the literature on general training investments first proposed by Becker (1964), and expanded by Acemoglu (1997), Acemoglu and Pischke (1998), Acemoglu and Pischke (1999), Autor (2001), and Moen and Rosén (2004). We contribute to this literature by showing empirically and theoretically that external learning is particularly important for mid-career workers, and by explicitly modeling the training sector and the market for external learning services.

Our paper also relates to the literature on the impacts of on-the-job human capital on lifecycle earnings (Rubinstein and Weiss, 2006; Barlevy, 2008; Lucas, 2009; Yamaguchi, 2010; Burdett, Carrillo-Tudela, and Coles, 2011; Huggett, Ventura, and Yaron, 2011; Bowlus and Liu, 2013; Bagger et al., 2014; Gregory, 2019; Karahan, Ozkan, and Song, 2022). While several studies focus on disentangling the roles of human capital, search dynamics, and shocks on earnings growth, we instead focus on the contributions of internal and external learning. In addition, by explicitly considering the time and monetary costs of internal and external learning building on the seminal literature on the tradeoff between learning and work (Ben-Porath, 1967; Heckman, 1976; Rosen, 1976), our work contrasts with many of these papers, which model on-the-job human capital as exogenous or a by-product of work.

[5] More broadly, this paper also relates to the literature on knowledge diffusion across the broader population (Jovanovic, 2014; Lucas and Moll, 2014; Perla and Tonetti, 2014; de la Croix, Doepke, and Mokyr, 2016; Benhabib, Perla, and Tonetti, 2021).

By examining how different learning sources shape 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, Laun, Meghir, and Pistaferri, 2021; Jarosch, 2021; Engbom, Moser, and Sauermann, 2022; Arellano-Bover and Saltiel, 2023). This literature shows that there is a substantial heterogeneity in human capital acquisition across firms, and that this heterogeneity is an important determinant of lifecycle earnings dynamics. However, the drivers of firms' learning environments are still poorly understood. We contribute to this literature by providing empirical and theoretical support for a key driver: the provision of internal and external learning opportunities.

Finally, our paper also relates to the extensive labor literature on the impacts of on-the-job learning opportunities on workers' earnings (see Heckman, Lalonde, and Smith, 1999; Kluve, 2010; McKenzie, 2017; Card, Kluve, and Weber, 2018; What Works - Centre for Local Economic Growth, 2016, for reviews),[6] and particularly to the literature showing that the productivity and earnings gains from 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 this literature is between coworker-based and classroom-based on-the-job learning, which broadly align with our internal and external learning categories. We contribute to this literature by suggesting that the effectiveness of these learning opportunities depends on worker characteristics, such as experience and human capital levels.

3 Data and empirical findings

In this section, we first present two novel facts that highlight the importance of internal and external learning for firms and workers. We then review the literature examining the impacts of internal and external learning on human capital and wages.

3.1 Data

To document our findings, we use firm- and worker-level surveys from the EU, the US, Germany, and several OECD countries. Table 3.1 provides brief descriptions of each data source, along with links to the appendices for more detailed information.

[6]Several studies document significant effects of internal and external learning on workers' human capital accumulation and earnings. In Section 3.4, we provide a summary of this evidence.

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 $\sim$ 95,000 firms per wave, most of them with 5+ workers	Appendix B.1
US Survey of Employer-Provided Training (US-SEPT)	United States 1995	Firm- and worker-level survey focusing on training investments and related costs in establishments with 50+ workers	Single cross-section $\sim$ 1,000 establishments & 1,000 workers	Appendix B.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 $\sim$ 25,000 workers per wave, excludes apprentices	Appendix B.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 $\sim$ 48,000 adults	Appendix B.4
OECD Program for the International Assessment of Adult Competencies (PIAAC)	40 OECD countries 2011–2017 (countries surveyed in different years)	Worker-level survey focusing on learning investments and skills	Single cross-section per country $\sim$ 230,000 adults total	Appendix B.5

3.2 Fact 1: Larger firms provide more learning options

First, using the EU-CVT data, we document the importance of both internal and external learning for firms and analyze how the availability of each type varies with firm size. We distinguish between internal and external learning by tracking where Continuing Vocational Training (CVT) activities take place and who conducts them. 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 occur inside firms and are conducted by internal trainers from those that take place outside firms or are conducted by external trainers. We consider "internal CVT courses" as reflecting internal learning and "external CVT courses" as reflecting external learning. Additionally, the survey measures "other types of CVT activities," which include four categories: 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 categorize participation in conferences and lectures as external learning, while the other activities are considered internal learning. Table B.1 shows that a large portion of firms in each country offer "internal CVT courses," "external CVT courses," and "other types of CVT activities," indicating that both internal and external learning are prevalent across firms.

In Table 3.2, we examine how the prevalence of internal and external learning opportunities varies across firms of different sizes. We find that larger firms (100+ workers) are more likely (71%) to offer both internal and external learning opportunities compared to smaller firms with 5–19 (32%) or 20–99 (46%) workers. Notably, the differences in the share of firms offering a single type of learning are much smaller across firm sizes, suggesting that larger firms tend to offer a broader range of learning opportunities, encompassing both internal and external learning. These patterns align with the findings of Engbom (2017), Arellano-Bover (2020), and Arellano-Bover and Saltiel (2023), which show that workers in more productive firms acquire skills more quickly, and Gregory (2019), who emphasizes the importance of having diverse learning options for firms' learning environments.

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

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

We assess the robustness of these patterns in Appendix C. In Table C.1, we show that the positive correlation between firm size and learning opportunities is robust to controlling for industry, socioeconomic, and country-year fixed effects. Moreover, in Table C.2, we show that workers in larger firms dedicate more hours to both types of learning on average.[7] Finally, in Table C.4, we replicate our findings using the US-SEPT data. This table shows that both the share of employees exposed to informal and formal training—corresponding roughly (though not perfectly) to internal and external learning, respectively[8]—along with the number of hours employees spend on each activity, generally increase with firm size. However, caution is warranted when interpreting these findings, as they are based on aggregate statistics.

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

Using the German BIBB/BAuA and US NHES data, we now explore the significance of the two learning sources for workers and how this evolves over their lifetimes. We construct measures of internal and external learning reflecting recent engagement in learning from coworkers or superiors, or from entities outside the current firm, respectively.[9] In the German data, internal learning indicates whether an individual has acquired the skills or knowledge necessary to complete their current tasks through coworkers or superiors, whereas external learning indicates whether an individual received external on-the-job training in the previous 2–5 years, or acquired the skills or knowledge necessary to complete their current tasks through external training. In the US data, internal learning captures workers who reported receiving instruction or training from a coworker or supervisor in their most recent work-experience program, defined as a job with learning attributes, whereas external learning captures workers who reported taking classes or training from a company, association, union, or private instructor in their most recent work-experience program, or ever earned a training certificate from an employment-related training program.[10] In Appendix B.3 and Appendix B.4, we provide further details on the questions and answers used to construct each variable in the German and US surveys, respectively.[11]

[7] This positive correlation between learning hours and firm size is also robust to controlling for industry, socioeconomic, and country-year fixed effects (see Table C.3).

[8] In particular, formal training is defined as planned, structured training with a defined curriculum (such as classes or seminars), while informal training refers to unstructured, unplanned learning (e.g., learning from a coworker or supervisor).

[9] In principle, formal schooling also fits this characterization of external learning, as it involves knowledge from outside the firm. However, since less than 10% of adult education in the EU involves formal schooling (with over 90% being on-the-job learning) (Ma et al., 2024), we exclude it when defining external learning.

[10] Specifically, a work-experience program encompasses jobs with learning attributes such as internships, co-ops, or apprenticeships. Approximately 25% of the US sample reported involvement in such programs. In Figure D.1, we show that our results are robust to limiting the sample to individuals reporting participation in a work-experience program, and to decomposing across learning components involving a "work-experience" program and those that do not.

[11] The measures of both internal and external learning primarily capture the flow (rather than the stock) of learning investments, by capturing skill acquisition in the current job or recent work-experience programs,

We construct potential years of work experience for both Germany and the US using age and educational level: $Potential\ experience = Age - Years\ of\ \ schooling - 6$. Our sample includes employed individuals with 1 to 45 years of potential experience, as the number of observations outside this range is small. Summary statistics in Table B.2 show that the shares of workers reporting each of the two sources of learning are sizeable in both settings. In Germany, 31% of workers report internal learning and 68% report external learning, while in the US, these figures are 23% and 44%, respectively.

We now document the prevalence of each learning source over the lifecycle. In Figure 3.1, we show how workers' engagement in internal and external learning changes with potential experience in Germany and the US. In both settings, the prevalence of internal learning decreases with potential experience, while external learning follows an inverted U-shape. These patterns are robust to several alternate specifications[12] and to controlling for demographic variables, firm characteristics, occupation, and industry fixed effects (see Table D.1). Furthermore, we find that the results hold when using workers' tenure instead of potential experience (see Figure D.2). This suggests that age alone does not explain these patterns, highlighting the importance of work experience and human capital. This is corroborated when we further explore the correlations between tenure and internal and external learning in Table D.2, since the observed patterns hold after including age fixed effects.

In Figure D.10, we extend our analysis to other OECD countries using the PIAAC data. These scatterplots pool data from all countries in the survey and depict the proportion of workers reporting internal and external learning across different experience bins, after controlling for country fixed effects.[13] The findings mirror the lifecycle patterns documented

which evolve as workers progress in their careers. An exception is a component of external learning in the US, which includes ever earning a training certificate. This component still provides insight into external learning flows: a sharp increase in this variable at certain ages indicates larger learning flows. In panel (c) of Figure D.1, we show that the inverted U-shape pattern for external learning persists even when excluding this component, although with more noise.

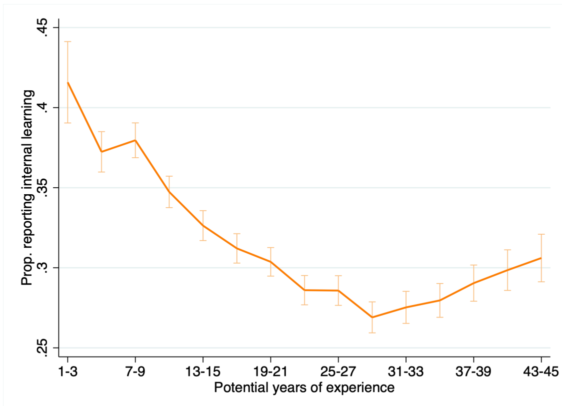
[12]These alternate specifications, detailed in Appendix D.1.1, include using age as an alternative measure of experience, and decomposing the data by one-year experience bins, gender, educational level, survey wave (Germany), cohort (Germany), and firm size (Germany). We also examine the correlation between internal and external learning, finding a negative correlation in Germany and a positive correlation in the US. These correlation results are available upon request.

[13]This data includes 27 countries, excluding Germany and the US due to data constraints. A full list of these countries can be found in Appendix B.5. External learning in this data is measured by participation in seminars, workshops, or other courses led by outsiders in the last 12 months. Internal learning is measured by participation in courses led by coworkers or supervisors in the last 12 months, or by reporting learning skills from coworkers more than once a month in the current job. Years of experience are directly measured, capturing realized years of paid work up to the measurement date. Analogous plots using potential years of experience, showing similar patterns, are available upon request. Table B.3 provides summary statistics

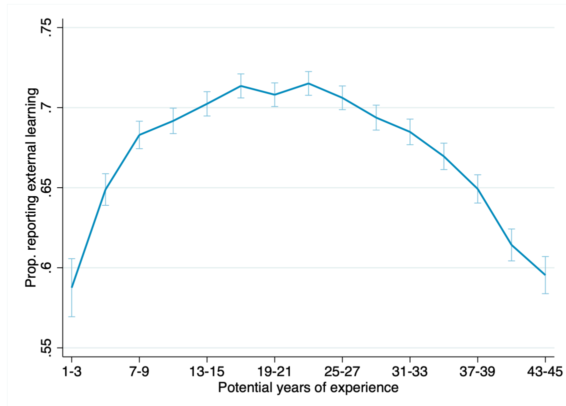
Figure 3.1: Prevalence of internal and external learning throughout workers' lifecycles

Germany

(a) Internal learning

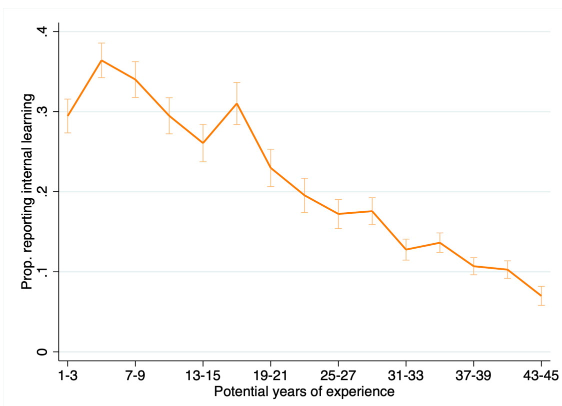


(b) External learning

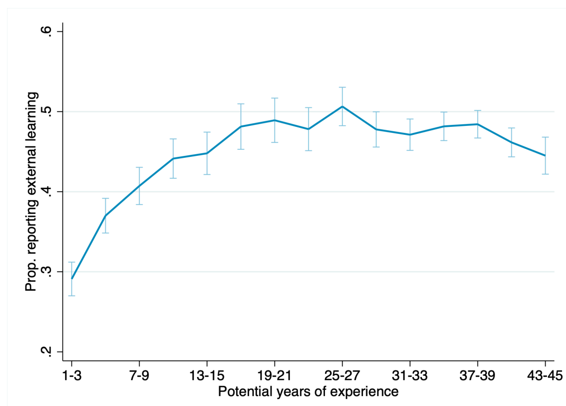


United States

(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.

previously and are robust to controlling for several demographic variables and firm characteristics, along with occupation and industry fixed effects (see Table D.3). However, the effects become noisier due to the small sample sizes in each country.

The lifecycle patterns documented in this section align with findings from various studies indicating that younger workers are more sensitive to peer learning than older workers. For example, Nix (2017) shows that coworker learning spillovers are more pronounced for

for this data.

younger workers, with no effects for workers over 40. Jarosch et al. (2021) find that the positive effects of peers' wages on future wages are substantially stronger among younger workers. Emanuel, Harrington, and Pallais (2023) document that sitting proximity increases how much junior software engineers learn from senior colleagues. Azoulay et al. (2010) show that academics, and especially younger ones, experience a decline in publication rates following the unexpected death of collaborators. Finally, Glaeser (1999) finds that young workers are more likely to be drawn to urban areas due to the increased opportunities for learning from skilled workers.

3.4 Review of literature on the link between internal and external learning, human capital, and wages

Our focus on internal and external learning to explain lifecycle wage growth is rooted in numerous studies showing significant and lasting effects of these learning types on wages and productivity. In Table E.1, we summarize the findings from 19 studies, primarily focusing on the US to align with our model's calibration in the following section.[14]

These studies document overwhelmingly positive and significant impacts of internal and external learning on wages and productivity. While identification strategies vary across studies, it is notable that those using instrumental variables or natural experiments report consistently robust positive effects (e.g., Bartel (1995) on external learning, Emanuel et al. (2023) on internal learning). Furthermore, the table highlights that these effects are present across diverse settings and time periods, underscoring the widespread importance of both internal and external learning for human capital and wages.

4 A model of human capital with two sources of learning

We now build a model featuring a novel two-source learning framework in order to shed light on our empirical findings and assess the importance of both internal and external learning in shaping lifecycle human capital acquisition and wage dynamics.

[14] In these studies, external learning generally refers to formal on-the-job training courses, while internal learning refers to informal on-the-job learning through interactions with superiors or coworkers. Internal learning is sometimes measured using proxies such as the educational level, wages, or other outcomes of coworkers, as well as other indicators of worker interaction. While we focus on evidence for the US in relation to external learning, we also draw from studies in other contexts for internal learning, as the literature on internal learning is less extensive.

4.1 Model setup

The model follows an overlapping generations structure, featuring two sectors: a training sector that employs workers (trainers) providing external learning services, and a final-good production sector. The production sector incorporates firm heterogeneity in productivity and labor market frictions, as in Cahuc, Postel-Vinay, and Robin (2006). Within this sector, workers and firms engage in Nash bargaining to determine compensation and jointly decide on learning investments to maximize the match value.

This framework captures key interactions between lifecycle human capital accumulation and labor market dynamics. First, labor market frictions create incentives for firms to invest in general skills. Without these frictions and in the absence of long-term contracts, human capital gains are immediately reflected in wages, eliminating firms' incentives to invest (Acemoglu, 1997; Acemoglu and Pischke, 1999; Moen and Rosén, 2004). Second, Nash bargaining allows firms and workers to share the returns from learning investments, consistent with theories where workers finance learning through lower wages (Becker, 1964; Ben-Porath, 1967; Mincer, 1974; Acemoglu and Pischke, 2003), as well as evidence on wage-setting through bargaining (Hall and Krueger, 2012; Caldwell and Harmon, 2019). Finally, the search setting emphasizes the importance of firm heterogeneity for career progression: matches with firms offering better or worse learning environments are enduring and have lasting effects on human capital and wages, aligning with recent work on the long-term impact of initial matches (Arellano-Bover, 2020) and the role of luck in wage dynamics (Huggett et al., 2011).

In what follows, we focus on the model's stationary equilibrium, where the age and human capital distributions of the workforce remain constant within firms. Workers are indexed by i, and firms by their productivity level z.

4.1.1 Workers

The model features overlapping generations of workers who enter the labor market at age $a = 1$ and retire at age J. Each cohort is a continuum of size 1, with new entrants replacing retirees. Workers can choose to work in either the production or training sector, and accumulate human capital throughout their lives, which determines their productivity. Each worker has one unit of time per period. In the production sector, this time is allocated to work, internal learning, and external learning, whereas in the training sector, workers spend all their time working. Workers derive linear utility from consuming a homogeneous final good and discount future utility at rate ρ.

4.1.2 Human capital accumulation

Human capital, h, evolves in a ladder-like fashion over a worker's life:

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

Workers begin with the lowest level of human capital, h_1, and move up the ladder through internal and external learning, as well as learning-by-doing.[15] All forms of learning contribute to general human capital, which is transferable if workers change jobs.[16]

We denote worker i's human capital as h_{m_i}, where m_i indexes the worker's position on the human capital ladder. For a worker i in firm z, the probability of moving up the ladder by one step is given by:

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

$s^E(h_{m_i}, \mathbb{H}(z))$ represents the increase in human capital per unit of learning time, where $\mathbb{H}(z) = \{h_{m_i}\}_{i \in \mathbb{I}(z)}$ is the human capital distribution of firm z's workforce (denoted by $\mathbb{I}(z)$), capturing its internal learning environment. l denotes the total time allocated to learning, with $0 < \gamma < 1$ capturing its diminishing marginal returns. $\epsilon > 0$ captures an exogenous probability of advancing up the human capital ladder, akin to learning-by-doing.[17]

The human capital increment, $s^E(h_{m_i}, \mathbb{H}(z))$, depends on the time allocated to internal and external learning, as well as the probability of matching with a coworker or trainer capable of effectively teaching the worker. Specifically, workers can only learn from coworkers or trainers with higher human capital than their own, consistent with the findings of Herkenhoff et al. (2024), who show that workers learn more from knowledgeable colleagues, and Jarosch et al. (2021), who show that higher-paid peers have a greater influence on workers' future wages. The functional form of the human capital increment is given by:

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

[15] In Appendix G.3, we allow initial human capital to differ across workers to account for differences in schooling levels, and find that the results remain similar to those of the baseline model.

[16] We focus on general human capital for tractability and consistency with existing evidence, which suggests that general human capital is considerably more important than firm-specific human capital for wage growth (Altonji and Shakotko, 1987; Lazear, 2009; Kambourov and Manovskii, 2009). Moreover, research on workplace learning (see Manuti, Pastore, Scardigno, Giancaspro, and Morciano, 2015, for a review) indicates that both internal and external learning are vital to workplace learning and develop transferable skills. Thus, rather than focusing on transferability differences, our theory emphasizes differences in the knowledge pool each learning source draws from, and generates a series of testable predictions that align with the observed data (see Appendix H).

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

The probability of matching with a coworker who has higher human capital when learning internally, denoted by $p(h_{m_i}, \mathbb{H}(z))$, depends on the share of such coworkers in the firm and is given by $p(h_{m_i}, \mathbb{H}(z)) = 1 - F(h_{m_i}; \mathbb{H}(z))$, where $F(\cdot; \mathbb{H}(z))$ represents the cumulative distribution of human capital in firm z. Similarly, the probability of matching with a trainer who has higher human capital when learning externally, denoted by $p_e(h_{m_i})$, is given by $p_e(h_{m_i}) = 1 - F_t(h_{m_i})$, where $F_t(\cdot)$ is the cumulative distribution of trainers' human capital, weighted by the number of training units each trainer produces. The shares of time devoted to internal and external learning are denoted by g and $(1-g)$, respectively, while $\sigma > 0$ captures the elasticity of substitution between the two types of learning. $A(z)$ and $A_e(z)$ represent the efficiencies of internal and external learning in firm z. Modeling these efficiencies as functions of firm productivity, z, captures the faster human capital accumulation observed in more productive firms, potentially due to scale effects or complementary physical capital.[18] This modeling approach helps explain empirical differences in learning time by firm size.

Workers forgo production time to engage in internal or external learning. However, external learning involves an additional cost of q per unit of time for purchasing training services from a higher-skilled trainer.[19] Workers incur no costs when coworkers learn from them and can teach multiple coworkers simultaneously. This reflects the nature of internal learning, which often occurs during everyday work interactions, such as during lunch or at the water cooler.[20]

Lastly, following Mincer (1989) and Blundell, Costa Dias, Goll, and Meghir (2021), we allow human capital to depreciate. A worker with human capital h_{m_i} faces a probability of $\frac{\delta_h(h_{m_i}-h_1)}{h_{m_i}-h_{m_i-1}}$ of moving down one step on the human capital ladder each period, where δ_h is the depreciation rate. This setting ensures that human capital cannot fall below h_1.

4.1.3 Training sector

The training sector is frictionless, allowing unemployed workers to freely transition and utilize the training technology without constraints.[21] Trainers meet workers engaged in external

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

[19]In Appendix G.5, we introduce a fixed learning cost to allow for changes in the extensive margin of learning, finding similar results to the baseline model.

[20]In Appendix G.2, we consider a scenario where internal trainers incur time costs for mentoring, and find that equilibrium properties and quantitative results remain similar to the baseline model.

[21]We allow only unemployed individuals to switch sectors, ensuring that the human capital levels of trainers and production workers overlap, consistent with NHES data. The training sector is modeled as frictionless since many trainers are self-employed or work as independent contractors (Hamlin, Ellinger, and Beattie,

learning randomly, and can effectively teach only those with lower human capital.[22] In addition, trainers' production scales with their human capital, enabling high-skill individuals to teach multiple students simultaneously. 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) among external learners. In equilibrium, the price of training services, q, ensures that the total demand for external training services matches the total supply provided by all trainers.

4.1.4 Production sector

The production sector features a frictional labor market with a unit measure of heterogeneous firms producing a homogeneous good. Firm productivity levels follow a Pareto distribution, $z \sim \Phi(z)$. Firms post vacancies v and match with workers through random search. Following Acemoglu and Hawkins (2014), vacancy posting costs are convex and given by $c_v v^{1+\gamma_v}/(1+\gamma_v)$, where $\gamma_v > 0$ ensures that more productive firms hire more workers while coexisting with less productive ones. The total number of vacancies is given by $V = \int v(z) \mathrm{d}\Phi(z)$, where $v(z)$ represents the optimal number of vacancies of firm z.

At the beginning of each period, labor matches end at an exogenous rate κ, sending separated workers to the unemployment pool alongside new entrants. Before job search begins, unemployed workers may choose to switch to the training sector and become trainers, while trainers may opt to switch to the production sector, joining the unemployment pool. The number of unemployed workers at the job search phase is denoted by U. Unemployed workers encounter job vacancies at a meeting rate λ_U, determined endogenously by $\lambda_U = \chi(V/U)$, where the function $\chi(\cdot)$ governs the matching process. Upon encountering a vacancy, an unemployed worker accepts it with certainty, under the assumption that the value of unemployment is equivalent to employment at the least productive firm.[23] Unemployment persists if no vacancy is encountered. We do not incorporate on-the-job search in our baseline model to avoid additional notation and complexity. The equilibrium properties and quantitative results remain very similar when on-the-job search is included (see Appendix G.1).

When a job match is established, worker i's output in firm z is the product of firm produc-

2008; Evans and Lines, 2014).

[22] This assumption amplifies the returns to human capital in the training sector, raising the average human capital level of trainers above that of production workers.

[23] As suggested by Bagger et al. (2014), this approach circumvents the complexities associated with varying reservation wages among workers with different human capital levels and ages. In Appendix G.4, we explicitly model unemployment benefits and losses, and find that the results are similar to the baseline.

tivity and worker human capital: $y = zh_{m_i}$. Thus, the production function is supermodular: higher productivity firms generate more output per unit of labor, and human capital and firm productivity act as complements, as illustrated by Acemoglu and Pischke (1998) and Bagger et al. (2014). Worker-firm pairs jointly optimize learning investments to maximize the match value and use Nash bargaining to determine the allocation of match surplus and the resulting wage rate. Below, we characterize these results, along with optimal vacancy posting.

4.1.4.1 Learning investments

Firms and workers jointly choose total learning time and the shares of this allocated to internal and external learning to maximize the match value. For a worker i of age a_i and human capital h_{m_i} in firm z, the match value is given by:

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

The first line represents the current output value minus the learning costs stemming from forgone production and payments to external trainers. The first term of the second line captures the spillovers generated by worker i for coworkers in firm z by impacting the firm's internal learning environment. This is quantified by comparing each coworker j's match value when worker i is present in firm z's workforce in the current period to the value if worker i were absent, holding all else constant. Similarly, and to prevent double-counting, the second term in the second line deducts the value of spillovers worker i perceives from other coworkers in the firm.[24] The final two lines capture the discounted future values based on whether worker i climbs the human capital ladder or not, which depend on the future match value, the value of unemployment, $V_U^a(h_m)$ (presented in Equation (A.7)), and the value of working as a trainer, $V_{TR}^a(h_m)$ (presented in Equation (A.8)). The expectation therein accounts for the uncertainty of human capital depreciation, which could cause the

[24] This term conceptually captures the idea that workers may take a wage cut to fund their human capital acquisition and compensate mentors in the firm (Jarosch et al., 2021).

worker to descend the ladder. $\mathbb{H}'(z) = \Gamma(\mathbb{H}(z))$ represents the law of motion for the human capital distribution in firm z.

In Appendix A.1 we provide details on the derivation of this match value and show that it can be simplified to

$$
\begin{aligned}
M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) \approx \max_{l,g} \Bigg[& \underbrace{z h_{m_i}(1-l)}_{\text{output value}} - \underbrace{(1-g)l \times q p_e(h_{m_i})}_{\text{pay for external trainers}} \Bigg] \\
& + \underbrace{\int_{j \in \mathbb{I}(z), j \neq i} \frac{\partial s}{\partial p} \left[p(h_{m_j}, \mathbb{H}(z)) - p(h_{m_j}, \mathbb{H}(z) \backslash \{i\}) \right] \times O^{a_j+1}(h_{m_j}, z, \mathbb{H}'(z)) dj}_{\text{spillovers generated by worker } i \text{ accruing to other workers in firm}} \\
& + \underbrace{s(h_{m_i}, \mathbb{H}(z)) \times \rho \mathbb{E} \left[(1-\kappa) M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\} \right]}_{\text{future value if successfully climb human capital ladder}} \\
& + \underbrace{[1 - s(h_{m_i}, \mathbb{H}(z))] \times \rho \mathbb{E} \left[(1-\kappa) M^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i}), V_{TR}^{a_i+1}(h_{m_i})\} \right]}_{\text{future value if don't successfully climb human capital ladder}} .
\end{aligned} \tag{4}
$$

This expression has two noteworthy changes compared to Equation (3).

First, the term capturing worker i's spillovers to its coworkers now explicitly captures worker i's impact on the firm's learning environment, as $\frac{\partial s}{\partial p} \left[p(h_{m_j}, \mathbb{H}(z)) - p(h_{m_j}, \mathbb{H}(z) \backslash i) \right]$ reflects the change in coworker j's learning probability due to worker i's presence. Relatedly, $O^{a_j+1}(h_{m_j}, z, \mathbb{H}'(z))$ captures the gain in worker j's match value from a one-step increase in human capital:[25]

$$
O^{a_j+1}(h_{m_j}, z, \mathbb{H}'(z)) = \rho \mathbb{E} \left[(1-\kappa) \Delta M^{a_j+1}(h_{m_j+1}, z, \mathbb{H}'(z)) + \kappa \Delta \max\{V_U^{a_j+1}(h_{m_j+1}), V_{TR}^{a_j+1}(h_{m_j+1})\} \right].
$$

Since workers make learning investments in every period, these spillovers represent a flow benefit.[26] The value of these spillovers is positive for high human capital workers who enhance the firm's learning environment, while negative for those with low human capital. This reflects a reapportioning of production value towards higher-skilled workers, and is illustrated in the context of an analytical example in Appendix A.1.4.

Second, the term representing the spillovers worker i perceives from other workers in the firm drops out, as it is approximately zero for each worker. This occurs because the positive

[25] $\Delta M^{a_j+1}(h_{m_j+1}, z, \mathbb{H}'(z)) = M^{a_j+1}(h_{m_j+1}, z, \mathbb{H}'(z)) - M^{a_j+1}(h_{m_j}, z, \mathbb{H}'(z))$ represents the increase in worker j's match value from a one-step increase in human capital assuming she remains at firm z in the next period. Similarly, $\Delta \max\{V_U^{a_j+1}(h_{m_j+1}), V_{TR}^{a_j+1}(h_{m_j+1})\} = \max\{V_U^{a_j+1}(h_{m_j+1}), V_{TR}^{a_j+1}(h_{m_j+1})\} - \max\{V_U^{a_j+1}(h_{m_j}), V_{TR}^{a_j+1}(h_{m_j})\}$ captures the analogous gain assuming the worker becomes unemployed in the next period. As before, the expectation reflects uncertainty of human capital depreciation.

[26] In particular, when considering spillover effects for the current period, the human capital distribution $\mathbb{H}(z) \backslash i$ represents firm z without worker i in the current period, while future human capital distributions are held constant. This ensures that spillover effects are treated as a flow.

impact of mentors and the negative impact of all other workers cancel each other out. We formally show these results, discuss their intuition, and provide an analytical example to illustrate the mechanisms at play in Appendix A.1.

We determine the optimal total learning time and the allocation between internal and external learning by taking the first-order conditions of Equation (4). The results are summarized in the following proposition.

Proposition 1 (Optimal learning investments)**.** *For worker i at firm z, the optimal learning investments (if they are internal solutions) follow:*
(i) The optimal total learning time l satisfies:

$$l = \left(\frac{\gamma s^E(h_{m_i}, \mathbb{H}(z)) O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z))}{z h_{m_i} + (1-g) q p_e(h_{m_i})} \right)^{1/(1-\gamma)}. \tag{5}$$

(ii) The optimal share of learning time allocated to internal learning g satisfies:

$$\begin{aligned} l q p_e(h_{m_i}) = &\left[s^E(h_{m_i}, \mathbb{H}(z))^{\frac{1}{\sigma}} \left(\frac{(A_e(z) p_e(h_{m_i}))^{\frac{\sigma-1}{\sigma}}}{(1-g)^{\frac{1}{\sigma}}} - \frac{(A(z) p(h_{m_i}, \mathbb{H}(z)))^{\frac{\sigma-1}{\sigma}}}{g^{\frac{1}{\sigma}}} \right) \right] \\ &\times O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) l^{\gamma}. \end{aligned} \tag{6}$$

Proof: See Appendix A.3.

Result (i) defines the optimal total learning time, which balances the benefits and costs of human capital accumulation. The numerator in Equation (5) captures the expected benefits from learning, which include gains in match value (if the worker stays in the firm) and unemployment value (if separated). The denominator reflects the marginal cost of learning due to forgone production and trainer payments. With diminishing returns to learning ($\gamma < 1$), this equation pins down the optimal learning time.

Result (ii) characterizes the optimal share of time allocated to internal learning. The left-hand side of Equation (6) reflects the benefit (in terms of cost reductions) of internal learning, which is cheaper than external learning. The right-hand side represents the loss associated with internal learning due to lower efficiency if the external learning success rate ($p_e(h_{m_i})$) is high. A higher internal learning success rate ($p(h_{m_i}, \mathbb{H}(z))$) or a higher cost of external learning services (q) leads to a greater share of learning time allocated to internal learning.

This latter result is key to understanding the lifecycle switch between learning modes. Young workers, with low human capital, enter the production sector and match with firms where many coworkers have higher human capital. Given the high probability of successful internal learning and its relatively low cost, young workers focus on internal learning. As workers age and accumulate human capital, the pool of higher-skilled coworkers shrinks, while the relative cost of forgone production—and thus internal learning—increases. This prompts a shift to external learning, since trainers tend to have higher average human capital than production workers.

In addition, it is worth noting that while firms and workers jointly determine learning levels to maximize the match value, these investments are inefficiently low compared to the social optimum because they ignore benefits to future employers and coworkers after separation. In Section 6.2, we evaluate the impact of subsidies matching those offered by several US states and covering a portion of external learning costs to address this inefficiency.

4.1.4.2 Wages

Following Cahuc et al. (2006) and Bagger et al. (2014), the wage for a worker i of age a_i and human capital h_{m_i} working in firm z, is proportional to the current-period revenue she generates, which includes net output and spillovers to coworkers:

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

$X^{a_i}(h_{m_i}, z, \mathbb{H}(z))$ captures the spillovers generated by worker i for coworkers in firm z, as shown in the second line of Equation (4).[27] r represents the piece rate capturing the worker's share of revenue, which is determined through Nash bargaining at the time of hiring and held constant for the duration of the match.

Specifically, the piece rate r is determined by maximizing the Nash product when the job match is formed:

$$\max_r \; [V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z)) - V_U^{a_i}(h_{m_i})]^{\beta} \, J^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))^{1-\beta}. \tag{7}$$

The worker's bargaining power is denoted by β, and the firm's by $1-\beta$. $V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))$ is the worker's value from the match, $V_U^{a_i}(h_{m_i})$ is her reservation value (unemployment), and $J^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))$ is the firm's value from the match, with zero reservation value for the firm. We present the formulas for the worker's and firm's values in Appendix A.2.[28]

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

[28] The worker and firm values sum to the match value, $M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) = J^{a_i}(r, h_{m_i}, z, \mathbb{H}(z)) + V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))$.

The first-order condition for the Nash bargaining solution implies that r satisfies:

$$\frac{V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z)) - V_U^{a_i}(h_{m_i})}{\beta} = \frac{J^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))}{1-\beta} = M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) - V_U^{a_i}(h_{m_i}). \quad (8)$$

We use this equation to solve for the piece rate, $r^{a_i}(h_{m_i}, z, \mathbb{H}(z))$, which depends on the worker's age and human capital level upon hiring and varies across firms.[29]

4.1.4.3 Vacancy posting

Each firm z determines the optimal number of vacancies $v(z)$ to post to maximize its share of the match surplus generated from hiring workers:

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

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

$$c_v v(z)^{\gamma_v} = \sum_{m=1}^{M} \sum_{a=1}^{J} \frac{\lambda_U (1-\beta)}{V} \left[M^a(h_m, z, \mathbb{H}(z)) - V_U^a(h_m) \right] D_U^a(h_m). \quad (10)$$

The left-hand side of this equation represents the marginal cost of posting a vacancy, which increases with the number of vacancies. The right-hand side represents the expected benefit from posting a vacancy, derived from hiring unemployed workers, with $(1-\beta)$ determining the firm's share of the surplus increment, as described in Equation (8).

4.1.5 Equilibrium and solution

In Appendix A.4, we define the general equilibrium of the model. In Appendix A.5, we describe the numerical algorithm used to solve it.

4.2 Calibration

We calibrate the model to the US to assess the role of internal and external learning in workers' human capital accumulation and wage growth. We adopt a Cobb-Douglas job matching function between job seekers and vacancies (Shimer, 2005), which implies a meeting

[29] Cahuc et al. (2006) analytically solve for the piece rate and show that it increases with the worker's outside option. In our model with endogenous human capital, we use numerical methods to solve for the value of employment at any given piece rate r, and then use this to find the piece rate that delivers the worker's bargained value.

rate of $\chi(x) = c_m x^{\phi}$, where $0 < \phi < 1$ represents the elasticity of job matches with respect to vacancies. The efficiencies of internal and external learning are parameterized 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 effectiveness across firms of varying productivity levels.[30] Firm productivity is assumed to follow a Pareto distribution, $\Phi(z) = 1 - z^{-\zeta}$, where ζ is the shape parameter.

4.2.1 Externally calibrated parameters

We draw several parameters directly from the literature, as listed in panel A of Table 4.1. A model period represents one quarter. We set the discount rate to $\rho = 0.99$ and assume a working life of $J = 160$ periods, equivalent to 40 years. Without loss of generality, the step size of the human capital ladder is set to $\gamma_h = 0.05$, meaning each step increases human capital by 5%. The lower bound of human capital, h_1, is normalized to 1. The elasticity of matches to vacancies is set at $\phi = 0.3$, following Shimer (2005). We set the curvature of vacancy costs at $\gamma_v = 1$, as in Acemoglu and Hawkins (2014), implying quadratic vacancy costs.[31]

Due to a lack of US-specific evidence, we rely on estimates from other countries for workers' bargaining power (β). Estimates of β vary widely. For instance, Bagger et al. (2014) find β between 0.29 and 0.32 for Denmark, while Gregory (2019) estimates β at 0.66 for Germany. We set $\beta = 0.5$ to reflect equal bargaining power, consistent with these estimates.[32]

4.2.2 Internally calibrated parameters

We estimate 11 parameters: the efficiencies of internal and external learning ($\bar{A}, \bar{A}_e, \alpha$), elasticity of substitution between learning modes (σ), learning-by-doing probability (ϵ), diminishing returns to learning time (γ), human capital depreciation rate (δ_h), constants in the matching and vacancy cost functions (c_m, c_v), firm productivity shape parameter (ζ), and separation rate (κ).

These parameters are estimated jointly using the method of moments, minimizing the squared differences between model and data moments. We target 11 data moments, the first

[30] As discussed earlier, allowing the efficiencies of internal and external learning to increase with firm productivity helps explain empirical differences in learning time by firm size, and captures other aspects of learning environments. To test robustness, we considered a linear functional form, $A(z) = \bar{A}(1 + \alpha z)$ and $A_e(z) = \bar{A}_e(1 + \alpha z)$, and found similar results. These results are available upon request.

[31] There is little consensus on the convexity of vacancy costs in the literature. We tested the robustness of our results by considering $\gamma_v = 2$ and found similar results. These results are available upon request.

[32] Appendix G.6 shows that our findings are robust to different values of β.

7 of which are: the unemployment rate (1994-2007) and labor market tightness (2000-2007) from FRED; the tail parameter of the firm size distribution estimated by Axtell (2001); the share of workers remaining employed in the next quarter from Donovan, Lu, and Schoellman (2023); the ratio of all workers' average external learning time to that of new workers (1 year of experience) from NHES data;[33] and the average quarterly returns to experience attributed to human capital for workers with 0–40 and 0–25 years of experience from Lagakos, Moll, Porzio, Qian, and Schoellman (2018).[34]

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)
γ_v	Curvature of vacancy cost	1	Acemoglu and Hawkins (2014)
β	Workers' bargaining power	0.5	Estimates in literature
	Panel B: Internally calibrated parameters		
$\bar{A}$	Constant in efficiency of internal learning	0.50	Joint estimation
$\bar{A}_e$	Constant in efficiency of external learning	0.23	Joint estimation
α	Elasticity of learning efficiency to firm productivity	0.73	Joint estimation
σ	Elasticity of substitution between internal/external learning	3.63	Joint estimation
ϵ	Exogenous human capital gain	0.02	Joint estimation
γ	Degree of diminishing returns in learning	0.41	Joint estimation
δ_h	Depreciation rate of human capital	0.03	Joint estimation
c_m	Constant in matching function	0.46	Joint estimation
c_v	Constant in vacancy cost function	0.32	Joint estimation
ζ	Shape parameter of firm productivity distribution	4.32	Joint estimation
κ	Exogenous separation rate	0.03	Joint estimation

Since the NHES data lacks employer and learning time information, we compute the remaining 4 moments using the 1995 US-SEPT dataset (see Appendix B.2 for details of this data).

[33] Since NHES only captures learning at the extensive margin, we gauge the share of time spent on external learning based on participation rates. Given the importance of this moment for estimating the elasticity of substitution, we consider a higher elasticity value in Appendix G.6 and show our findings are robust.

[34] We focus on the returns attributed to human capital, as wage growth in our baseline model is driven by human capital acquisition. Consistent with the literature (Altonji, Smith, and Vidangos, 2013; Bagger et al., 2014), which attributes roughly 50% of wage growth to human capital, we target 50% of the observed wage growth. Appendix G.1 includes a model extension with job ladders and targets the full observed wage growth, obtaining similar quantitative results.

These moments include the shares of time spent on internal and external learning, the ratio of average learning time in firms with 100+ workers to that in firms with 50–99 workers, and the ratio of older workers' average total learning time to that of all workers.[35]

Our choice of targeted moments is guided by the relationship between the parameters and these moments. For example, the vacancy cost constant (c_v) directly impacts labor market tightness, while the matching function constant (c_m) determines the unemployment rate given this tightness. The efficiencies of internal and external learning ($\bar{A}$, $\bar{A}_e$) are linked to the shares of time spent on each type of learning. A higher elasticity of learning efficiency to firm productivity (α) indicates a larger disparity in learning environments across firms, while a higher elasticity of substitution between learning modes (σ) suggests a more pronounced shift from internal to external learning as workers age. Given the efficiencies of learning, the relative learning time of young versus old workers, along with lifetime and youth wage growth, provides insights into the extent of diminishing returns to learning (γ), learning-by-doing probability (ϵ), and human capital depreciation (δ_h). In Table F.1, we present the Jacobian matrix, which shows how changes in each parameter affect the corresponding moments, further highlighting these connections.

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)	0.55	0.55
3. Shape parameter of firm employment distribution[36]	1.24	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	0.0023	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 internal learning	0.033	0.033
8. Share of total time spent on external learning	0.014	0.014
9. Ratio of total learning time in 100+ worker firms to 50–99 worker firms	1.20	1.13
10. Ratio of all to new workers' average time spent on external learning	1.55	1.51
11. Ratio of old to all workers' average learning time	0.52	0.51

Panel B of Table 4.1 presents the values of the internally calibrated parameters, which are

[35] Older workers are defined as those aged 55+ in the data and as those in the last 10 years of their working life in the model.

[36] Although our estimated shape parameter for the productivity distribution is $\zeta = 4.32$, the model predicts a thick tail for the employment distribution with a shape parameter of 1.24. This arises from the presence of an unemployment value, per which firm value increases more than proportionally with productivity. In addition, faster human capital accumulation further increases firm value in more productive firms.

reasonable and align well with existing literature. For instance, the elasticity of learning efficiency to firm productivity ($\alpha = 0.73$) suggests greater learning efficiency in more productive firms, consistent with Engbom (2021). The elasticity of substitution between learning modes ($\sigma = 3.63$) indicates moderate substitutability. The degree of diminishing returns to learning ($\gamma = 0.41$) is similar to the estimate by Manuelli and Seshadri (2014). Each worker has a 2% chance per period of advancing on the human capital ladder via learning-by-doing. The calibrated quarterly human capital depreciation rate ($\delta_h = 0.03$) matches estimates from Blundell et al. (2021) based on British labor surveys, while the quarterly job destruction rate ($\kappa = 0.03$) aligns with the employment-unemployment transition rates reported by Shimer (2012). With these parameters, our model fits the targeted data moments well, as shown in Table 4.2. In the next section, we further validate our model and calibration by comparing its predictions regarding workers' learning patterns with the data.

4.3 Model predictions and validation: workers' learning patterns

To further validate our framework and calibration, we explore key properties of the model's stationary equilibrium and compare them to the data. We focus on two learning patterns consistent with the evidence presented earlier: (1) the *lifecycle pattern of worker learning*, per which young workers prioritize internal learning and shift towards external learning as they age, and (2) the *firm productivity pattern of worker learning*, per which workers in more productive firms spend more time on both internal and external learning.

4.3.1 The lifecycle pattern of worker learning

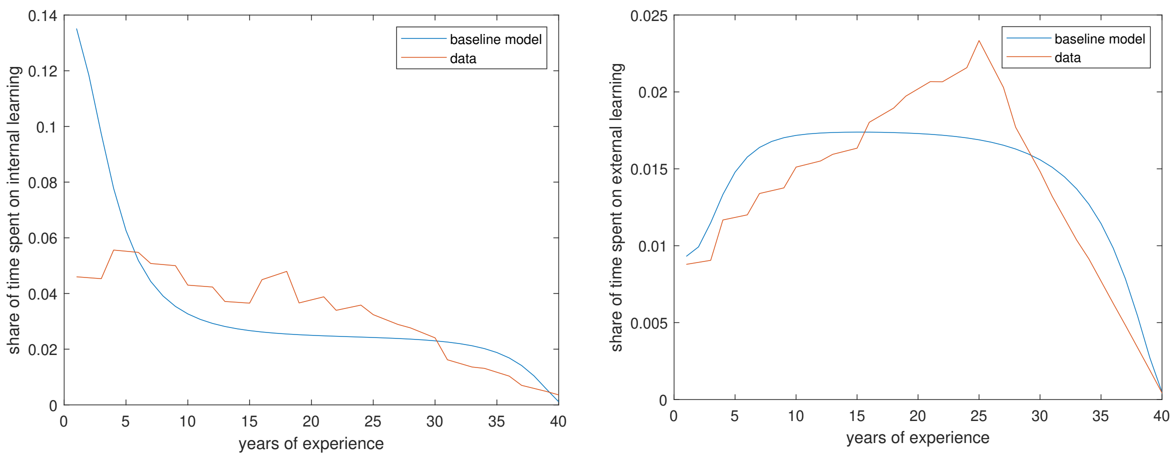
The *lifecycle pattern of worker learning* arises in the model from shifts in the incentives for internal and external learning throughout a worker's career. Early in their careers, workers have a long career ahead and encounter many coworkers with higher human capital, encouraging them to invest in internal learning. As workers climb the human capital ladder, however, the pool of higher-skilled coworkers shrinks, prompting a transition to external learning, since trainers have higher average human capital levels.[37] Eventually, external learning also declines as workers age further, due to the proximity to retirement and continued human capital accumulation which raises the opportunity costs of learning while reducing its success rate.

[37]Figure F.1a shows the human capital distribution for production workers and trainers, indicating that trainers have higher human capital on average. This arises because the returns to human capital are larger in the training sector.

Figure 4.1 illustrates this result by showing that, as workers gain experience, the model predicts a decline in internal learning, and an initial increase, followed by a decline, in external learning. This aligns well quantitatively with the data on learning profiles across the lifecycle, except for very young workers, where the model overestimates internal learning time. This suggests that other institutional factors, not accounted for in our model, may influence the learning patterns of novice workers.

Figure 4.1: 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 and the data. To construct the data moments, we multiply the participation rates from panels (c) and (d) in Figure 3.1 by age-specific hours per participant from the US-SEPT data, distinguishing between internal and external learning. Since the US-SEPT data provides hours per participant for only four age groups, we interpolate the data using a linear trend to estimate age-specific hours. These values are then normalized to ensure that the overall average aligns with our targeted moments.

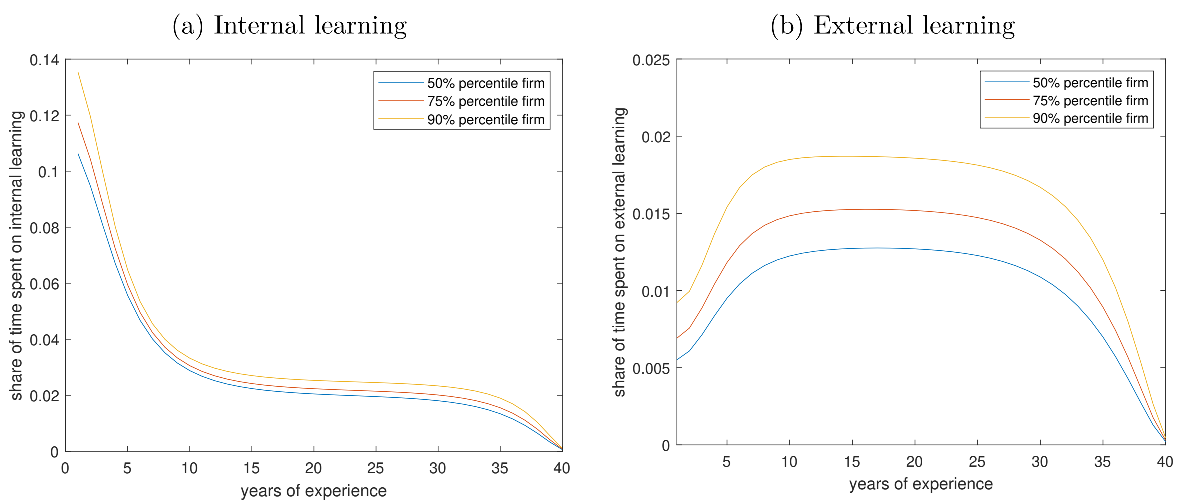
In Appendix H.1, we provide evidence supporting three testable predictions underlying these lifecycle learning patterns: (1) trainers possess higher average human capital levels than production workers, making them better suited for teaching experienced workers; (2) the proportion of workers engaging in learning-by-doing increases with human capital, aligning with the rise in work hours and the decline in external and internal learning as workers age; and (3) internal learners tend to handle less complex tasks compared to external learners, consistent with their lower human capital and the greater ease of receiving training from coworkers.[38]

[38] To further validate our modeling of external learning, we compare the external learning costs suggested by our model to those in the data. According to the US-SEPT, external training payments (including tuition reimbursements, direct payments, and subsidies) amount to 13% of trainees' wage costs due to lost production time. In our model, these payments account for 15% of trainees' wage costs, showing notable

4.3.2 The firm productivity pattern of worker learning

The *firm productivity pattern of worker learning* reflects the higher returns to skill acquisition in more productive firms. Figure 4.2 illustrates this by showing the average time shares allocated to internal and external learning across the lifecycle for firms in the 50th, 75th, and 90th percentiles of productivity. More productive firms foster better learning environments by offering a greater variety of learning opportunities.

Figure 4.2: 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.

Although we cannot directly compare these results with data since the US-SEPT lacks information on learning patterns by both age and firm size, this finding aligns with empirical evidence from Section 3.2, as well as Figure D.9, which shows that internal and external learning increase with firm size throughout the lifecycle for German workers. Additionally, these results are consistent with the findings of Engbom (2017), Arellano-Bover (2020), and Arellano-Bover and Saltiel (2023), who show that workers in more productive firms acquire skills more quickly, and with Gregory (2019), who highlights the importance of having diverse training options for firms' learning environments.

The higher returns to skill acquisition prevalent in more productive firms stem from the supermodularity of the production function, enhanced learning efficiency, and an improved pool of coworkers due to positive assortative matching. This latter point is illustrated in Figure F.1b, which shows that more productive firms have a higher concentration of high-

similarity.

skill workers, consistent with evidence of positive sorting between employers and employees in the US (Barth, Bryson, Davis, and Freeman, 2016; Abowd, McKinney, and Zhao, 2018; Song, Price, Guvenen, Bloom, and Wachter, 2019). In our model, this sorting emerges from the greater learning investments and more favorable learning environments found in more productive firms, allowing workers to ascend the human capital ladder more rapidly.

5 Role of internal and external learning in human capital and wages

To evaluate the impact of internal and external learning on lifecycle human capital and wage dynamics, we conduct counterfactual analyses by disabling each learning source in the model. To do this, we set the efficiency terms for internal and external learning, $\bar{A}$ and $\bar{A}_e$, to zero in the baseline calibration. As a result, the time allocated to internal or external learning no longer influences the probability of climbing the human capital ladder.

5.1 Role of internal and external learning in human capital

Table 5.1 summarizes the average time spent on each learning source and the average level of human capital in the stationary equilibrium of the baseline model, as well as in the stationary equilibria of models without each learning source. We find that internal learning is 1.57 times more important than external learning for lifecycle human capital: without external learning, average human capital decreases by 7%, whereas without internal learning, it decreases by 11%. Without both learning sources, average human capital falls by 18%, though some

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

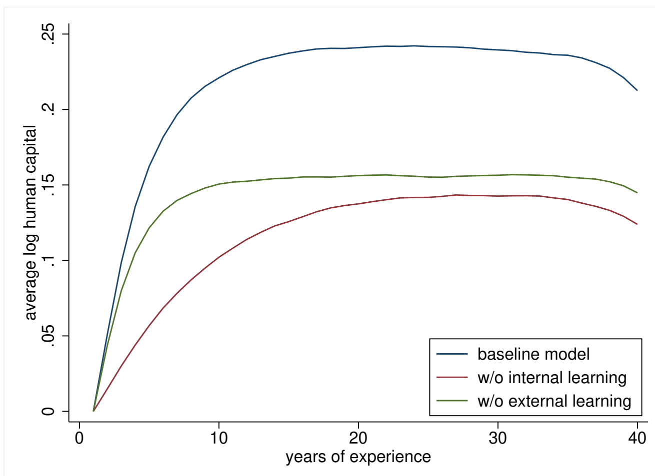
	Workers' share of time spent on learning		
	External learning	Internal learning	Avg human capital
Calibrated economy	1.44%	3.31%	1.27
W/o internal learning	1.58%	0	1.13
W/o external learning	0	3.60%	1.18
W/o both sources of learning	0	0	1.04
W/o any form of learning	0	0	1

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

human capital accumulation still occurs through learning-by-doing. In the absence of all learning sources, average human capital remains fixed at $h_1 = 1$, which is 21% lower than in the calibrated economy.

In Figure 5.1, we further explore the impact of internal and external learning on human capital formation by plotting the lifecycle trajectory of human capital in the baseline model and the two counterfactual scenarios. We find that while human capital growth is significantly reduced in the absence of either learning source compared to the baseline model, internal learning plays a more critical role during youth, whereas external learning becomes more important as workers age.

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.

5.2 Role of internal and external learning in wages

We now examine how wage growth and dispersion evolve over the lifecycle in the baseline and counterfactual scenarios. To understand the factors influencing wage dynamics, we break down wages into several components:

$$\log w^{a_i}(h_{m_i}, z, \mathbb{H}(z)) = \underbrace{\log r}_{\text{piece rate}} + \underbrace{\log z}_{\text{firm productivity}} + \underbrace{\log h_{m_i}}_{\text{worker's own productivity}} + \underbrace{\log\left(\frac{zh_{m_i} + X^{a_i}(h_{m_i}, z, \mathbb{H}(z))}{zh_{m_i}}\right)}_{\text{spillovers generated by worker } i \text{ to other workers in firm}}$$
$$+ \underbrace{\log\left(\frac{zh_{m_i}(1-l) - (1-g)lqp_e(h_{m_i}) + X^{a_i}(h_{m_i}, z, \mathbb{H}(z))}{zh_{m_i} + X^{a_i}(h_{m_i}, z, \mathbb{H}(z))}\right)}_{\text{learning costs}}. \quad (11)$$

Figure 5.2 plots lifetime wage growth in the baseline and counterfactual scenarios, along with the different components driving it.[39] In the baseline model, lifetime wage growth (blue line) is driven by three human capital components:

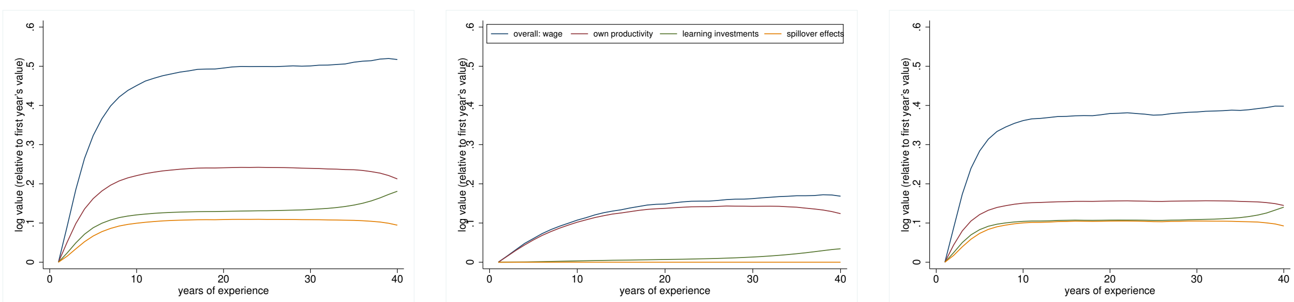
1. Worker's own productivity (red line): As workers accumulate human capital, they become more productive, leading to higher wages. This is the standard human capital gain discussed in the literature.

2. Spillovers to coworkers (yellow line): As workers climb the human capital ladder, they are more likely to serve as "teachers" within the firm, generating spillovers and receiving compensation.

3. Learning costs (green line): Learning investments and its associated costs decline with human capital and age, further boosting wages.

Figure 5.2: Wage growth and its components in 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 components (own productivity, spillovers to coworkers, and learning investments and its associated costs) by plotting the average log value of each of these (normalized by 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.

We find that own productivity accounts for 52% of total wage gains from human capital over 25 years of experience in the baseline model, while spillover effects and reduced learning costs together account for 48%, with learning costs contributing 26 percentage points and spillovers contributing 22 percentage points. This indicates that including the non-own productivity dimension of human capital—reflecting the broader benefits of human capital acquisition beyond the worker's own output—is crucial for understanding wage dynamics.[40]

[39] Since the model abstracts from job ladders, the first two components of Equation (11) – the piece rate and firm productivity— remain largely constant throughout workers' lifetimes and are therefore omitted for clarity.

[40] In Appendix H.2, we provide empirical evidence supporting the importance of the non-own productivity

Both internal and external learning are important to this result. In the absence of internal learning, spillover effects disappear, and the evolution of learning costs is slower due to less time spent on costly external learning. As a result, wage growth is mainly driven by increases in the worker's own productivity. Nevertheless, in the absence of external learning, lifetime learning costs and spillover effects are also reduced, as the high-cost external learning is eliminated, and the reduced human capital accumulation limits coworkers' effectiveness as "teachers." Thus, considering both internal and external learning is key for understanding lifecycle wage dynamics: both learning types influence learning costs; and internal learning allows knowledgeable workers to be compensated for fostering valuable spillovers among their colleagues, while external learning equips workers with new knowledge, enhancing their ability to generate these beneficial spillovers.

Additionally, Figure F.2 illustrates how wage dispersion evolves over the lifecycle in both the baseline and counterfactual scenarios, highlighting that internal and external learning are crucial for explaining the lifecycle rise in wage dispersion and its underlying drivers. Without internal learning, wage dispersion grows slowly as advancing up the human capital ladder becomes very costly. Without external learning, wage dispersion increases more rapidly, but differences in own productivity are understated as workers catch up to their peers more quickly. This further underscores the importance of both internal and external learning in shaping lifecycle wage dynamics.

6 Model applications and importance of two-source learning framework

We now use the model to examine the effects of remote work and training subsidies, highlighting the importance of considering both internal and external learning. To do this, we compare the effects of remote work and external learning subsidies on human capital and wages in the baseline model and in models without internal or external learning. For comparability, these alternate models are calibrated to match the targeted moments in Table 4.2.[41]

dimension of human capital, and specifically coworker spillovers, across various occupations in the US. Using data from the Department of Labor's O*NET project, we find that coworker tutoring is prevalent in many non-teaching occupations, especially those performed by highly skilled workers.

[41] To ensure comparability with the baseline model, we recalibrate the models without external or internal learning to match the total learning time in Table 4.2. The elasticity of substitution between learning modes (σ) is not included in these calibrations since it only applies when both learning sources are present. Consequently, we also exclude a related targeted moment—the ratio of all to new workers' time spent on external learning—which heavily depends on σ, as shown in Table F.1, and cannot be computed without

Table 6.1 summarizes the results of these exercises. Panel A shows the effects of remote work and training subsidies on human capital and wage growth in the baseline model, while panels B and C present the corresponding results for the alternate models. We now discuss these findings in detail.

Table 6.1: Impact of shocks in baseline and alternate model setups

	Benchmark (no shocks)	**Impact of shocks (rel. to baseline)**		
		Shift to remote work		50% subsidy for
		20% remote	100% remote	external learning
Panel A: Baseline model				
Average human capital	1.27	-0.45%	-3.10%	6.34%
On-the-job human capital gains	0.27	-2.11%	-14.70%	30.02%
Δ log wage, 0–5 yrs of exp	0.323	-5.05%	-25.30%	32.93%
Δ log wage, 5–25 yrs of exp	0.176	-1.08%	-10.10%	60.95%
Panel B: Model without internal learning				
Average human capital	1.32	0%	0%	7.80%
On-the-job human capital gains	0.32	0%	0%	32.24%
Δ log wage, 0–5 yrs of exp	0.170	0%	0%	68.35%
Δ log wage, 5–25 yrs of exp	0.217	0%	0%	9.96%
Panel C: Model without external learning				
Average human capital	1.23	-0.68%	-3.66%	0%
On-the-job human capital gains	0.23	-3.61%	-19.47%	0%
Δ log wage, 0–5 yrs of exp	0.355	-5.03%	-27.34%	0%
Δ log wage, 5–25 yrs of exp	0.166	-2.47%	-12.87%	0%

Notes: This table presents the following in the benchmark economy (with no shocks), the scenarios with 20% and 100% remote work, and the scenario with 50% subsidy to external learning: the average human capital level, on-the-job human capital gains (built by subtracting the initial human capital from the average human capital in the baseline scenario), and the average wage growth at 0–5 years of experience, and 5–25 years of experience. Panels A–C refer to the baseline model, the model without internal learning, and the model without external learning, respectively. The results in the alternate models report the percent changes relative to the corresponding baseline results.

6.1 Impact of remote work

To examine the effects of remote work, we draw on the findings of Barrero et al. (2021), who document a rise in remote work from 5% to 20% of workdays before and after the pandemic, and the findings of Emanuel et al. (2023), who report a 15% decrease in feedback from nearby colleagues after the shift to remote work. Based on this evidence, we infer a 2.25% decrease in the success rate of internal learning in the model (15% reduction in coworker contact

external learning.

$\times$ 15% reduction in feedback). We also consider an extreme scenario with 100% remote work, leading to a 14.25% decrease in the success rate of internal learning (95% reduction in coworker contact $\times$ 15% reduction in feedback).

The second and third columns of panel A in Table 6.1 show the impact of remote work in the baseline model. The post-pandemic rise in remote work results in a 0.45% decline in average human capital, equivalent to 2.11% of the human capital gained through on-the-job learning. In the scenario where all work is remote, average human capital decreases by 3.1%, representing 14.7% of on-the-job human capital gains. This suggests that remote work significantly hinders on-the-job human capital formation, with potential long-lasting effects, as younger workers who learn less will be less equipped to mentor future generations.

The increase in remote work also reduces lifetime wage growth. For workers with 0–5 years of experience, average wage growth declines by 5.05% due to the post-pandemic increase in remote work, and by 25.30% if all work is conducted remotely. At these early career stages, this decline is largely due to reduced productivity resulting from fewer opportunities for senior mentoring. This aligns with the findings of Emanuel et al. (2023), who show that reduced coworker feedback disproportionately affects young workers who rely heavily on guidance from more-experienced colleagues. For workers with 5–25 years of experience, average wage growth declines by 1.08% with the post-pandemic rise in remote work and by 10.10% if all work is remote. At these later career stages, the reduction in spillover compensation due to fewer junior colleagues to mentor accounts for much of this decline.[42] This suggests that remote work affects not only the wages of younger workers through reduced learning, but also senior workers' wages through reduced compensation from spillovers.

The second and third columns of panels B and C present the impact of remote work in the models without internal and external learning, respectively. In the model without internal learning, remote work has no effect on human capital or wage growth. In contrast, in the model without external learning, both human capital and wage growth losses are significantly overstated compared to the baseline model because there is no option to compensate for the loss of coworker mentoring by increasing external learning. This suggests that focusing on just one form of learning can skew the impact of shocks that affect internal and external learning asymmetrically, as it overlooks the potential for workers to adjust their learning strategies in response.

[42] In particular, we find that reduced spillovers explain 85% of the decline in wage growth for workers with 5–25 years of experience.

6.2 Impact of external learning subsidy

We now use the model to assess the impact of an external learning subsidy. This is motivated by the widespread use of policies reimbursing firms for formal training costs, such as professional trainers or lost production time, which align with external learning costs. After reviewing training subsidy policies in 21 US states (summarized in Table F.2), we focus on a 50% subsidy for external learning, which is the median policy among these states.

The last column of panel A in Table 6.1 shows the impact of the training subsidy in the baseline model. We find that a 50% subsidy for external learning increases average human capital by 6.34%, and raises wage growth by 32.93% for workers with 0–5 years of experience and by 60.95% for those with 5–25 years of experience. These gains follow not only from the reduced cost of external learning, but also from a better distribution of human capital within firms, as more workers learn from highly skilled trainers. This highlights the potential of targeted learning subsidies to boost human capital formation and stimulate aggregate productivity and economic growth.

The last column of panels B and C presents the impact of the training subsidy in the models without internal and external learning, respectively. In the model without external learning, these subsidies have no effect. In contrast, in the model without internal learning, the subsidy's effects on human capital are larger than in the baseline model, and the wage growth impacts heavily concentrate in the first five years of a worker's career. This happens because in this alternate model, young workers depend more on external learning to compensate for the lack of internal learning, greatly amplifying the effects of subsidies that reduce the cost of external learning. This result underscores the limitations of focusing on only one learning type when evaluating the effect of formal training subsidies: ignoring internal learning overstates the benefits for young workers, potentially creating unrealistic benchmarks for policymakers who often target this demographic group with learning policies.

7 Model extensions and robustness

In Appendix G, we examine the robustness of our results by considering several model extensions and alternative parameterizations. In particular, we assess robustness to: (1) incorporating on-the-job search; (2) adding time costs for internal trainers; (3) allowing initial human capital to vary across workers; (4) factoring in unemployment benefits; (5) adding a fixed learning cost to capture changes in the extensive margin of learning; and (6) considering different values of workers' bargaining power (β) and the elasticity of substitution between learning modes (σ). Detailed descriptions of each extension are provided in Appendices G.1–G.6.

We summarize the findings of each extension for: (1) the importance of each learning source for aggregate human capital in Table G.1; (2) the contribution of different components to wage growth in Table G.2; and (3) the impacts of remote work and training subsidies in Table G.3. Our quantitative results are highly consistent across all scenarios.

8 Conclusions

This paper explores how internal and external learning shape lifecycle human capital and wage dynamics. We document two novel facts underscoring the importance of these learning sources for both firms and workers. First, both internal and external learning are widely provided by firms, with larger firms offering better learning environments by providing a greater variety of learning options. Second, both sources of learning are important to workers, but have markedly different lifecycle patterns: internal learning declines with experience, while external learning follows an inverted U-shape over experience.

To shed light on the mechanisms driving these empirical findings and assess the importance of internal and external learning to lifecycle wage dynamics, we develop a model with a two-source learning framework. In the model, the incentives for each learning type change with age and shifts in workers' positions within their firm's human capital distribution. Consistent with our empirical findings, the calibrated model shows that internal learning is more critical for younger workers, while external learning becomes more significant as they age. The model also highlights the importance of considering both internal and external learning for fully capturing human capital and wage growth dynamics. Together, these two learning sources influence the non-own productivity dimension of human capital—encompassing learning costs and spillover effects among coworkers—and shape the effects of learning policies and shocks, given workers' ability to reallocate their learning time across sources.

Our findings have several implications. First, our results show that the role of human capital in lifecycle wage growth is multifaceted. In particular, we find that interpreting wage growth solely as a reflection of productivity may overlook the significance of non-own productivity compensation stemming from learning costs and coworker spillovers, especially for more-experienced workers. Second, our results show that policies that bridge the communication gap caused by remote work or create alternative learning avenues can yield both short-term and long-term benefits by enhancing the learning of remote workers, and bolstering their potential to serve as mentors in the future. Finally, our findings suggest that other forms of human capital, such as schooling, may also play a pivotal role in incentivizing on-the-job human capital accumulation by enriching the knowledge pool in the economy. Further exploring the interplay between various learning sources across different stages of the lifecycle is an interesting avenue for future research.

References

Abowd, J. M., K. L. McKinney, and N. L. Zhao (2018): "Earnings Inequality and Mobility Trends in the United States: Nationally Representative Estimates from Longitudinally Linked Employer-Employee Data," *Journal of Labor Economics*, 36, S183–S300. 28

Acemoglu, D. (1997): "Training and Innovation in an Imperfect Labour Market," *The Review of Economic Studies*, 64, 445–464. 1, 5, 13

Acemoglu, D. and W. Hawkins (2014): "Search with Multi-worker Firms," *Theoretical Economics*, 9, 583–628. 16, 22, 23

Acemoglu, D. and J.-S. Pischke (1998): "Why Do Firms Train?" *The Quarterly Journal of Economics*, 113, 79–119. 1, 5, 17

——— (1999): "Beyond Becker: Training in Imperfect Labour Markets," *The Economic Journal*, 109, 112–142. 5, 13

——— (2003): "Minimum wages and on-the-job training," *Research in Labor Economics*, 22, 159–202. 13

Akcigit, U., S. Caicedo, E. Miguelez, S. Stantcheva, and V. Sterzi (2018): "Dancing with the Stars: Innovation Through Interactions," *NBER Working Paper Series*. 1, 5

Altonji, J. G. and R. A. Shakotko (1987): "Do Wages Rise with Job Seniority?" *Review of Economic Studies*, 54, 437–459. 14

Altonji, J. G., A. A. Smith, and I. Vidangos (2013): "Modeling Earnings Dynamics," *Econometrica*, 81, 1395–1454. 23

Arellano-Bover, J. (2020): "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size," *Manuscript*. 3, 6, 8, 13, 27

Arellano-Bover, J. and F. Saltiel (2023): "Differences in On-the-Job Learning Across Firms," *Manuscript*. 3, 6, 8, 27

Autor, D. H. (2001): "Why Do Temporary Help Firms Provide Free General Skills Training?" *Quarterly Journal of Economics*, 116, 1409–1448. 5

Axtell, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293, 1818–1820. 23

Azoulay, P., J. Wang, and J. S. G. Zivin (2010): "Superstar Extinction," *The Quarterly Journal of Economics*, 125, 549–589. 5, 12

Bagger, J., F. Fontaine, F. Postel-Vinay, and J. M. Robin (2014): "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," *American Economic Review*, 104, 1551–1596. 1, 5, 16, 17, 20, 22, 23, 47, Supplemental Appendix - 23, Supplemental Appendix - 24

Barlevy, G. (2008): "Identification of Search Models Using Record Statistics," *The Review of Economic Studies*, 75, 29–64. 5

Barrero, J. M., N. Bloom, and S. J. David (2021): "Why Working from Home Will Stick," *NBER Working Paper Series*. 4, 32

Barth, E., A. Bryson, J. C. Davis, and R. Freeman (2016): "It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States," *Journal of Labor Economics*, 34, S67–S97. 28

Becker, G. S. (1962): "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70, 9–49. 1

——— (1964): *Human Capital*, Chicago: University of Chicago Press. 5, 13

Ben-Porath, Y. (1967): "The Production of Human Capital Over the Life Cycle," *Journal of Political Economy*, 75, 352–365. 1, 5, 13

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

Blundell, R. W., M. Costa Dias, D. Goll, and C. Meghir (2021): "Wages, Experience and Training of Women," *Journal of Labor Economics*, 39, 275–315. 15, 25

Bowlus, A. J. and H. Liu (2013): "The Contributions of Search and Human Capital to Earnings Growth over the Life Cycle," *European Economic Review*, 64, 305–331. 5

Burdett, K., C. Carrillo-Tudela, and M. G. Coles (2011): "Human Capital Accumulation and Labor Market Equilibrium," *International Economic Review*, 52, 657–677. 5

Burstein, A. and A. Monge-Naranjo (2009): "Foreign Know-How, Firm Control, and the Income of Developing Countries," *The Quarterly Journal of Economics*, 124, 149–195. 1, Supplemental Appendix - 31

Cahuc, P., F. Postel-Vinay, and J. M. Robin (2006): "Wage Bargaining with On-the-job Search: Theory and Evidence," *Econometrica*, 74, 323–364. 13, 20, 21, Supplemental Appendix - 21

Caicedo, S., R. E. Lucas, and E. Rossi-Hansberg (2019): "Learning, Career Paths, and the Distribution of Wages," *American Economic Journal: Macroeconomics*, 11, 49–88. 1, 5

Caldwell, S. and N. Harmon (2019): "Evidence from Coworker Networks," *Job Market Paper*. 13

Card, D., J. Kluve, and A. Weber (2018): "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations," *Journal of the European Economic Association*, 16, 894–931. 6

de la Croix, D., M. Doepke, and J. Mokyr (2016): "Clans, Guilds, and Markets:

Apprenticeship Institutions and Growth in the Pre-Industrial Economy," *The Quarterly Journal of Economics*, 133, 1–70. 5

Donovan, K., W. J. Lu, and T. Schoellman (2023): "Labor Market Dynamics and Development," *The Quarterly Journal of Economics*, 138, 2287–2325. 23

Emanuel, N., E. Harrington, and A. Pallais (2023): "The Power of Proximity Office Interactions Affect Online Feedback and Quits, Especially for Women and Young Workers," *NBER Working Paper Series*. 4, 5, 32, 33

Engbom, N. (2017): "Worker Flows and Wage Growth over the Life-Cycle: A Cross-Country Analysis," *Manuscript*. 3, 8, 27

——— (2021): "Labor Market Fluidity and Human Capital Accumulation," *Manuscript*. 6, 15, 25

Engbom, N., C. Moser, and J. Sauermann (2022): "Firm Pay Dynamics," *NBER Working Paper Series*. 6

Evans, C. and D. Lines (2014): "Which hat do I say I'm wearing?: identity work of independent coaching practitioners," *European Journal of Training and Development*, 38, 764–779, publisher: Emerald Group Publishing Limited. 16

Faberman, R. J., A. I. Mueller, A. Sahin, and G. Topa (2017): "Job Search Behavior among the Employed and Non-Employed," *NBER Working Paper Series*. Supplemental Appendix - 21

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015): "The Next Generation of the Penn World Table," *American Economic Review*, 105, 3150–3182. Supplemental Appendix - 23

Fitzenberger, B. and R. Völter (2007): "Long-run Effects of Training Programs for the Unemployed in East Germany," *Labour Economics*, 14, 730–755. 6

Friedrich, B., L. Laun, C. Meghir, and L. Pistaferri (2021): "Earnings Dynamics and Firm-Level Shocks," *NBER Working Paper Series*. 6

Garicano, L. (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108, 874–904. 1, 5

Garicano, L. and E. Rossi-Hansberg (2004): "Inequality and the Organization of Knowledge," *The American Economic Review*, 94, 197–202. 1, 5

——— (2006): "Organization and Inequality in a Knowledge Economy," *Quarterly Journal of Economics*, 1383–1435. 1, 5

Glaeser, E. (1999): "Learning in Cities," *Journal of Urban Economics*, 46, 254–277. 12

Gregory, V. (2019): "Firms as Learning Environments: Implications for Earnings Dynamics and Job Search," *Manuscript*. 1, 5, 6, 8, 15, 22, 27, Supplemental Appendix - 24

Hall, R. E. and A. B. Krueger (2012): "Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search," *American Economic Journal: Macroeconomics*, 4, 56–67. 13

Hamlin, R. G., A. D. Ellinger, and R. S. Beattie (2008): "The emergent coaching industry: a wake-up call for HRD professionals," *Human Re-*

source Development International, 11, 287–305, publisher: Routledge _eprint: https://doi.org/10.1080/13678860802102534. 15

HECKMAN, J. J. (1976): "A Life-Cycle Model of Earnings, Learning, and Consumption," *Journal of Political Economy*, 84, S11–S44. 5

HECKMAN, J. J., R. J. LALONDE, AND J. A. SMITH (1999): "The Economics and Econometrics of Active Labor Market Programs," *Handbook of Labor Economics*, 3, 1865–2097. 6

HERKENHOFF, K., J. LISE, G. MENZIO, AND G. M. PHILLIPS (2024): "Production and Learning in Teams," *Econometrica*, Forthcoming. 1, 5, 14, 41

HORNSTEIN, A., P. KRUSELL, AND G. L. VIOLANTE (2011): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *American Economic Review*, 101, 2873–98. Supplemental Appendix - 23

HUGGETT, M., G. VENTURA, AND A. YARON (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101, 2923–2954. 5, 13

JAROSCH, G. (2021): "Searching for Job Security and the Consequences of Job Loss," *NBER Working Paper Series*. 6

JAROSCH, G., E. OBERFIELD, AND E. ROSSI-HANSBERG (2021): "Learning from Coworkers," *Econometrica*, 89, 647–676. 1, 5, 12, 14, 17

JOVANOVIC, B. (2014): "Misallocation and Growth," *American Economic Review*, 104, 1149–1171. 5

KAMBOUROV, G. AND L. MANOVSKII (2009): "Occupational Specificity of Human Capital," *International Economic Review*, 50, 63–115. 14

KARAHAN, F., S. OZKAN, AND J. SONG (2022): "Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job Ladder Risk vs. Human Capital," *FRB St. Louis Working Paper Series*. 5

KLUVE, J. (2010): "The Effectiveness of European Active Labor Market Programs," *Labour Economics*, 17, 904–918. 6

LAGAKOS, D., B. MOLL, T. PORZIO, N. QIAN, AND T. SCHOELLMAN (2018): "Life Cycle Wage Growth across Countries," *Journal of Political Economy*, 126, 797–849. 23

LAZEAR, E. P. (2009): "Firm-Specific Human Capital: A Skill-Weights Approach," *Journal of Political Economy*, 117, 699–726. 14

LUCAS, R. E. (2009): "Ideas and growth," *Economica*, 76, 1–19. 5

LUCAS, R. E. AND B. MOLL (2014): "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 122, 1–51. 5

LUTTMER, E. G. J. (2014): "An Assignment Model of Knowledge Diffusion and Income Inequality," *Manuscript*. 1, 5, Supplemental Appendix - 31

MA, X., A. NAKAB, AND D. VIDART (2024): "Human Capital Investment and Development: The Role of On-the-job Training," *Journal of Political Economy Macroeconomics*, 2, 107–148. 1, 9

MANUELLI, R. E. AND A. SESHADRI (2014): "Human Capital and the Wealth of Nations," *American Economic Review*, 104, 2736–2762. 1, 25

Manuti, A., S. Pastore, A. F. Scardigno, M. L. Giancaspro, and D. Morciano (2015): "Formal and Informal Learning in the Workplace: A Research Review," *International Journal of Training and Development*, 19, 1–17. 14

McKenzie, D. (2017): "How Effective are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence," *World Bank Research Observer*, 32, 127–154. 6

Mincer, J. (1989): "Human Capital and the Labor Market: A Review of Current Research," *Educational Researcher*, 18. 15

Mincer, J. A. (1974): "The Human Capital Earnings Function," in *Schooling, Experience, and Earnings*, NBER, 83–96. 13

Moen, E. and A. Rosén (2004): "Does Poaching Distort Training?" *The Review of Economic Studies*, 71, 1143–1162. 1, 5, 13

Monge-Naranjo, A. (2019): "Markets, Externalities, and the Dynamic Gains of Openness," *International Economic Review*, 60, 1131–1170. 15

Nix, E. (2017): "Learning Spillovers in the Firm," *Manuscript*. 1, 5, 11

Perla, J. and C. Tonetti (2014): "Equilibrium Imitation and Growth," *Journal of Political Economy*, 122, 1–21. 5

Rosen, S. (1972): "Learning and Experience in the Labor Market," *The Journal of Human Resources*, 7, 326–342. 15

——— (1976): "A Theory of Life Earnings," *Journal of Political Economy*, 84, S45–S67. 5

Rubinstein, Y. and Y. Weiss (2006): "Chapter 1 Post Schooling Wage Growth: Investment, Search and Learning," *Handbook of the Economics of Education*, 1, 1–67. 1, 5

Shimer, R. (2005): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95, 25–49. 21, 22, 23

——— (2012): "Reassessing the Ins and Outs of Unemployment," *Review of Economic Dynamics*, 15, 127–148. 25

Song, J., D. Price, F. Guvenen, N. Bloom, and T. V. Wachter (2019): "Firming Up Inequality," *Quarterly Journal of Economics*, 134, 1–50. 28

Wallskog, M. (2023): "Entrepreneurial Spillovers Across Coworkers," *Manuscript*. 5

What Works - Centre for Local Economic Growth (2016): "Evidence Review 1: Employment Training," Tech. rep., London School of Economics. 6

Yamaguchi, S. (2010): "Job Search, Bargaining, and Wage Dynamics," *Journal of Labor Economics*, 28, 595–631. 5

A Appendix: Additional model results

A.1 Derivation of match value in Equation (4)

A.1.1 Incorporating spillovers in match value

Without accounting for spillovers, the match value between a worker i of age a_i and human capital h_{m_i}, and a firm with productivity z, employing a workforce $\mathbb{I}(z)$ with human capital distribution $\mathbb{H}(z)$, includes output net of training costs, and discounted expected future value depending on whether worker i gains human capital in the current period:

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

This match value omits spillovers within the firm, which occur because coworkers affect the firm's human capital distribution, impacting its internal learning environment and each worker's likelihood of climbing the human capital ladder. To incorporate spillovers in worker i's match value, we add the gains accruing to workers in this firm due to worker i's presence, and to avoid double counting, we deduct the spillovers perceived by worker i from other workers. These terms are captured in the second line of Equation (3), which expands Equation (A.1).

Herkenhoff et al. (2024) model spillovers in firms with a maximum of two workers. Building on their approach, we examine multi-worker firms while maintaining analytical tractability by: (1) assuming a continuum of workers to eliminate randomness, ensuring each hire faces the same learning environment in steady state; and (2) stipulating that production is independent across workers, with spillovers occurring only through human capital accumulation. As shown below, since spillovers perceived by worker i from others are approximately zero, our approach captures the marginal impact of hiring worker i on the total match value of all other workers in the firm.

A.1.2 Measuring the size of spillovers

We measure worker j's impact on worker i's match value by comparing worker i's match value when worker j is present in firm z's workforce in the current period with the match value if worker j were absent, holding all other factors constant. Using the match value in Equation (A.1), worker j's spillover effect on worker i's match value can be written as:

$$\begin{aligned}
&M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) - M^{a_i}(h_{m_i}, z, \mathbb{H}(z)\backslash\{j\}) = [zh_{m_i}(1 - l^{a_i}(h_{m_i}, z, \mathbb{H}(z))) - (1 - g^{a_i}(h_{m_i}, z, \mathbb{H}(z)))l^{a_i}(h_{m_i}, z, \mathbb{H}(z)) \times qp_e(h_{m_i})] \\
&- [zh_{m_i}(1 - l^{a_i}(h_{m_i}, z, \mathbb{H}(z)\backslash\{j\})) - (1 - g^{a_i}(h_{m_i}, z, \mathbb{H}(z)\backslash\{j\}))l^{a_i}(h_{m_i}, z, \mathbb{H}(z)\backslash\{j\}) \times qp_e(h_{m_i})] \\
&+ [s(h_{m_i}, \mathbb{H}(z)) - s(h_{m_i}, \mathbb{H}(z)\backslash\{j\})] \times \rho\mathbb{E}\left[(1-\kappa)M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\}\right] \\
&- [s(h_{m_i}, \mathbb{H}(z)) - s(h_{m_i}, \mathbb{H}(z)\backslash\{j\})] \times \rho\mathbb{E}\left[(1-\kappa)M^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) + \kappa \max\{V_U^{a_i+1}(h_{m_i}), V_{TR}^{a_i+1}(h_{m_i})\}\right] \\
&\approx \underbrace{\frac{\partial s(h_{m_i}, \mathbb{H}(z))}{\partial p(h_{m_i}, \mathbb{H}(z))}[p(h_{m_i}, \mathbb{H}(z)) - p(h_{m_i}, \mathbb{H}(z)\backslash\{j\})]}_{\text{Worker } j\text{'s marginal impact on worker } i\text{'s learning probability}} \times \underbrace{\rho\mathbb{E}\left[(1-\kappa)\Delta M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}'(z)) + \kappa\Delta \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\}\right]}_{\text{Increase in match value from raising worker } i\text{'s human capital}}
\end{aligned} \tag{A.2}$$

The last equality follows from the envelope theorem, as the optimal training time $l^{a_i}(h_{m_i}, z, \mathbb{H}(z)))$ and time allocation $g^{a_i}(h_{m_i}, z, \mathbb{H}(z)))$ maximize worker i's match value. Since we assess worker j's spillover effect in the current period, the future human capital distribution, $\mathbb{H}'(z)$, is treated as given.[43]

This spillover expression consists of two components. The first component reflects the change in coworker i's learning probability due to worker j's presence in the firm. The second component reflects the gain in worker i's match value from a one-step increase in human capital, capturing the net present value (NPV) of such an increase. Since workers make learning investments in every period, these spillovers provide a flow benefit analogous to that in a standard capital accumulation problem, where agents invest in capital at prevailing capital prices, reflecting the NPV of future capital rents.

A.1.3 Showing that spillovers generated by other workers equal zero

We now show that the spillovers accruing to worker i from all other workers in firm z are zero. Define $O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) = \left[(1-\kappa)\Delta M^{a_i+1}(h_{m_i+1}, z, \mathbb{H}'(z)) + \kappa\Delta \max V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\right]$, and use Equation (A.2) to note that these aggregate spillovers can be written as:

$$\begin{aligned}
&\underbrace{\int_{j\in\mathbb{I}(z), j\neq i} M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) - M^{a_i}(h_{m_i}, z, \mathbb{H}(z)\backslash\{j\})dj}_{\text{spillovers generated by other workers in firm accruing to worker } i} \\
&\approx \int_{j\in\mathbb{I}(z), j\neq i} \frac{\partial s}{\partial p}[p(h_{m_i}, \mathbb{H}(z)) - p(h_{m_i}, \mathbb{H}(z)\backslash\{j\})] \times O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z))dj \\
&= \frac{\partial s}{\partial p} O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z)) \int_{j\in\mathbb{I}(z), j\neq i} [p(h_{m_i}, \mathbb{H}(z)) - p(h_{m_i}, \mathbb{H}(z)\backslash\{j\})]\, dj
\end{aligned} \tag{A.3}$$

[43]While worker j's absence from the firm's workforce in the current period could affect the future human capital distribution, tracking this effect presents computational challenges. Furthermore, given the continuum of workers and the fact that the human capital distribution depends on all previous periods, the impact of a single worker in the current period is negligible.

This is approximately zero because the last integral in the expression, which represents the aggregate change in worker i's probability of matching with a coworker possessing higher human capital after the removal of every other worker j, equals zero.

To show this, we start by expressing the probability that worker i matches with a coworker of higher human capital as $p(h_{m_i}, \mathbb{H}(z)) = \frac{Z(h_{m_i},z)}{D(z)}$, where $D(z)$ is the measure of workers in firm z, and $Z(h_{m_i}, z)$ is the measure of workers with human capital higher than h_{m_i}. Thus, the aggregate change in worker i's probability of matching with a coworker with higher human capital is:

$$\int_{j\in\mathbb{I}(z),j\neq i} [p(h_{m_i}, \mathbb{H}(z)) - p(h_{m_i}, \mathbb{H}(z)\backslash\{j\})]\, dj = \int_{j\in\mathbb{I}(z),j\neq i} \left[\frac{Z(h_{m_i}, z)}{D(z)} - \frac{Z(h_{m_i}, z)\backslash\{j\}}{D(z)\backslash\{j\}}\right] dj$$

We split this integral into two parts for workers with human capital above and below that of worker i, noting that removing a worker j with $h_{m_j} \leq h_{m_i}$ does not affect $Z(h_{m_i}, z)$:

$$= \int_{j\in\mathbb{I}(z),j\neq i,h_{m_j}\leq h_{m_i}} \left[\frac{Z(h_{m_i}, z)}{D(z)} - \frac{Z(h_{m_i}, z)}{D(z)\backslash\{j\}}\right] dj + \int_{j\in\mathbb{I}(z),j\neq i,h_{m_j}>h_{m_i}} \left[\frac{Z(h_{m_i}, z)}{D(z)} - \frac{Z(h_{m_i}, z)\backslash\{j\}}{D(z)\backslash\{j\}}\right] dj.$$

Combining the two integrals, we get:

$$= \frac{D(z)}{D(z)\backslash\{1\}} \int_{j\in\mathbb{I}(z),j\neq i} \left(\mathbf{1}_{\{h_{m_j}>h_{m_i}\}} \frac{1}{D(z)} - \frac{p(h_{m_i}, \mathbb{H}(z))}{D(z)}\right) dj = \frac{D(z)}{D(z)\backslash\{1\}} [p(h_{m_i}, \mathbb{H}(z)) - p(h_{m_i}, \mathbb{H}(z))] = 0. \tag{A.4}$$

The last term follows because the integral sums to zero across all workers j. Here, $D(z)\backslash\{1\}$ represents the workforce size after the removal of one atomistic worker.

The first term in Equation (A.4) above captures the two forces mediating how worker j's presence in the firm affects the probability that worker i matches with a coworker of higher human capital: (1) an increase in the mass of potential mentors if worker j's human capital exceeds that of worker i, which increases this probability; and (2) an increase in the size of the workforce in firm z, which reduces this probability. When integrated across the entire workforce of the firm, these two effects cancel each other out. Specifically, the first force raises the probability of matching with a coworker with higher human capital by $\frac{1}{D(z)}$ for each eligible worker, but since only a measure $p(h_{m_i}, \mathbb{H}(z))D(z)$ of workers are eligible for this positive impact, its aggregate effect is given by $p(h_{m_i}, \mathbb{H}(z))$. Conversely, the second force reduces the probability of matching with a coworker with higher human capital by $\frac{p(h_{m_i},\mathbb{H}(z))}{D(z)}$ for every worker in the firm's workforce (measure $D(z)$), yielding an aggregate effect of $p(h_{m_i}, \mathbb{H}(z))$ which cancels out the first force exactly.

A.1.4 Analytical example to illustrate spillovers in match value

To further illustrate the intuition behind the match value and the role of spillovers, we present a simple discrete analytical example involving a firm with productivity z employing four workers, indexed by $i = 1, 2, 3, 4$, in the stationary equilibrium. These indices represent their human capital and seniority, such that $h_{m_4} > h_{m_3} > h_{m_2} > h_{m_1}$ and $a_4 > a_3 > a_2 > a_1$.

Table A.1 summarizes the demographic and learning details of these workers. To save on notation, we denote the internal learning probability of worker i as p_i; the expected gain in worker i's future match value from a one-step increase in human capital as O_i; and the overall probability of worker i climbing the human capital ladder as s_i.

Table A.1: Example of fictional firm with a discrete number of workers

Worker i	Age	h_{m_i}	p_i	$p_i \backslash \{j\}$ with $j =$ 1	2	3	4	$p_i - p_i \backslash \{j\}$ with $j =$ 1	2	3	4	Spillovers **to** other workers from i $\sum_{j \neq i} \left[(p_j - p_j \backslash \{i\}) \, O_j \frac{\partial s_j}{\partial p_j} \right]$	Spillovers **from** other workers to i $O_i \frac{\partial s_i}{\partial p_i} \sum_{j \neq i} [p_i - p_i \backslash \{j\}]$
1	a_1	h_{m_1}	1	X	1	1	1	X	0	0	0	$-1/3 \times O_2 \frac{\partial s_2}{\partial p_2} - 1/6 \times O_3 \frac{\partial s_3}{\partial p_3} + 0 \times O_4 \frac{\partial s_4}{\partial p_4}$	$O_1 \frac{\partial s_1}{\partial p_1} [0 + 0 + 0] = 0$
2	a_2	h_{m_2}	2/3	1	X	1/2	1/2	$-1/3$	X	1/6	1/6	$0 \times O_1 \frac{\partial s_1}{\partial p_1} - 1/6 \times O_3 \frac{\partial s_3}{\partial p_3} + 0 \times O_4 \frac{\partial s_4}{\partial p_4}$	$O_2 \frac{\partial s_2}{\partial p_2} [-1/3 + 1/6 + 1/6] = 0$
3	a_3	h_{m_3}	1/3	1/2	1/2	X	0	$-1/6$	$-1/6$	X	1/3	$0 \times O_1 \frac{\partial s_1}{\partial p_1} + 1/6 \times O_2 \frac{\partial s_2}{\partial p_2} + 0 \times O_4 \frac{\partial s_4}{\partial p_4}$	$O_3 \frac{\partial s_3}{\partial p_3} [-1/6 - 1/6 + 1/3] = 0$
4	a_4	h_{m_4}	0	0	0	0	X	0	0	0	X	$0 \times O_1 \frac{\partial s_1}{\partial p_1} + 1/6 \times O_2 \frac{\partial s_2}{\partial p_2} + 1/3 \times O_3 \frac{\partial s_3}{\partial p_3}$	$O_4 \frac{\partial s_4}{\partial p_4} [0 + 0 + 0] = 0$

The first panel of this table provides the ages, human capital levels, and internal learning probabilities of each worker. The second panel shows the internal learning probability each worker would have if each other worker were absent from firm z. The third panel presents the difference between these counterfactual probabilities and the actual internal learning probability, capturing the effect of each other worker j on worker i's learning probability.

The fourth panel presents the spillovers generated by worker i accruing to all other workers in the firm. These values reflect how worker i's presence affects the probability that other workers will climb the human capital ladder, multiplied by their marginal value of human capital. Consequently, workers with lower human capital generate negative spillovers, reducing their match value (and wages) since they deteriorate the firm's learning environment. Conversely, workers with higher human capital create positive spillovers that raise their match value (and wages) since they improve the firm's learning environment.

The final panel presents the spillovers accruing to worker i from other workers in the firm. These spillovers are zero for each worker, as the positive impact of mentors and the negative impact of other workers cancel each other out.

Using this information, we can express the match values for each worker, denoted for simplicity as $M^{a_i}(h_{m_i}, z)$ as follows:

- Worker 1:

$$M^{a_1}(h_{m_1}, z) = \underbrace{zh_{m_1}(1-l_1)}_{\text{output value}} - \underbrace{(1-g_1)l_1 \times qp_e(h_{m_1})}_{\text{pay for external trainers}} + \underbrace{-{}^1\!/_3 \times O_2 \frac{\partial s_2}{\partial p_2} - {}^1\!/_6 \times O_3 \frac{\partial s_3}{\partial p_3}}_{\text{spillovers to other workers from worker 1}} - \underbrace{0}_{\text{spillovers from other workers to worker 1}}$$
$$+ \underbrace{s_1 \times \rho \mathbb{E}\left[(1-\kappa)M^{a_1+1}(h_{m_1+1}, z) + \kappa \max\{V_U^{a_1+1}(h_{m_1+1}), V_{TR}^{a_1+1}(h_{m_1+1})\}\right]}_{\text{future value if worker 1 successfully climbs human capital ladder}}$$
$$+ \underbrace{[1-s_1] \times \rho \mathbb{E}\left[(1-\kappa)M^{a_1+1}(h_{m_1}, z) + \kappa \max\{V_U^{a_1+1}(h_{m_1}), V_{TR}^{a_1+1}(h_{m_1})\}\right]}_{\text{future value if worker 1 doesn't successfully climb the human capital ladder}}$$

- Worker 2:

$$M^{a_2}(h_{m_2}, z) = \underbrace{zh_{m_2}(1-l_2)}_{\text{output value}} - \underbrace{(1-g_2)l_2 \times qp_e(h_{m_2})}_{\text{pay for external trainers}} + \underbrace{-{}^1\!/_6 \times O_3 \frac{\partial s_3}{\partial p_3}}_{\text{spillovers to other workers from worker 2}} - \underbrace{0}_{\text{spillovers from other workers to worker 2}}$$
$$+ \underbrace{s_2 \times \rho \mathbb{E}\left[(1-\kappa)M^{a_2+1}(h_{m_2+1}, z) + \kappa \max\{V_U^{a_2+1}(h_{m_2+1}), V_{TR}^{a_2+1}(h_{m_2+1})\}\right]}_{\text{future value if worker 2 successfully climbs human capital ladder}}$$
$$+ \underbrace{[1-s_2] \times \rho \mathbb{E}\left[(1-\kappa)M^{a_2+1}(h_{m_2}, z) + \kappa \max\{V_U^{a_2+1}(h_{m_2}), V_{TR}^{a_2+1}(h_{m_2})\}\right]}_{\text{future value if worker 2 doesn't successfully climb the human capital ladder}}$$

- Worker 3:

$$M^{a_3}(h_{m_3}, z) = \underbrace{zh_{m_3}(1-l_3)}_{\text{output value}} - \underbrace{(1-g_3)l_3 \times qp_e(h_{m_3})}_{\text{pay for external trainers}} + \underbrace{{}^1\!/_6 \times O_2 \frac{\partial s_2}{\partial p_2}}_{\text{spillovers to other workers from worker 3}} - \underbrace{0}_{\text{spillovers from other workers to worker 3}}$$
$$+ \underbrace{s_3 \times \rho \mathbb{E}\left[(1-\kappa)M^{a_3+1}(h_{m_3+1}, z) + \kappa \max\{V_U^{a_3+1}(h_{m_3+1}), V_{TR}^{a_3+1}(h_{m_3+1})\}\right]}_{\text{future value if worker 3 successfully climbs human capital ladder}}$$
$$+ \underbrace{[1-s_3] \times \rho \mathbb{E}\left[(1-\kappa)M^{a_3+1}(h_{m_3}, z) + \kappa \max\{V_U^{a_3+1}(h_{m_3}), V_{TR}^{a_3+1}(h_{m_3})\}\right]}_{\text{future value if worker 3 doesn't successfully climb the human capital ladder}}$$

- Worker 4:

$$M^{a_4}(h_{m_4}, z) = \underbrace{zh_{m_4}(1-l_4)}_{\text{output value}} - \underbrace{(1-g_4)l_4 \times qp_e(h_{m_4})}_{\text{pay for external trainers}} + \underbrace{{}^1\!/_6 \times O_2 \frac{\partial s_2}{\partial p_2} + {}^1\!/_3 \times O_3 \frac{\partial s_3}{\partial p_3}}_{\text{spillovers to other workers from worker 4}} - \underbrace{0}_{\text{spillovers from other workers to worker 4}}$$
$$+ \underbrace{s_4 \times \rho \mathbb{E}\left[(1-\kappa)M^{a_4+1}(h_{m_4+1}, z) + \kappa \max\{V_U^{a_4+1}(h_{m_4+1}), V_{TR}^{a_4+1}(h_{m_4+1})\}\right]}_{\text{future value if worker 4 successfully climbs human capital ladder}}$$
$$+ \underbrace{[1-s_4] \times \rho \mathbb{E}\left[(1-\kappa)M^{a_4+1}(h_{m_4}, z) + \kappa \max\{V_U^{a_4+1}(h_{m_4}), V_{TR}^{a_4+1}(h_{m_4})\}\right]}_{\text{future value if worker 4 doesn't successfully climb human capital ladder}}$$

Adding up these match values for all workers in firm z gives the aggregate workforce value:

$$
\begin{aligned}
\sum_i M^{a_i}(h_{m_i}, z) = & \underbrace{\sum_i z h_{m_i}(1-l_i)}_{\text{output values}} - \underbrace{\sum_i (1-g_i) l_i \times q p_e(h_{m_i})}_{\text{pay for external trainers}} \\
& + \underbrace{\sum_i s_i \times \rho \mathbb{E}\left[(1-\kappa) M^{a_i+1}(h_{m_i+1}, z) + \kappa \max\{V_U^{a_i+1}(h_{m_i+1}), V_{TR}^{a_i+1}(h_{m_i+1})\}\right]}_{\text{future values if workers successfully climb human capital ladder}} \\
& + \underbrace{\sum_i [1-s_i] \times \rho \mathbb{E}\left[(1-\kappa) M^{a_i+1}(h_{m_i}, z) + \kappa \max\{V_U^{a_i+1}(h_{m_i}), V_{TR}^{a_i+1}(h_{m_i})\}\right]}_{\text{future values if workers don't successfully climb human capital ladder}}
\end{aligned}
$$

This expression shows that spillover effects cancel out when we aggregate them across all workers, as they represent a reapportioning of production value towards higher human capital workers and away from lower human capital workers. This expression confirms that there is no double-counting, and that the aggregate workforce value ultimately comes from production.

A.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 of productivity z, employing workforce $\mathbb{I}(z)$ with human capital distribution $\mathbb{H}(z)$, and perceiving a revenue share r, is given by:[44]

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

(A

Firm's value. The value for a firm with productivity z, employing 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 receives a revenue share r, is given by:

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

(A.6

[44] As before, the spillover effects from other workers in the firm accruing to worker i are approximately zero and thus excluded from both the worker's and firm's values.

Unemployment value. Following Bagger et al. (2014), we assume unemployment is equivalent to employment at the least productive firm. The value for an unemployed worker of age a_i and human capital h_{m_i} is:

$$V_U^{a_i}(h_{m_i}) = \min_z M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) \tag{A.7}$$

Trainer's value. The value for a worker of age a_i and human capital h_{m_i} of being a trainer is:

$$V_{TR}^{a_i}(h_{m_i}) = h_{m_i} q F_e(h_{m_i - 1}) + \rho \mathbb{E} \max\{V_U^{a_i+1}(h_{m_i}), V_{TR}^{a_i+1}(h_{m_i})\}. \tag{A.8}$$

A.3 Proof of Proposition 1

Firms and workers jointly maximize the match value in Equation (4). 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 $\frac{\gamma s^E(h_{m_i}, \mathbb{H}(z)) O^{a_i+1}(h_{m_i}, z, \mathbb{H}'(z))}{z h_{m_i} + (1-g) q p_e(h_{m_i})} < 1$ holds, which ensures internal solutions (and is satisfied in our calibration for all worker and firm types), the first-order condition with respect to l implies:

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

The first-order condition with respect to g yields:

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

Combining the above equations implies:

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

All else being equal, the partial derivatives of this equation show that g increases with $p(h_{m_i}, \mathbb{H}(z))$ and decreases with q.

A.4 Definition of equilibrium

Define $D^a(h_m, z)$ and $D_t^a(h_m)$ as the measures of employed workers of age a and human capital h_m at firm z and in the training sector, respectively. The equilibrium of the model is defined as follows:

Definition 1. *The general equilibrium consists of meeting rates λ_U, employment distributions $\{D_U^a(h_m), D^a(h_m, z), D_t^a(h_m)\}$, firms' vacancy postings $v(z)$, learning investments $\{g^a(h_m, z, \mathbb{H}(z)), l^a(h_m, z, \mathbb{H}(z))\}$, wage rates $w^a(h_m, z, \mathbb{H}(z))$, and the price of external training services q. These variables satisfy:*

1. *The wage rate* $w^a(h_m, z, \mathbb{H}(z))$ *is determined by the bargaining process described in Section 4.1.4.2.*

2. *Learning investments* $\{g^a(h_m, z, \mathbb{H}(z)), l^a(h_m, z, \mathbb{H}(z))\}$ *satisfy Equations (5) and (6) in Proposition 1, where the learning probabilities from internal and external sources are determined by the employment distributions within each firm* ($D^a(h_m, z)$) *and the training sector* ($D^a_t(h_m)$)*.*

3. *Firms' optimal vacancy postings* $v(z)$ *are given by Equation* (10)*.*

4. *The meeting rate* λ_U *is determined by the ratio of unemployed workers* ($U = \sum_a \sum_m D^a_U(h_m)$) *to the total number of vacancies* ($V = \int v(z)d\Phi(z)$)*.*

5. *The employment distributions* $\{D^a_U(h_m), D^a(h_m, z), D^a_t(h_m)\}$ *are consistent with vacancy postings* $v(z)$*, learning investments* $\{g^a(h_m, z, \mathbb{H}(z)), l^a(h_m, z, \mathbb{H}(z))\}$*, exogenous job separations, and unemployed workers' sectoral choices.*

6. *The training price* q *clears the market for external training services such that the total demand for external training services (aggregated external learning time across all production workers) equals the total supply (external learning units provided by all external trainers).*

A.5 Solving algorithm

1. **Firm productivity and human capital grids**: We first divide the firm productivity distribution into 500 equally sized bins, selecting a representative firm at the midpoint of each bin. The human capital ladder consists of 40 steps,[45] starting at the lowest level $h_1 = 1$, with a step size of $\gamma_h = 0.05$.

2. **Initial guesses**: We then start with initial guesses for several variables, including the match values for each firm and worker type, $M^a(h_m, z, \mathbb{H}(z))$, the value of working in the training sector $V^a_{TR}(h_m)$, labor market tightness θ, vacancy distribution $G(z) = \int_0^z v(y)d\Phi(y)/\int_0^\infty v(y)d\Phi(y)$, and the price of external learning services q. From match values for the lowest productivity firm, we can also compute unemployment values for each worker type.

3. **Optimal learning investments**: Given the match, unemployment, and trainer values,

[45] In the baseline equilibrium, the average level of human capital is around 5 steps, with almost no workers exceeding 10 steps on the ladder.

we solve for the optimal learning investments for each firm and worker type, $\{g^a(h_m, z, \mathbb{H}(z)), l^a(h_m, z, \mathbb{H}(z))\}$ using Equations (5) and (6) in Proposition 1. Also, we find the employment distributions $\{D_U^a(h_m), D^a(h_m, z), D_t^a(h_m)\}$ using these values, labor market tightness, and vacancy distribution.

4. **Update values and vacancies**: We then use the optimal learning investments and labor market tightness to compute the implied match and trainer values using Equations (4) and (A.8) through backward induction, starting from the end of working life. Given the employment distributions and updated match values, we compute the optimal number of vacancies posted by each firm $v(z)$ using Equation (10), and labor tightness $\theta = \int v(z) d\Phi(z) / \sum_m \sum_a D_U^a(h_m)$. We also adjust the price of external training services if demand exceeds (or falls short of) supply.

5. **Convergence and wages**: We iterate on Steps 3–4 until convergence in match values, trainer values, labor market tightness, vacancy distribution, and the price of external learning services. Finally, we compute wage rates in the stationary equilibrium using the Nash bargaining process specified in Section 4.1.4.2.

Supplemental Appendix

B Data details

B.1 EU Continuing Vocational Training Survey

The European Union's Continuing Vocational Training Enterprise Survey (EU-CVT) gathers data from firms in the EU, and focuses on their investments in continuing vocational training for their staff within the last calendar year. CVT surveys have been conducted in 1993, 1999, 2005, 2010, and 2015. Due to data availability, we focus on the 2005, 2010, and 2015 surveys (CVT3, CVT4, and CVT5). These surveys sample 78,000, 101,000, and 111,000 firms, respectively, from all EU member states and Norway. This data mainly focuses on firms with 5 or more workers.

B.1.1 Internal and external learning

Firms are classified as offering internal (external) learning if they provide "internal (external) CVT courses" or "other forms of CVT" that utilize the internal (external) knowledge pool. We now examine each of these categories in detail.

- CVT courses refer to planned and organized education or training activities financed at least partially by the enterprise, with an explicit goal of learning. Specifically, these activities aim to help employees with "the acquisition of new competences or the development and improvement of existing ones." These courses are typically conducted away from the active workplace, such as in a classroom or training institution, are highly organized, and include content explicitly designed for a group of learners, such as a curriculum. 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 the workplace, but can also include participation in conferences, trade fairs, and similar events. Specifically, we use the following types of "other forms of CVT":[46]

[46] Among the "other types of CVT activities," the survey also includes self-directed learning, which is more

– Guided on-the-job training: Training in the workplace using the normal tools of work with a tutor present. This is categorized as internal learning.

– Job rotation, exchanges, secondments, or study visits: Visits to other departments or enterprises for learning purposes. This is categorized as internal learning.

– Learning or quality circles: Groups of employees who come together on a regular basis with the primary aim of learning. This is categorized as internal learning.

– Participation in conferences, workshops, trade fairs, and lectures. This is categorized as external learning.

B.1.2 Summary statistics in EU-CVT data

Table B.1: Share of firms providing CVT courses and other types of CVT activities

	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
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 the five largest countries 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.

B.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 gathered information from firms and randomly selected employees in establishments with 50 or more workers in the US.[47] The employer portion of the survey examined the intensity and costs of employer-provided formal training, while the employee portion focused on the time spent on both formal and informal training. The

akin to learning-by-doing and thus not considered here.

[47] 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.

survey sampled 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, so we rely on aggregate statistics to support the results in Section 3.2. We also use this data to calibrate the model in Section 4.

B.2.1 Internal and external learning

To construct measures of external and internal learning in this dataset, we rely on information about formal and informal training investments, which roughly (though not perfectly) correspond to each of these, respectively.

- Formal training involves planned activities with a structured format and defined curriculum. Examples include attending company-led classes, participating in seminars conducted by professional trainers, or watching pre-planned audio-visual presentations.
- Informal training is unstructured, unplanned, and easily adaptable to situations or individuals. Examples include colleagues showing each other how to use equipment or supervisors teaching job-related skills.[48]

B.3 German Worker Qualification Survey

The Worker Qualification Survey, conducted by the BIBB (Federal Institute for Vocational Education, Bonn), together with the IAB (until 1999) and BAuA (after 1999), spans 7 waves conducted in 1979, 1985, 1992, 1999, 2006, 2012, and 2018. Each wave covers a representative sample of 20,000 to 35,000 members of the labor force (excluding apprentices). The survey offers comprehensive data for analyzing both the cross-sectional and temporal evolution of qualifications and working conditions of the German workforce.

This data has two important limitations. First, there is variation in the questions asked across survey waves, affecting the comparability of the skill acquisition measures over time. However, this should not necessarily affect the aggregate lifecycle patterns of on-the-job learning. More importantly, Table D.1 indicates that our results are robust to controlling for survey wave fixed effects. Second, the survey's response rate is relatively low, reaching levels as low as 44%. To address this, we adjust our results using the weighting schemes provided by BIBB to account for selection probabilities and selective refusals.

[48] It is worth noting that in the survey, informal training is distinct from self-learning activities such as learning-by-doing or learning through hobbies.

B.3.1 Internal learning, external learning, experience, and wages

- Internal learning is constructed as a binary variable indicating whether an individual acquired the skills or knowledge required for their current job through colleagues or superiors. The relevant question remains relatively consistent across surveys, except for: (1) the 1979 survey, which did not differentiate between learning-by-doing and internal learning and is therefore excluded; (2) the 2006 survey, which asked about receiving professional development through coaching from superiors; and (3) the 2011/2012 and 2017/2018 surveys, which did not include a related question and are therefore excluded. There is a slight change in the questions in the 1979–1999 surveys: individuals could list all sources of skill acquisition in 1979 and 1986, while in 1992 and 1999, they listed only the two main sources.

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

- Potential experience is constructed as: $Age - Years\ of\ schooling - 6$. $Years\ of\ schooling$ is constructed using: $Year\ of\ graduation\ of\ highest\ degree - Birth\ year - 6$.

- The number of years with the current employer is derived directly from the corresponding survey variable for the years prior to 1998/1999. For later waves, it is calculated as the difference between the current year and the year the individual started with their current employer. Promotions within the same company are not considered as employer switches. For self-employed workers or business owners, this variable reflects the years since the start of running the current business or occupation.

- Hourly wages are constructed using data on the monthly wage and regular hours worked. In earlier surveys, monthly wages were recorded as ordinal values, with respondents selecting from wage range options. In later surveys, exact wage figures were provided. Therefore, for early waves, individual wages are imputed using the midpoint of the reported wage range. Wages are adjusted for inflation using the German CPI with base year 2015, and currency is adjusted to reflect the switch to the Euro.

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

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

B.4.1 Internal learning, external learning, experience, and wages

- Internal learning is constructed as a binary variable taking a value of one for workers who reported receiving instruction or training from a coworker or supervisor in their last work-experience program, and zero for all other surveyed workers.[49]

- External learning is constructed as a binary variable taking 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. For all other workers, this variable takes a value of zero.[50]

[49] Work-experience programs encompass job roles with learning attributes such as internships, co-ops, practicums, clerkships, externships, residencies, clinical experiences, apprenticeships, or other learning components. In panel (a) of Figure D.1, we show that our documented lifecycle pattern of internal learning is robust to considering only individuals reporting participation in a work-experience program.

[50] In panels (c) and (d) of Figure D.1, we illustrate the robustness of our documented lifecycle pattern of external learning when decomposing across the two learning components: taking classes during a last work-experience program, or ever receiving an employment-related training certificate. In panel (b) of

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

- Hourly wages are constructed using data on yearly work earnings, weeks worked, and regular hours worked. Yearly work earnings and weeks worked are recorded in an ordinal fashion and are thus imputed using the midpoint of the reported range. Wages are reported in dollars of 2016.

B.4.2 Summary statistics in German BIBB/BAuA and US NHES data

Table B.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.

Figure D.1, we show that our documented lifecycle pattern of external learning is robust to considering only individuals reporting participation in a work-experience program.

B.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 across more than 40 OECD countries surveyed in different rounds. In total, the survey covers a sample of around 230,000 individuals.

While PIAAC provides rich information for constructing measures of internal and external learning for workers in participating countries, it has important limitations. First, the sample sizes in each country are relatively small, ranging from about 700 to 2,500 observations suitable for our analysis, further reduced when dividing samples across experience bins. Second, privacy laws in some countries exclude several important variables. After refinements and constructing our main variables of interest, the countries included in our analysis from each survey round are:

- 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.

B.5.1 Internal learning, external learning, experience, and wages

- Internal learning is constructed as a binary variable taking a value of one for workers who either reported participating in courses led by coworkers/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: $Age - Years\ of\ schooling - 6$. $Years\ of\ schooling$ follows from the mapping between educational attainment and years of schooling provided by PIAAC. We omit workers with an educational attainment of less than secondary since we cannot directly map this into years of schooling.

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

B.5.2 Summary statistics in PIAAC data

Table B.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.

C Robustness of Fact 1

C.1 EU-CVT data

Table C.1: 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 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 C.2: 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 C.3: 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 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C.2 US-SEPT Data

Table C.4: 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.

D Robustness of Fact 2

D.1 German BIBB/BAuA and US NHES data

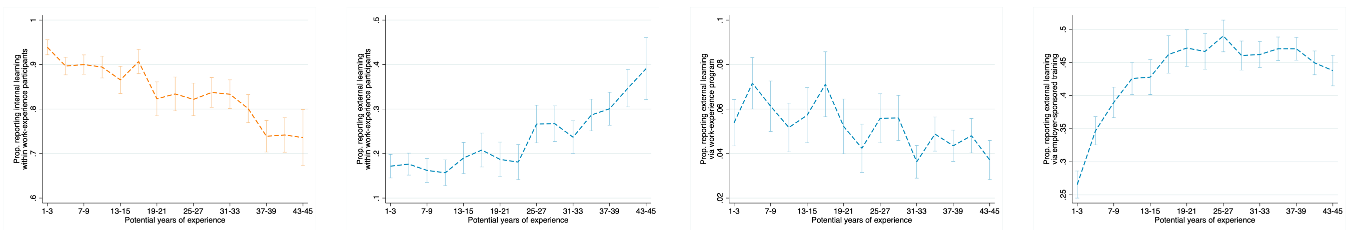
Table D.1: Correlations between internal and external learning and potential experience

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 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure D.1: Decomposing different types of internal and external learning in the US

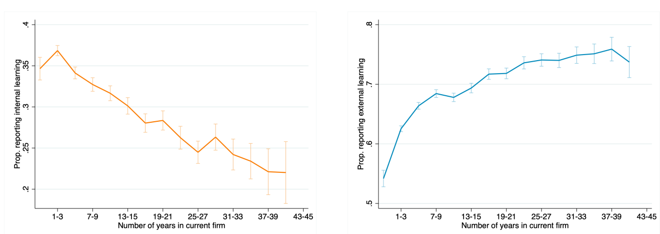
(a) Prevalence of internal learning within work-experience program participants | (b) Prevalence of external learning within work-experience program participants | (c) External learning via participation in work-experience program | (d) External learning via certificate in employer-sponsored training



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 D.2: 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 D.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

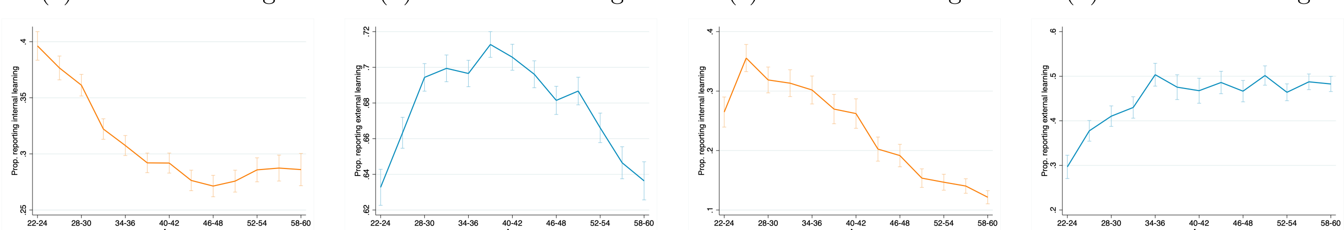
Notes: This table shows the coefficients from regressing internal and external learning on years with current employer (tenure) and years with current employer squared in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed, and have between 1 and 45 years of potential experience. Regressions are weighted using the 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D.1.1 Alternate specifications

Figure D.3: Sources of learning by age

Germany — (a) Internal learning; (b) External learning

United States — (c) Internal learning; (d) External learning

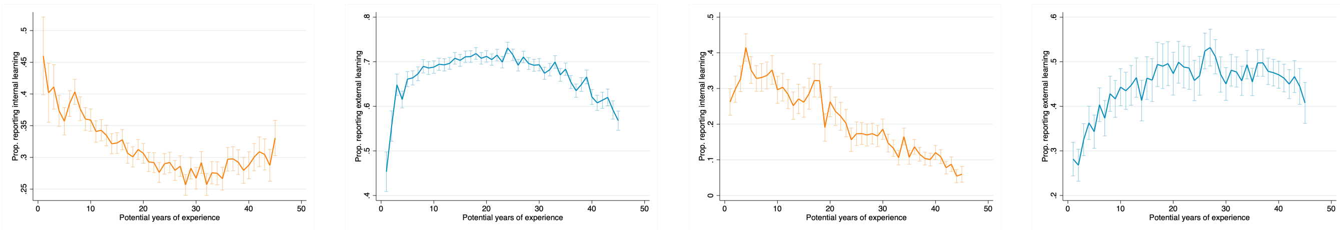


Notes: These figures plot the proportion of workers reporting engaging in internal and external learning across different age bins in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed, have between 1 and 45 years of potential experience, and are 21–60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

Figure D.4: 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

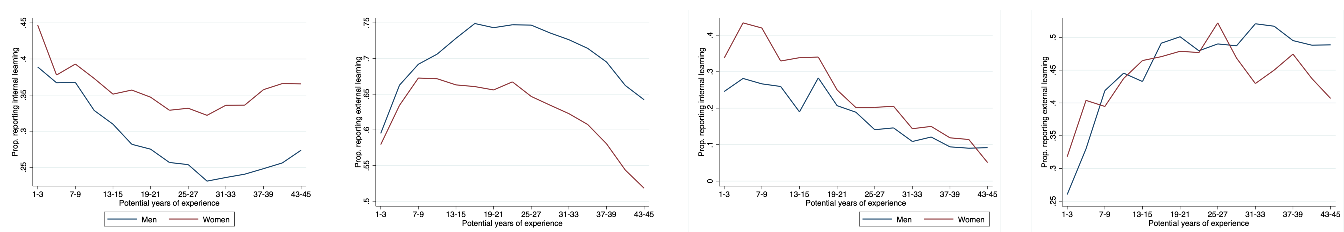


Notes: These figures plot the proportion of workers reporting engaging in internal and external learning across different levels of potential experience in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have between 1 and 45 years of potential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

Figure D.5: Sources of learning throughout workers' lifecycles by gender

Germany | **United States**

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

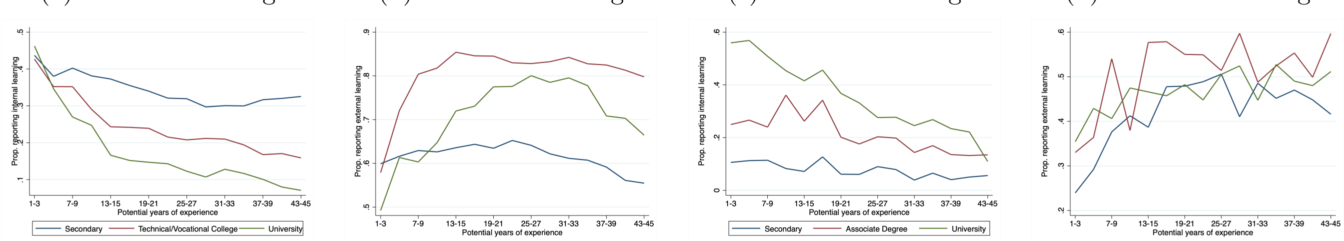


Notes: These figures plot the proportion of male and female workers reporting engaging in internal and external learning across different potential experience bins in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have between 1 and 45 years of 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 D.6: Sources of learning throughout workers' lifecycles by educational level

Germany | **United States**

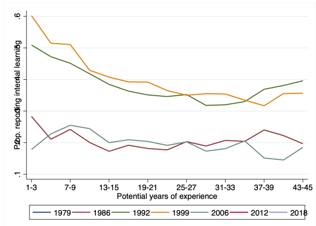
(a) Internal learning (b) External learning (c) Internal learning (d) External learning



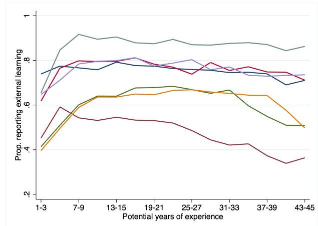
Notes: These figures plot the proportion of workers with different educational attainment levels reporting engaging in internal and external learning across different potential experience bins in the German BIBB/BAuA and US NHES data. The samples encompass individuals who are currently employed and have 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 D.7: Sources of learning throughout workers' lifecycles by survey wave in Germany

(a) Internal learning



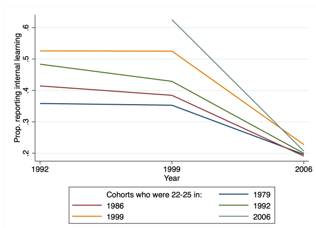
(b) External learning



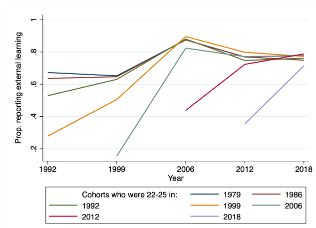
Notes: These figures plot the proportion of workers reporting engaging in internal and external learning across different potential experience bins in different survey waves of the German BIBB/BAuA data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of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

Figure D.8: Sources of learning throughout workers' lifecycles by cohort in Germany

(a) Internal learning



(b) External learning

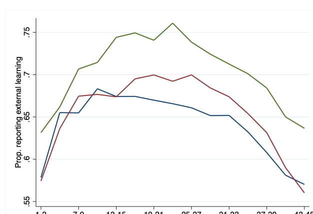


Notes: These figures plot the proportion of workers of different cohorts reporting engaging in internal and external learning across different potential experience bins in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of potential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

Figure D.9: Sources of learning throughout workers' lifecycles by firm size in Germany

(a) Internal learning

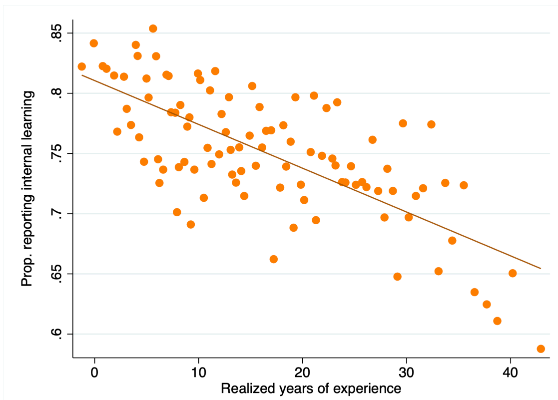
(b) External learning

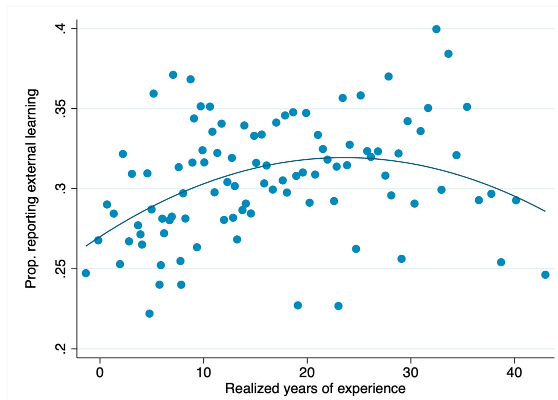


Notes: These figures plot the proportion of workers employed in firms of different sizes reporting engaging in internal and external learning across different potential experience bins in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed and have between 1 and 45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

D.2 OECD PIAAC data

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





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, 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.

Table D.3: 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 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E Literature review of internal and external learning

Table E.1: Studies examining the effects of internal and external learning on wages and human capital accumulation

Panel A: Results concerning external learning

Study	Period	Dataset	Sample	External learning measure	Outcome, Estimation method	Effect sizes
Lillard and Tan (1986)	1983	CPS	Men	Incidence of formal OJT	Log wage, OLS	0.044***
Lynch (1992)	1980–1983	NLSY79	Young non-college graduates	Weeks of formal OJT	Log wage,, OLS+SC Log wage, FE	0.002 -0.0002
Bishop (1994)	1982–1987	EOPP-NCRVE, NFIBS	New hires	Incidence of formal OJT	Log starting wage, OLS Log current wage, OLS Subjective starting prod., OLS Subjective current prod., OLS	0.019 -0.013 0.095*** -0.003
Bartel (1995)	1986–1990	Personnel records for Manufacturing Co.	Professional employees	Incidence of formal OJT Days of OJT	Log wage, IV+FE	0.106*** 0.016***
Veum (1995)	1986–1990	NLSY79	Young workers	Hours of formal OJT Incidence of formal OJT	Log wage, OLS Log wage, FE Log wage, OLS Log wage, FE	0.004 0.008 0.0728* 0.0897**
Krueger and Rouse (1998)[1]	1991–1995	Records for manufacturing co. Records for service co.	Production workers All workers	Incidence of formal OJT	Log wage, FE	0.004*** -0.002
Loewenstein and Spletzer (1999)	1988-1991	NLSY79	Young workers	Number of OJT Spells	Log wage, OLS	0.044**
Parent (1999)	1979–1991	NLSY79	Young workers	Years of formal OJT	Log wage, OLS Log wage, HT	0.1692*** 0.1216***
Marcotte (2000)	1966-1981 1979-1994	NLS NLSY79	White young men	Incidence of formal OJT	Log wage, OLS	0.086** 0.090**
Frazis and Loewenstein (2005)[2]	1979-2000	NLSY79	Young workers	Hours of formal OJT	Log wage, FE Log wage, FE (cube root spec.)	0.003 0.036
Hamil-Luker (2005)	1977–1987 1988–1998	NLS young women NLSY79	Young women Young Women	Incidence of formal OJT	Log wage, RE	0.05*** 0.07***
Blanchflower and Lynch (2007)	1979–1988	NLSY79	Young non-college graduates	Incidence of formal OJT	Log wage, OLS	0.08

Panel B: Results concerning internal learning

Study	Period	Dataset	Sample	Internal learning measure	Outcome, Estimation method	Effect sizes
Lillard and Tan (1986)	1983	CPS	Men	Incidence of informal OJT	Log wage, OLS	0.224***
Azoulay et al. (2010)	1975–2006	Faculty roster + Superstar list	US medical school faculty	Death of superstar coauthor Sudden death of superstar coauthor Anticipated death of superstar coauthor	Publication rate, OLS	-0.074*** -0.105*** -0.047
Nix (2017)	1985–2012	Swedish employer-employee	Prime-age men	Average education of coworkers	Log wage, OLS Log wage, FE Log wage, richer FE[3]	0.191*** 0.049*** 0.031***
Akcigit et al. (2018)	1977–2010	European patents	Inventors	No. ↑ quality co-inventors No. ↑ quality co-inventors in same firm	Patent citations, FE Patent citations, FE	0.021*** 0.00004***
Jarosch et al. (2021)	1993–2010	German employer-employee	All workers	Log mean wage of coworkers Log mean wage of coworkers w/ ↑ wages Log mean wage of coworkers w/ ↓ wages	Log wage in 1 year, FE Log wage in 5 years, FE Log wage in 1 year, FE Log wage in 5 years, FE Log wage in 1 year, FE Log wage in 5 years, FE	0.071*** 0.17*** 0.09*** 0.22*** 0.019*** 0.060***
Emanuel et al. (2023)	2019–2020	Records for Fortune 500 firm	Software engineers	Office closure for Covid-19	No. comments/program, DiD No. comments/program, DiD+FE	-1.29*** -1.17**
Wallskog (2023)	1993–2013	US employer-employee	All workers in 18 states	Share of entrepreneur coworkers	Future entrepreneurship , OLS Future entrepreneurship, FE+controls	0.105*** 0.025***
Herkenhoff et al. (2024)	2001–2008	US employer-employee	Prime-age men in 11 states, EUE & w/ wage lower than mean of coworkers Prime-age men in 11 states, EUE[4] & w/ wage higher than mean of coworkers	Log mean wage of coworkers	Log wage in 2 years, FE	0.145*** 0.041***

Notes: *Estimation Method Acronyms:* OLS denotes ordinary least squares. FE denotes individual/team and other types of fixed effects. SC denotes various forms of selection correction (Heckit, Hurdle, etc). IV is instrumental variable approach. HT corresponds to the Hausmann-Taylor method. RE denotes random effects model. DiD denotes difference in differences. *Effects Considered:* We focus on studies examining the separate effects of internal and external learning, leaving out studies that consider the impact of internal and external learning composites. If multiple specifications or outcomes are explored in one study, these are noted separately. The estimates mostly refer to the effects of completed learning investments, though in some studies this may not be the case. In studies where the distinction between learning with current and previous employers is made, we consider learning completed with the current employer generally. [1] Although the training programs considered here were partially sponsored by the government, we still include them as part of firm-provided training since they targeted workers that were already employees at the firm, the content was tailored to the specific needs and requests of the firm, and some of the costs were assumed by the firm. [2] This study reports the marginal effects of training at the median. No standard errors are provided. [3] The richer FEs in this study consider both individual FE, and County x Industry x Year FE. [4] For this study, EUE refers to individuals who transitioned from employment at a firm in year t into unemployment (for at least one quarter) in year $t+1$, and then back into employment at a different firm in $t+2$.

Appendix E References

Akcigit, U., S. Caicedo, E. Miguelez, S. Stantcheva, and V. Sterzi (2018): "Dancing with the Stars: Innovation Through Interactions," *NBER Working Paper Series.* 1, 5

Azoulay, P., J. Wang, and J. S. G. Zivin (2010): "Superstar Extinction," *The Quarterly Journal of Economics*, 125, 549–589. 5, 12

Bartel, A. P. (1995): "Training, Wage Growth, and Job Performance: Evidence from a Company Database," *Journal of Labor Economics*, 13, 401–425.

Bishop, J. H. (1994): "The Impact of Previous Training on Producicitivity and Wages," in *Training and the Private Sector*, ed. by L. M. Lynch, Chicago: University of Chicago Press, 161–200.

Blanchflower, D. G. and L. M. Lynch (2007): "Training at Work : A Comparison of U.S. and British Youths," in *Training and the Private Sector: International Comparisons*, ed. by L. M. Lynch, Chicago: University of Chicago Press, 233–260.

Emanuel, N., E. Harrington, and A. Pallais (2023): "The Power of Proximity Office Interactions Affect Online Feedback and Quits, Especially for Women and Young Workers," *NBER Working Paper Series.* 12

Frazis, H. and M. A. Loewenstein (2005): "Reexamining the Returns to Training: Functional Form, Magnitude, and Interpretation," *Journal of Human Resources*, 40, 453–476.

Hamil-Luker, J. (2005): "Women's Wages: Cohort Differences in Returns to Education and Training over Time," *Social Science Quarterly*, 86, 1261–1278.

Herkenhoff, K., J. Lise, G. Menzio, and G. M. Phillips (2024): "Production and Learning in Teams," *Econometrica*, Forthcoming. 1, 5, 14, 41

Jarosch, G., E. Oberfield, and E. Rossi-Hansberg (2021): "Learning from Coworkers," *Econometrica*, 89, 647–676. 1, 5, 12, 14, 17

Krueger, A. and C. Rouse (1998): "The Effect of Workplace Education on Earnings, Turnover, and Job Performance," *Journal of Labor Economics*, 16, 61–94.

Lillard, L. A. and H. W. Tan (1986): "Private Sector Training Who Gets It and What Are Its Effects," Tech. rep.

Loewenstein, M. A. and J. R. Spletzer (1999): "General and Specific Training Evidence and Implications," *Journal of Human Resources*, 34, 731–733.

Lynch, L. M. (1992): "Private-Sector Training and the Earnings of Young Workers," *The American Economic Review*, 82, 299–312.

Marcotte, D. E. (2000): "Continuing Education, Job Training, and the Growth of Earnings Inequality," *Industrial and Labor Relations Review*, 53, 602–623.

Nix, E. (2017): "Learning Spillovers in the Firm," *Manuscript.* 1, 5, 11

Parent, D. (1999): "Wages and Mobility: The Impact of Employer-Provided Training," *Journal of Labor Economics*, 17, 298–317.

Veum, J. R. (1995): "Sources of Training and Their Impact on Wages," *Industrial and Labor Relations Review*, 48, 812.

F Additional quantitative results

F.1 Identification of parameters

Table F.1: Jacobian matrix with elasticities of targeted moments to parameters

	Parameter										
Moment	A	A_e	α	σ	ϵ	γ	δ_h	c_m	c_v	ζ	κ
Unemployment rate	-0.08	-0.07	-0.08	0.03	-0.02	0.12	0.11	**-1.95**	0.31	0.45	**1.29**
Labor market tightness (#vacancies/#unemployed)	0.18	0.11	0.15	-0.06	0.04	-0.27	-0.23	**1.38**	**-0.65**	**-0.90**	**-0.96**
Shape parameter of firm employment distribution	-0.01	0.00	-0.04	0.00	0.00	0.00	0.02	0.00	0.00	**0.12**	0.01
Share of workers that remain employed in next quarter	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	**-0.03**
Average wage growth (per quarter), 0–40 yrs of exp	**1.16**	**0.53**	**0.67**	-0.16	0.18	**-1.36**	**-1.91**	-0.05	0.05	-0.38	0.00
Average wage growth (per quarter), 0–25 yrs of exp	**0.81**	0.22	0.38	-0.22	0.04	**-1.25**	**-2.00**	-0.21	-0.25	**-0.73**	-0.39
Share of total time spent on internal learning	**1.70**	-0.28	**0.51**	-0.04	0.07	-0.42	**-1.29**	-0.01	0.00	-0.42	-0.16
Share of total time spent on external learning	**-0.50**	**1.44**	0.24	**-0.59**	-0.06	**-1.07**	**-1.12**	-0.03	0.00	**-0.64**	-0.11
Ratio of learning time in 100+ to 50–99 worker firms	**-0.24**	-0.18	0.07	0.05	-0.04	0.16	0.04	-0.01	0.00	-0.12	-0.02
Ratio of all to new workers' avg external learning time	**0.56**	**-1.12**	-0.45	**3.15**	-0.14	-0.42	0.13	-0.02	0.00	0.41	0.10
Ratio of old to all workers' average learning time	0.40	-0.12	0.11	-0.07	-0.06	-0.14	**0.57**	0.00	0.00	-0.12	-0.02

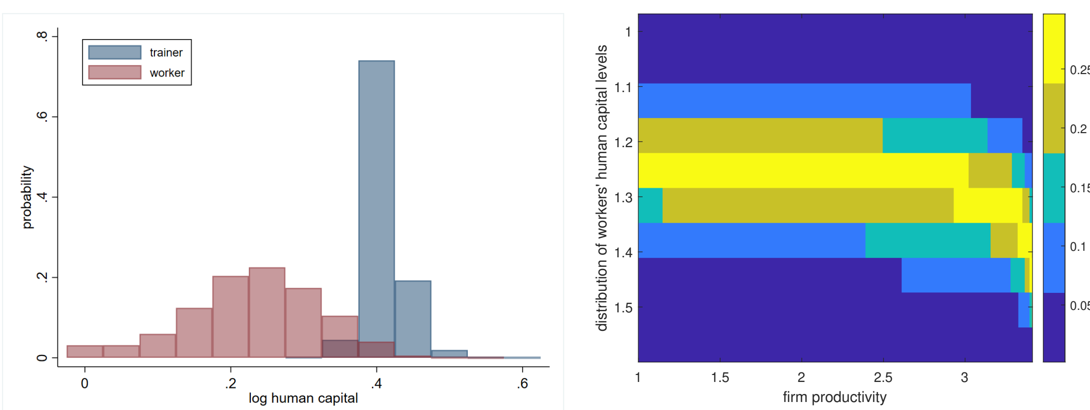
Notes: This table presents the elasticity of each moment targeted in the model calibration to each internally calibrated parameter. These elasticities are determined by calculating the percentage increase in each moment after a 10% change in the calibrated parameter value while holding the remaining 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.

F.2 Properties of equilibrium

Figure F.1: Distribution of human capital of trainers, production workers, and within firms

(a) Distribution of trainers' and workers' human capital

(b) Human capital distribution within firms

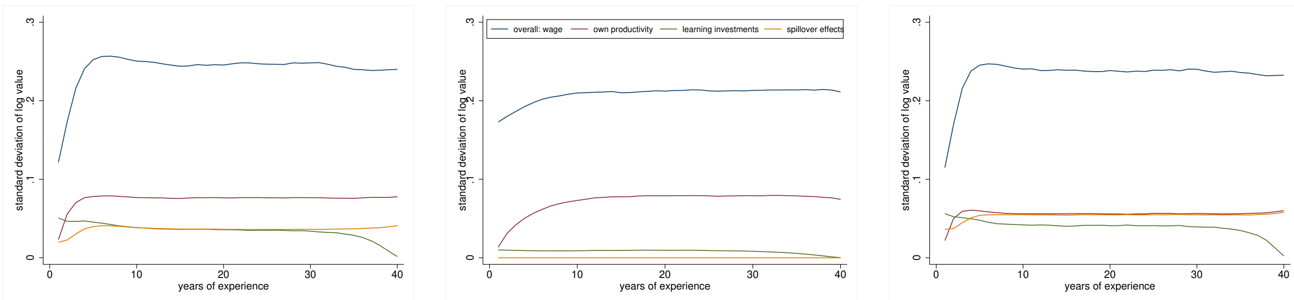


Notes: These figures illustrate the human capital distributions of production workers, external trainers, and firms. Panel (a) plots the share of workers of each human capital level within trainers and production workers. Panel (b) illustrates the human capital distribution within each firm by using 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.

F.3 Role of learning sources in wage dispersion

Figure F.2: 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 components (own productivity, spillovers to coworkers, and learning investments and its 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.

F.4 Training subsidies in the US

Table F.2: Training subsidies in the US

Country	Year	Subsidy or incentive to employer
Alabama	2014 - present	75% of training costs reimbursed
Arizona	2015 - 2020	50-75% of training costs reimbursed
Colorado	2018 - present	60% of training costs reimbursed
Florida	1993 - present	50-75% of training costs reimbursed
Georgia	1994 - present	50% of training costs tax deductible
Hawaii	1991 - present	50% tuition costs reimbursed
Illinois	1992 - present	50% of training costs reimbursed
Kentucky	1984 - present	50% of training costs reimbursed
Maryland	1989 - present	50% of training costs reimbursed
Massachusetts	2008 - present	50% of training costs reimbursed
Mississippi	2013 - present	50% of training costs reimbursed
Montana	2005 - present	Funding of $5,000 for training
Nebraska	2005 - present	Funding of $800-4,000 for training
New Hampshire	2007 - present	50% of training costs reimbursed
New Jersey	1992 - present	50% of training costs reimbursed
New Mexico	1972 - present	50-75% of training costs reimbursed
Pennsylvania	1999 - present	Funding of $600-1,200 per trainee
Rhode Island	2006 - present	50% of training costs reimbursed
Washington	1983 - present	50% of training costs reimbursed
Wisconsin	2012 - present	50% of training costs reimbursed
Wyoming	1997 - present	Funding of $1,000 per trainee

Notes: The government targets formal training, which corresponds roughly (though not perfectly) to external learning, as discussed in Appendix B.2.

G Model extensions and robustness

G.1 Incorporating on-the-job search

We now extend the model to include on-the-job search, following Cahuc et al. (2006). In this extension, employed workers can search for new job opportunities alongside unemployed workers. We introduce $0 \leq \eta < 1$ to represent the search intensity of employed workers relative to unemployed workers, who have a normalized search intensity of 1.

When an employed worker receives an external job offer from a poaching firm with productivity z', there are three possible scenarios determining whether the worker transitions to the new firm and how their revenue fraction r, which governs their wage, is adjusted:

(a) **No move, no change in revenue share**
If the poaching firm cannot offer the worker a value higher than the worker's current value, $V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z)) > M^{a_i}(h_{m_i}, z', \mathbb{H}(z'))$, the worker will stay with its current firm, and the revenue share does not change.

(b) **Move to poaching firm**
If the match with the poaching firm is more valuable than the match with the current firm, $M^{a_i}(h_{m_i}, z', \mathbb{H}(z')) > M^{a_i}(h_{m_i}, z, \mathbb{H}(z))$, the worker will move to the poaching firm and will use the current firm's match value as the outside option when negotiating the new revenue share through Nash Bargaining:

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

(c) **No move, update revenue share using poaching firm**
If the match with the current firm is more valuable than the match with the poaching firm, but the worker's current value is low enough to make the poacher's offer appealing, $M^{a_i}(h_{m_i}, z', \mathbb{H}(z')) < M^{a_i}(h_{m_i}, z, \mathbb{H}(z))$ and $M^{a_i}(h_{m_i}, z', \mathbb{H}(z')) > V^{a_i}(r, h_{m_i}, z, \mathbb{H}(z))$, the current firm negotiates a new surplus division to retain the worker. The worker will use the poaching firm's match value as the outside option when negotiating the new revenue share through Nash Bargaining:

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

With the on-the-job search extension, the match and worker's values now incorporate the benefits from job-to-job transitions. We compute the optimal learning levels by maximizing the match value, and solve for the revenue fraction r that delivers the negotiated worker's value. This fraction r then defines the corresponding wage.

We set the on-the-job search intensity η to 0.3, following Faberman, Mueller, Sahin, and Topa (2017). All other externally calibrated parameters remain as in the baseline. For internally calibrated parameters, we still look to match the targeted moments in Table 4.2, but now target the overall average wage growth per quarter for both 0–40 and 0–25 years of experience, not just wage growth from human capital. With on-the-job search, our model predicts an average quarterly wage growth of 0.46% for workers with 0–40 years of experience, aligning closely with the data (0.45%). All other model moments still align well with the targeted data moments.

A summary of the results in this extension is provided in panel A of Tables G.1–G.3, with two noteworthy changes. First, in the extended model, the overall importance of human capital to wage growth, both through own productivity and non-own productivity embodied in learning costs and spillovers, declines. Nevertheless, with learning costs and spillovers capturing 51% of the wage gains from human capital over 25 years of experience in this scenario, the *relative* importance of these different components of human capital remains very close to the baseline results, where this figure is 48%. Second, average human capital declines by 1.12% under the 20% remote work scenario, compared to a smaller decline of 0.45% in the baseline model. This greater reduction in the extended model is partly due to remote work hampering coworker learning, which significantly affects the incentives for highly productive firms to post job vacancies. These firms employ more skilled workers, and have higher levels of coworker learning than less productive firms. This disincentive is more pronounced in the extended model with on-the-job search because poaching leads to a higher concentration of skilled workers in highly productive firms. Consequently, remote work also causes a more severe decline in wage growth from spillover effects in the extended model compared to the baseline model.

G.2 Incorporating time costs for internal trainers

In the baseline model, we assumed that internal learning incurred no costs for the colleagues acting as internal trainers. This assumption stemmed from the nature of internal learning, which often occurs informally. However, it is possible that mentoring imposes time costs on internal trainers. To incorporate this, we modify the per-period output value in Equation (4) for worker i in firm z to be

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

$\bar{l}_{int}(z, \mathbb{H}(z))$ is the average time workers in firm z spend on internal learning, which also captures the time worker i spends mentoring internal learners under the assumption of random matching. $F_{int}(h_{m_i-1}; \mathbb{H}(z))$ is the proportion of workers (weighted by their internal learning time) in firm z with human capital levels lower than h_{m_i}, representing the share of workers in the firm whom worker i actively mentors. ν governs the time loss experienced by internal trainers due to mentoring.

To calibrate ν, 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.[51] The other targeted moments are the same as in Table 4.2. The calibration results indicate that $\nu = 0.19$, suggesting that internal trainers do incur some time costs, although these are not very significant.

[51] The US-SEPT data provides information on training expenditures related to the wages and salaries of in-house trainers in 1995, which we combine with the wage and salary costs of trainees to calculate the proportion of internal trainers' wage costs in relation to learners' wage costs.

A summary of the results in this extension is provided in panel B of Tables G.1–G.3. All results closely resemble those of the baseline model.

G.3 Incorporating heterogeneity in initial human capital

In the baseline model, we assumed that all workers had the same level of human capital when entering the labor market. Now, we introduce variation in initial human capital across workers to reflect differences in schooling levels. To achieve this, we match the initial distribution of human capital levels to the distribution of schooling years observed in the NHES data. Specifically, we standardize the human capital of individuals with the lowest education level (secondary school) in the NHES data to 1. Then, we determine the human capital level of all other workers by considering their years of schooling relative to this baseline level.[52] We recalibrate all internally calibrated parameters to match the targeted moments presented in Table 4.2.

A summary of the results in this extension is provided in panel C of Tables G.1–G.3. Aligning the initial distribution of human capital levels with the distribution of schooling years raises average human capital levels, and raises the effects of training subsidies on lifetime wage growth as trainers now have higher levels of human capital. All other results remain very close to the baseline results.

G.4 Incorporating unemployment benefits

In the baseline model, unemployment was considered equivalent to employment at the least productive firm following Bagger et al. (2014) and in order to circumvent the complexities associated with varying reservation wages among workers of different types. We now refine our treatment of unemployment by explicitly factoring in unemployment benefits and assuming that the unemployment value is given by:

$$V_U^{a_i}\left(h_{m_i}\right) = bh_{m_i} + \rho\mathbb{E}\max\{V_U^{a_i+1,s}\left(h_{m_i}\right), V_{TR}^{a_i+1}(h_{m_i})\}.$$

b captures unemployment benefits (or losses) per unit of human capital. $V_U^{a_i,s}\left(h_{m_i}\right) = V_U^{a_i}\left(h_{m_i}\right) + \beta\lambda_U \int \frac{v(z)}{V}\max\{M^{a_i}(h_{m_i}, z, \mathbb{H}(z)) - V_U^{a_i}\left(h_{m_i}\right), 0\}d\Phi(z)$ represents the value of staying in the production sector while unemployed before job search begins. This adjustment in the unemployment value may result in workers rejecting certain job offers when the unemployment value exceeds the match surplus.

We calibrate the newly introduced parameter b along with all previously internally calibrated parameters to match the targeted moments presented in Table 4.2. To calibrate b, we introduce a new targeted moment: the ratio of the minimum wage to the mean wage, which is reported as 0.6 by Hornstein, Krusell, and Violante (2011).[53]

[52] We assume a 5% return to schooling, which follows from the returns to advanced degrees in the Penn World Table (Feenstra, Inklaar, and Timmer, 2015), and facilitates computation as a one step increase in the human capital ladder corresponds to a 0.05 increase in human capital in our quantitative model.

[53] To ensure a more accurate comparison, we maintain the same vacancy distribution as in the baseline model,

A summary of the results in this extension is provided in panel D of Tables G.1–G.3. This extension substantially increases the impact of training subsidies on human capital and wage growth. This occurs because when learning is subsidized, younger workers are more likely to accept job offers and start accumulating human capital, even in low-productivity firms. In contrast, older workers, having accumulated sufficient human capital, tend to accept offers only from high-productivity firms offering higher wages. All other results remain very close to the baseline results.

G.5 Incorporating a fixed learning cost to consider extensive margin of learning

In the baseline model, learning differences stemmed solely from the intensive margin, as all workers (except those retiring soon) engaged in learning. To incorporate changes along the extensive margin, we introduce a fixed learning cost. Specifically, in each period, if the expected increase in match surplus from learning is less than f, representing upfront fixed learning costs, both the firm and the worker will opt out of learning activities.

We calibrate the newly introduced parameter f along with all previously internally calibrated parameters to match the targeted moments presented in Table 4.2. To calibrate f, we introduce a new targeted moment: the proportion of workers actively learning with their current employer, which is 0.96 according to the US-SEPT.

A summary of the results in this extension is provided in panel E of Tables G.1–G.3. This extension substantially reduces the impact of training subsidies on human capital and wage growth. This is because the subsidies focus solely on reducing the variable costs of external learning and are insufficient to exceed the fixed costs for many workers. All other results remain very close to the baseline results.

G.6 Alternative parameterizations

We assess the robustness of our results to considering different values for two key parameters driving wage bargaining and human capital formation in the model. First, we change the workers' bargaining power (β) to 0.3 and 0.7 (the baseline value is 0.5), matching the estimates in Bagger et al. (2014) and Gregory (2019), respectively. Second, we increase the substitutability between internal and external learning by setting the elasticity of substitution (σ) to 5 (the baseline value is 3.63). For each scenario, we recalibrate all other internally calibrated parameters to match the targeted moments presented in Table 4.2.

A summary of the results in this extension is provided in panels F–H of Tables G.1–G.3. These results closely resemble those of the baseline model in all cases.

considering that unemployment benefits might cause workers to reject certain job offers, potentially leading to a degenerate vacancy distribution.

G.7 Summary of robustness results

Table G.1: Learning and human capital in the model extensions

	Workers' share of time spent on learning		
	External learning	Internal learning	Avg Human Capital
	Panel A: Incorporating on-the-job search		
Calibrated economy	1.39%	3.37%	1.37
W/o internal learning	1.44%	0	1.17
W/o external learning	0	3.98%	1.23
	Panel B: Incorporating time costs for internal trainers		
Calibrated economy	1.41%	3.32%	1.28
W/o internal learning	1.62%	0	1.13
W/o external learning	0	3.49%	1.18
	Panel C: Incorporating heterogeneity in initial human capital		
Calibrated economy	1.47%	3.21%	1.44
W/o internal learning	1.66%	0	1.24
W/o external learning	0	3.28%	1.29
	Panel D: Incorporating unemployment benefits		
Calibrated economy	1.35%	3.34%	1.32
W/o internal learning	1.48%	0	1.16
W/o external learning	0	4.46%	1.23
	Panel E: Incorporating a fixed learning cost to consider extensive margin of learning		
Calibrated economy	1.40%	3.24%	1.28
W/o internal learning	1.75%	0	1.15
W/o external learning	0	3.72%	1.19
	Panel F: Setting workers' bargaining power $\beta = 0.3$		
Calibrated economy	1.36%	3.24%	1.25
W/o internal learning	1.50%	0	1.12
W/o external learning	0	3.54%	1.16
	Panel G: Setting workers' bargaining power $\beta = 0.7$		
Calibrated economy	1.47%	3.26%	1.29
W/o internal learning	1.69%	0	1.15
W/o external learning	0	3.63%	1.19
	Panel H: Setting elasticity of substitution between learning modes $\sigma = 5$		
Calibrated economy	1.15%	4.10%	1.29
W/o internal learning	1.65%	0	1.14
W/o external learning	0	4.67%	1.22

Notes: This table presents the share of time workers spend on internal and external learning, along with the average human capital level, in the calibrated economy of alternative model setups, and in corresponding counterfactual scenarios that shut down internal and external learning, respectively.

Table G.2: Decomposition of wage growth in the model extensions

	Decomposition of wage growth			
Δlog wage, 0–25 yrs of exp	own productivity	learning costs	spillovers	others
Panel A: Incorporating on-the-job search				
0.974	34.20%	15.90%	19.46%	30.44%
Panel B: Incorporating time costs for internal trainers				
0.505	49.09%	26.41%	21.25%	3.25%
Panel C: Incorporating heterogeneity in initial human capital				
0.453	50.18%	21.10%	22.88%	5.84%
Panel D: Incorporating unemployment benefits				
0.569	49.84%	28.83%	21.80%	-0.47%
Panel E: Incorporating a fixed learning cost to consider extensive margin of learning				
0.515	48.41%	26.52%	22.12%	2.94%
Panel F: Setting workers' bargaining power $\beta = 0.3$				
0.474	47.27%	27.14%	21.64%	3.95%
Panel G: Setting workers' bargaining power $\beta = 0.7$				
0.514	49.89%	25.36%	21.45%	3.30%
Panel H: Setting elasticity of substitution between learning modes $\sigma = 5$				
0.567	43.72%	29.60%	22.90%	3.78%

Notes: This table presents the decomposition of wage growth at 25 years of experience in alternative model setups.

Table G.3: Impact of shocks on human capital and wages in the model extensions

	Benchmark (no shocks))	Impact of shocks (rel. to baseline)		
		Shift to remote work		50% subsidy for external learning
		20% remote	100% remote	
Panel A: Incorporating on-the-job search				
Average human capital	1.37	-1.12%	-4.04%	6.71%
On-the-job human capital gains	0.37	-4.13%	-14.87%	24.67%
Δ log wage, 0–5 yrs of exp	0.579	-6.03%	-18.24%	27.55%
Δ log wage, 5–25 yrs of exp	0.394	-3.86%	-11.95%	47.14%
Panel B: Incorporating time costs for internal trainers				
Average human capital	1.28	-0.44%	-3.46%	6.16%
On-the-job human capital gains	0.28	-2.02%	-16.02%	28.51%
Δ log wage, 0–5 yrs of exp	0.328	-5.23%	-27.52%	28.44%
Δ log wage, 5–25 yrs of exp	0.177	-1.39%	-8.05%	59.09%
Panel C: Incorporating heterogeneity in initial human capital				
Average human capital	1.44	-0.63%	-3.65%	8.16%
On-the-job human capital gains	0.28	-3.29%	-19.01%	42.48%
Δ log wage, 0–5 yrs of exp	0.259	-6.06%	-32.13%	37.47%
Δ log wage, 5–25 yrs of exp	0.193	1.42%	-12.65%	72.97%
Panel D: Incorporating unemployment benefits				
Average human capital	1.32	-0.45%	-3.42%	13.24%
On-the-job human capital gains	0.32	-1.89%	-14.27%	55.23%
Δ log wage, 0–5 yrs of exp	0.349	-3.35%	-22.47%	66.09%
Δ log wage, 5–25 yrs of exp	0.221	-2.75%	-15.16%	127.64%
Panel E: Incorporating a fixed learning cost to consider extensive margin of learning				
Average human capital	1.28	-0.37%	-2.60%	3.04%
On-the-job human capital gains	0.28	-1.71%	-11.83%	13.87%
Δ log wage, 0–5 yrs of exp	0.350	-6.50%	-22.36%	17.18%
Δ log wage, 5–25 yrs of exp	0.165	6.76%	-5.12%	32.26%
Panel F: Setting workers' bargaining power $\beta = 0.3$				
Average human capital	1.25	-0.45%	-2.28%	5.91%
On-the-job human capital gains	0.25	-2.26%	-11.52%	29.91%
Δ log wage, 0–5 yrs of exp	0.302	-4.17%	-23.93%	33.60%
Δ log wage, 5–25 yrs of exp	0.172	-3.99%	-3.46%	64.07%
Panel G: Setting workers' bargaining power $\beta = 0.7$				
Average human capital	1.29	-0.55%	-2.57%	5.83%
On-the-job human capital gains	0.29	-2.46%	-11.43%	25.90%
Δ log wage, 0–5 yrs of exp	0.341	-4.83%	-20.67%	29.47%
Δ log wage, 5–25 yrs of exp	0.172	-2.50%	-9.58%	47.72%
Panel H: Setting elasticity of substitution between learning modes $\sigma = 5$				
Average human capital	1.29	-0.70%	-3.00%	5.53%
On-the-job human capital gains	0.29	-3.11%	-13.43%	24.73%
Δ log wage, 0–5 yrs of exp	0.408	-6.18%	-24.39%	29.05%
Δ log wage, 5–25 yrs of exp	0.159	3.36%	-3.70%	59.29%

Notes: This table presents the following in the benchmark economy (with no shocks), the scenarios with 20% and 100% remote work, and the scenario with 50% subsidy to external learning: the average human capital level, on-the-job human capital gains (built by subtracting the initial human capital from the average human capital in the baseline scenario), and the average growth at 0–5 years of experience, and 5–25 years of experience. These results refer to the baseline model with both internal and external learning.

H Evidence supporting testable predictions of the model

H.1 Evidence supporting mechanisms driving learning patterns

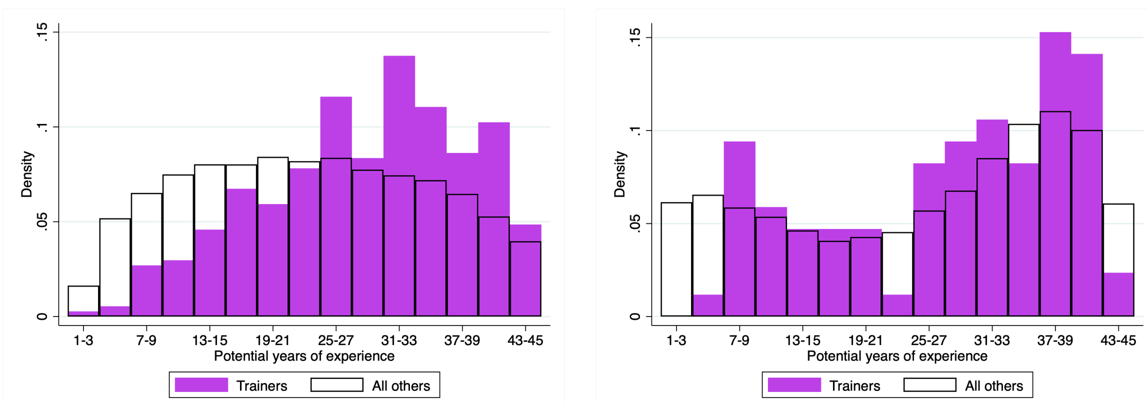
H.1.1 Human capital distribution of trainers is left-skewed

Our theory suggests that trainers possess higher human capital than production workers, which motivates mid-career workers to transition into external learning after exhausting learning opportunities within their firm. Figure H.1 supports this by plotting the histograms of potential experience for trainers and production workers in Germany and the US. Trainers are defined as workers engaged in occupations involving training, teaching, or instructional activities outside of school and university education. Production workers include all other workers except trainers.

The plots reveal that trainers in both Germany and the US concentrate at higher levels of potential experience compared to other workers.[54]

Figure H.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.

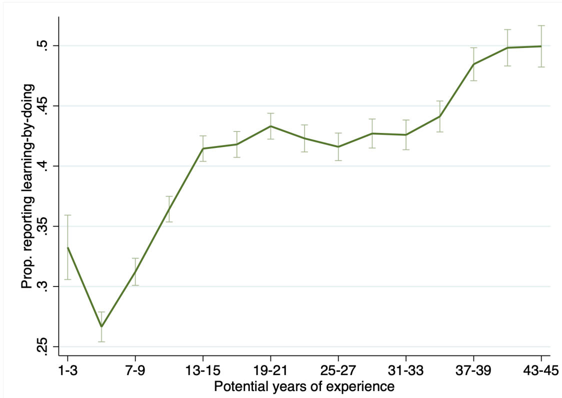
H.1.2 Proportion of workers who learn-by-doing rises with human capital

Another key prediction of our model is that the portion of workers who do not invest in internal or external learning and instead rely on learning-by-doing increases with human capital. We provide evidence for this using the German BIBB/BAuA data. We construct a measure of learning-by-doing capturing individuals who acquired the skills/knowledge

[54] Additional results show that these findings are consistent when comparing trainers to workers in professional and technical occupations, as well as to internal and external learners as defined in Section 3. Further, these results are robust to formally testing these distributional differences through quantile regressions of potential years of experience on the trainer variable (with production workers as the omitted category) even after including a battery of controls. These results are available upon request.

necessary to complete the tasks in their current job through the job itself. In Figure H.2, we find that learning-by-doing increases with potential experience, consistent with our model's prediction.[55]

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

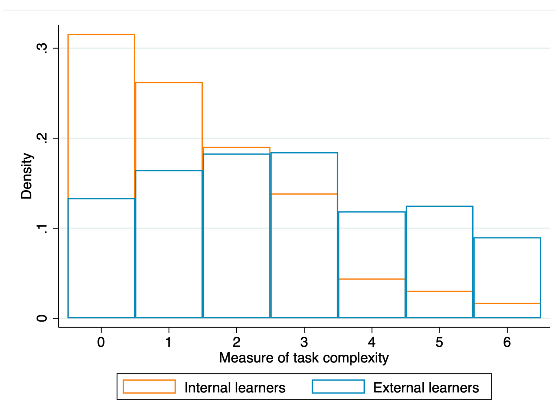


Notes: This figure plots the proportion of workers reporting engaging in learning-by-doing across different potential experience bins in the German BIBB/BAuA data. The sample encompasses individuals who are currently employed and have between 1 and 45 years of potential experience. The results are weighted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

H.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 therefore more easily trained by coworkers. To further test these human capital differences between internal and external learners, we examine how the skill content of the tasks performed varies between these two groups of workers.

Figure H.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.

[55] This positive correlation between experience and learning-by-doing remains statistically significant even after controlling for various demographic and firm-level variables. These results are available upon request.

Using data from Germany, we construct a measure of task complexity by counting the number of specific skills workers report using in their jobs. These skills include: math and stats; foreign language; computing; accounting, purchasing, financing, and taxes; marketing; and management and organization. Higher values of this measure indicate greater task complexity. Figure H.3 shows that the distribution of task complexity is more heavily concentrated at lower levels for individuals learning internally compared to those learning externally.[56]

H.2 Evidence supporting importance of coworker instruction

Our theory and quantitative findings indicate that wage compensation resulting from coworker spillovers, particularly those generated by high-skill workers who can effectively teach their colleagues, constitutes a significant component of lifecycle wage growth. We provide evidence supporting this by documenting that tutoring coworkers is a common task across various occupations, and it is often performed by high-skill workers.

To do this, we use data from the US Department of Labor's O*NET project, which characterizes the tasks associated with each occupation, along with the required mix of knowledge, skills, and abilities needed to perform them. This dataset encompasses information on 900 occupations, with analysts at O*NET assigning scores characterizing the importance of various work activities, skills, and other descriptors in each of them.

H.2.1 Tutoring tasks are important

First, we show that coworker tutoring is an important aspect of many non-teaching occupations.[57] We focus on three work activities and one skill that capture coworker tutoring tasks: (1) guiding, directing, and motivating subordinates activity; (2) coaching and developing others activity; (3) training and teaching others activity; and (4) instructing skill.

In Columns 2–4 of Table H.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 their average score.[58] We find that a significant proportion of non-teaching occupations assigns high importance to coworker tutoring tasks and skills. For instance, 57 occupations, representing 6.7% of all non-teaching occupations, attribute a medium or higher score to the importance of activities related to guiding, directing, and motivating subordinates. A

[56] These results are robust to formally testing these distributional differences through quantile regressions of task complexity on the external learning variable (with internal learning as the omitted category) even after including a battery of controls. These results are available upon request.

[57] In order to focus on non-teaching occupations, we exclude occupations linked to teaching in the education sector, as well as positions related to education administration. In addition, we drop occupations that solely focus on training within firms, like training and development managers and specialists, as well as those engaged in broader community-wide education, such as health educators or sports coaches.

[58] The scores for each activity and skill range from 0 to 7. Therefore, medium or higher scores correspond to values of 3.5 or higher.

breakdown of these occupations reveals that most of them are managerial and supervisory roles.[59] This aligns with existing literature suggesting that managerial inputs are crucial for on-the-job human capital development (Burstein and Monge-Naranjo, 2009; Luttmer, 2014).

Table H.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 2 and 3 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 4 reports the average score for these occupations in each of the considered activities or skills. Column 5 reports the average job zone score for these occupations.

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

We then examine the skill requirements of these tutoring-intensive occupations. Specifically, we utilize the "job zone" information provided in the O*NET data, which categorizes occupations into five bins based on the education, experience, and training required. A higher bin signifies a greater level of skill necessary for the occupation.

In Column 5 of Table H.1, we present the average job zone for occupations with a medium or higher score in each of the activities and skill sets of interest. We find that tutoring-intensive occupations typically demand high skill levels. For instance, among the 57 occupations reporting a medium or higher score in the importance of activities related to guiding, directing, and motivating subordinates, the average job zone score is 3.97 (the maximum is 5). This indicates that considerable education and experience are needed for these roles.

[59] This breakdown, along with the lists of the occupations reporting a medium or higher score in the importance of the other tutoring activities and skill, is available upon request.