The Impact of AI on Career Ladders

Eleanor Dillon* Dana Scott*

November 3, 2026

Abstract

Entry-level jobs do more than supply current labor: they produce experienced workers through learning and advancement. Because workers can take the skills they acquire elsewhere, training firms capture only part of the value they create and may provide too few junior jobs. We study whether generative AI deepens this distortion when it can substitute for junior work and complement senior labor. Using data from Lightcast career histories, we show that firms representing one-fifth of employment create 49.7 percent of junior starts, yet 81.2 percent of advancing entrants at these firms reach senior jobs elsewhere. We link these data to Microsoft 365 Copilot sales and exploit quasi-random variation in contract-renewal timing to identify effects on hiring composition and whether firms with stronger internal pipelines preserve more junior employment. An equilibrium model aggregates these responses, traces their effects on the future supply and allocation of experienced workers, and compares decentralized human-capital production with the social optimum.

*Microsoft Research. Email: eleanor.dillon@microsoft.com; danascott@microsoft.com.

1. Introduction

Entry-level jobs do more than supply firms with labor today. They also produce human-capital: through learning-by-doing, less-experienced employees acquire the skills needed for more advanced jobs (Arrow 1962; Becker 1962). Although any one firm can hire experienced workers trained elsewhere, the economy as a whole can produce them only if some firms create opportunities for workers to enter, learn, and advance. Because workers can carry the skills they acquire to other employers, firms may capture only part of the value created by these jobs (Acemoglu and Pischke 1998, 1999). The decentralized economy may therefore create too few junior jobs and produce too little human capital.

Generative AI may deepen this inefficiency. AI can automate many codifiable cognitive tasks traditionally assigned to junior workers (Eloundou et al. 2023), and early evidence points to slower hiring among early-career workers in AI-exposed jobs (Brynjolfsson et al. 2025a; Lichtinger and Hosseini Maasoum 2025). At the same time, work requiring context, judgment, and responsibility remains difficult to automate (Dell'Acqua et al. 2026). Substituting AI for junior labor can be privately efficient even as it reduces the pipeline that produces future experienced workers. The key question is whether firms cut junior hiring by more than is socially efficient because they do not internalize the value of trained workers to other employers.

In this paper we ask whether AI widens the gap between private and social incentives to create junior jobs, and what this implies for the future supply of experienced workers. We combine large-scale data on worker career histories with Microsoft enterprise sales data on the staggered introduction of Microsoft 365 Copilot. We first document a division of labor in human-capital production across firms: junior-intensive firms create internal promotion opportunities and become net exporters of experienced workers, while senior-intensive firms hire workers trained elsewhere. We then exploit staggered access to Copilot to estimate how AI changes junior hiring and whether firms with stronger internal pipelines preserve more of these jobs. Finally, we develop an equilibrium model that translates these firm-level responses into changes in the supply and allocation of experienced workers and compares the decentralized response with the social optimum.

Using rich data on worker career histories from Lightcast, we document three facts showing that human-capital production is concentrated and connected across firms. First, junior jobs are not spread evenly across firms: the most junior-intensive firms, which together account for one-fifth of employment, create 49.7 percent of all junior starts. Second, the same firms generate a disproportionate share of advancement into senior work but retain little of the resulting talent. They account for 42.4 percent of junior entrants who reach a senior job within five years, but 81.2 percent of those workers advance at another employer. The most junior-intensive quintile accounts for 49.5 percent of first-time senior exports but only 25.9 percent of first-time senior imports. The market for

experienced labor therefore connects firms that create entry opportunities to firms that ultimately employ the workers those opportunities produce.

Motivated by these facts, we develop a simple two-period framework of firms' hiring choices. We first characterize a baseline economy without AI. In the first period, firms choose how many junior workers to hire and train and how many experienced workers to recruit from the external market. Junior workers contribute to current production while acquiring skills. In the second period, the firm observes the quality of the workers it trained, who may remain with the firm as seniors or move to another employer. Because training reveals worker quality and develops firm-specific human capital, firms differ in how much they value producing their own senior workers rather than hiring them externally. We then introduce the option to adopt AI. AI substitutes for labor in junior tasks while complementing skilled labor in senior tasks, reducing the immediate value of junior hiring while increasing the value of experienced workers. The framework predicts that firms will shift hiring toward seniors, but that firms with more valuable internal pipelines will preserve more junior jobs. Whether total hiring rises or falls is ambiguous because greater senior productivity may offset the decline in junior labor demand. These predictions motivate our quasi-experimental tests of changes in junior and senior hiring, heterogeneity by firms' pre-AI pipelines, and subsequent changes in internal promotion and external senior hiring.

We next bring these predictions to data using the rollout of Microsoft 365 Copilot, a generative AI assistant embedded in workplace applications such as Word, Excel, Outlook, and Teams. Microsoft broadened Copilot's availability to enterprise customers in early 2024, but firms' access remained closely tied to the renewal dates of their existing Microsoft contracts. Because these renewal schedules were determined before broad availability, their timing provides plausibly exogenous variation in when firms gained access to Copilot. Linking Microsoft sales records to worker career histories, we use contract renewal timing as an instrument for Copilot access. The analysis tests whether access shifts the composition of new hiring away from junior workers and toward senior workers, whether it changes total hiring, and whether junior hiring declines less at firms with more valuable pre-existing internal pipelines. We then examine subsequent changes in internal promotion and external senior hiring. We expect to complete the first-stage validation, main hiring estimates, heterogeneity, and downstream worker-flow analyses in the coming weeks.

Related literature. This paper contributes to three distinct strands of economic literature. First, we contribute to the literature on on-the-job training and internal labor markets, beginning with the canonical learning-by-doing mechanism of Arrow (1962) and the distinction between general and specific human capital of Becker (1962). Our model builds on the framework of Acemoglu and Pischke (1998, 1999), who establish that firms will provide training to entry-level workers if they can capture part of the return to general training through information frictions. We connect

training decisions to the logic of internal promotions and career ladders (Doeringer and Piore 1971; Gibbons and Waldman 1999; Baker et al. 1994; Huitfeldt et al. 2023; Pastorino 2024). We synthesize these approaches by developing a model in which firms create junior-level jobs that both produce current output and generate future senior-level human capital. Early evidence suggests that AI technology has a greater effect on less-experienced workers and entry-level hiring (Brynjolfsson et al. 2025a,b; Lichtinger and Hosseini Maasoum 2025), motivating the distinct impacts at each career stage in our model. Our contribution is to show, theoretically and empirically, that the effect of AI on human-capital production depends on how much of the future return to junior employment firms can capture, linking firms' internal career structures to the external supply of experienced workers.

Second, we build on the task model of jobs developed by Autor et al. (2003) and Acemoglu and Autor (2011), the task-based theory of automation of Acemoglu and Restrepo (2022), and the link between tasks and job hierarchies of Garicano (2000) and Autor and Thompson (2025). Recent work has focused on the unique relationship of AI to task hierarchies: it can substitute for human labor in simple tasks performed by junior-level workers (Hui et al. 2024; Cui et al. 2025; Crane and Soto 2026) but can complement expert workers in complex tasks (Ide and Talamàs 2025; Agarwal et al. 2025b,a; Gans and Goldfarb 2026; Demirer et al. 2026). We advance this work by adding a dynamic human-capital margin to the task framework: automating junior tasks changes not only current labor demand but also the future supply of experienced workers, and our empirical design separates this substitution channel from changes in firm scale and senior-worker sourcing.

Third, our work connects to four contemporaneous papers that examine the theoretical relationship between AI and human capital development (Afrouzi et al. 2026; Friebel et al. 2026; Garicano and Rayo 2026; Ide 2025). Each of these papers asks how the talent pipeline might unravel if AI can replace human labor in entry-level tasks through which novice workers previously developed human capital. Afrouzi et al. (2026) develop a model of learning by doing with an endogenous task frontier. They show that high- and low-learning equilibria can coexist, so AI can either raise or reduce welfare. We instead ask which firms reduce human-capital production and show that their pre-AI training incentives predict substitution away from junior labor. Garicano and Rayo (2026) model training as a bilateral relationship between a master and a novice. They show that training survives when AI raises expert productivity sufficiently relative to novice productivity. We show that these private incentives may nevertheless generate too few junior jobs when training firms cannot capture the full value of the human capital they produce. Ide (2025) studies how novices acquire tacit knowledge by apprenticing with experts. He shows that liquidity constraints and a floor on novice compensation can lead to excessive automation. We derive similar underprovision from imperfect appropriability and use firm heterogeneity to predict where it occurs. Friebel et al. (2026) examine how AI changes the junior-senior hierarchy within a representative firm. By introducing heterogeneous firms and an external senior labor market, we distinguish changes in workforce composition from changes in firm

scale and internal-versus-external senior sourcing.

Our work complements these papers. They establish that AI can alter the viability of learning and training; we derive firm-level predictions about *which* firms reduce junior-job creation and bring those predictions to data using Copilot access linked to worker career histories. This allows us to distinguish substitution away from junior labor from changes in firm scale and from changes in the internal-versus-external sourcing of senior workers. Moreover, the implications of these firm-level predictions for the provision of junior-level job opportunities across the economy depends on the *ex ante* distribution of firm incentives. By bringing the theory to data, we learn about this distribution to derive policy-relevant aggregate implications.

Outline. Section 2 describes the data sources and selection of our main analysis sample. Section 3 establishes descriptive facts on the distribution of junior-level job opportunities across the economy. Section 4 develops a conceptual framework for how different firms will respond to the availability of AI and outlines empirical predictions. Section 5 uses quasi-experimental variation from the Copilot rollout to test these predictions. Section 6 concludes and describes next steps for this paper.

2. Data

2.1. Lightcast Career Profiles

Our primary data come from Lightcast Career Profiles, formerly marketed as Burning Glass Technologies resume data. Career Profiles are a proprietary collection of pseudonymized individual career histories assembled from publicly available professional profiles, third-party resume databases and job boards, recruiting-industry sources, and other commercial data providers. Each profile can contain a sequence of reported employment spells, including employer, job title, standardized occupation, and start and end dates, together with information on education, skills, and place of residence. Lightcast uses automated matching and classification procedures to deduplicate records that appear to describe the same individual and to standardize free-text employer names, job titles, educational institutions, and locations.

2.2. Classification of junior and senior jobs

We classify jobs, defined at the level of firm × standardized title, into junior and senior groups using the observable characteristics of workers entering each job. For every job, we aggregate worker-level signals based on prior labor-market experience, proximity to undergraduate completion, and job titles. We then fit a finite-mixture model by expectation maximization. The model estimates the prevalence of each signal within each latent group and combines the signals to calculate posterior

probabilities that a job belongs to each group. This approach allows the data to determine which combinations of worker characteristics distinguish entry, middle, and senior jobs, rather than relying on predetermined job labels. Because it does not require observing promotions or other within-firm mobility, it can be applied broadly, including to jobs at smaller firms. For each job, the model aggregates eight binomial outcomes, summarized in the table below.

TABLE 1. Signals Used in the Job-Level Classification

Signal	Definition	Direction
Low experience	≤ 2 years of estimated experience at job start	Junior
Recent degree entry	Starts within 1 year of most recent completed degree	Junior
Entry title	Entry, junior, or intern title	Junior
Lightcast seniority flag	Entry, training, unpaid	Junior
High experience	≥ 5 years of estimated experience at job start	Senior
Senior title	Senior-IC, manager, director, or leadership title	Senior
Lightcast seniority flag	Senior, manager, director, vice president, executive, owner, partner	Senior
Lightcast band flag	Manager, director, vice president, executive	Senior

The resulting groups exhibit clear orderings in worker characteristics, including on measures that are not directly used in the classification. Table 2 summarizes the composition of the two job types on measures related to observed and potential experience, tenure, and subsequent retention mobility. Median observed experience rises from 0.58 years in junior-level jobs to 5.8 years in senior-level jobs; median potential experience similarly rises from 1.08 to 6.33 years among workers for whom we observe a dated undergraduate degree. Higher-level jobs are also associated with greater tenure and attachment than entry-level jobs. Workers are more likely to enter senior-level jobs from another position within the same firm: the internal-entry share is 5.9% for junior jobs and 16.6% for senior jobs. Additionally, the median worker in an entry-level job stays for 1.17 years, compared to 2.09 years for workers in senior-level jobs. Together, these patterns indicate that the classification captures economically meaningful differences in job seniority and position within firms' career structures.

2.3. Microsoft 365 Copilot sales data

We study firm-level responses to the availability of general-purpose AI using the rollout of Copilot for Microsoft 365 (hereafter, "Copilot"), Microsoft's enterprise generative-AI assistant. Copilot combines large language models with organizational information accessible through Microsoft Graph and embeds the resulting capabilities in Word, Excel, PowerPoint, Outlook, Teams, and a cross-application chat interface. Users can employ natural-language prompts to draft and summarize text, analyze data, create presentations, synthesize email and meeting content, and retrieve information

TABLE 2. Worker Profiles by Junior/Senior Job Level

	Junior	Senior
Panel A: Sample and classification confidence		
Distinct workers	38,272,758	75,876,410
Weighted median workers per job	26.0	9.0
Panel B: Prior labor-market experience		
Prior experience (years)		
25th percentile	0.00	2.09
50th percentile	0.58	5.80
75th percentile	2.33	11.56
Prior experience (%)		
≤ 1 year	58.1	17.5
≥ 2 years	28.0	75.9
≥ 5 years	10.6	54.8
≥ 10 years	3.4	30.3
First observed job (%)	43.7	24.8
Panel C: Education-based potential experience		
Observed degree date (%)	27.7	38.1
Within two years of latest degree (%)	65.0	19.3
Years since last degree		
25th percentile	0.50	2.67
50th percentile	1.08	6.33
75th percentile	3.00	13.00
Panel D: Firm entry and tenure		
Hired from within firm (%)	5.9	16.6
Median job tenure (years)	1.17	2.09
Median subsequent firm tenure (years)	1.34	2.44
Panel E: Retention and subsequent mobility		
Firm retention (%)		
1+ years	60.8	81.1
3+ years	32.7	50.6
5+ years	22.7	36.4
Job retention (%)		
1+ years	58.0	78.6
3+ years	28.6	44.0
5+ years	19.5	29.7
Subsequent internal job move (%)	9.8	13.0
Median years to internal job move	1.08	1.83

across workplace documents and communications. The sales data identify when a firm first acquired Copilot licenses and therefore gained the ability to assign product access to its employees, rather than employees' realized usage of the tool.

Microsoft initially introduced the product to a small group of enterprise customers in 2023 and made it generally available to large enterprises in November 2023, initially with a 300-seat purchase minimum. In January 2024, Microsoft removed the previous minimum and expanded commercial availability. We consider the January 2024 event to be the main broad commercial rollout to firms of all sizes, but individual firms purchased Copilot licenses on a rolling basis after this date. We discuss firms' adoption timing in Section 5.

2.4. Sample selection

We construct the analysis sample from workers entering high-confidence junior jobs. We first identify each worker's first substantive job beginning between 2014 and 2017 and within one year after completion of a dated bachelor's degree. We exclude internships, require the job to last at least 90 days, and restrict entrants to workers with no more than two years of estimated prior labor-market experience. The origin job must have a junior probability of at least 0.80. These restrictions identify 653,084 junior entrants at 173,240 origin firms.

To attribute subsequent advancement to the origin job, we require the worker to remain in that job for at least 180 days and exclude senior transitions occurring before that exposure threshold. We also retain the common cohort observed in the three- and five-year constructions. This leaves 553,671 entrants at 152,577 firms. We follow each entrant for five years and identify the first qualifying senior job: a substantive non-internship job lasting at least 90 days, with a senior probability above 0.80 and a sufficiently large increase in senior probability relative to the origin job. The transition is internal if that job is at the origin firm and external otherwise.

Firm-level advancement rates are noisy for the large number of firms with only one or two observed entrants. We therefore require at least 10 eligible entrants per origin firm. This restriction leaves 6,299 firms and 314,218 entrants. Finally, we require positive observed worker-year exposure during 2014–2017 so that junior-hiring intensity can be defined; this restriction removes no additional firms. The final descriptive sample therefore contains 6,299 firms.

For these firms, junior-hiring intensity is measured using all high-confidence junior firm-episode starts—not only the degree-dated entrant cohort—during 2014–2017. The sample contains 7,754,452 junior starts and 91,009,777 worker-years. We rank firms by junior starts per 100 worker-years and form bins containing approximately equal worker mass. Thus, the degree-dated cohort supplies the advancement outcomes used to select and characterize firms, while the broader worker histories supply the measure of overall junior hiring.

3. Descriptive Facts

This section documents how firms vary along three distinct margins: how many junior opportunities they provide, how often their entrants advance, and whether they are net recipients or suppliers of workers making junior-to-senior job transitions. These measures distinguish the provision of junior opportunities, the progression associated with those opportunities, and the ultimate allocation of senior talent across firms.

To track how workers progress in the early stages of their careers, we focus on a high-confidence subsample of *labor market entrants* identified by a degree-dated entry signal. Under our definition, entrants have a dated bachelor's degree, have no more than two years of work experience prior to graduation, begin their first non-internship job within one year of graduation, and remain attached to this job for at leasts six months. We also require the entry job to have probability of being a junior job of at least 0.8. These sample restrictions are chosen to focus the sample on workers who are entering the labor market for the first time and to avoid confounding entry with incomplete profiles.

First, *junior hiring intensity* is the number of high-confidence junior job starts per 100 worker-years within a firm:

$$\text{Junior hiring intensity} \equiv 100 \times \frac{\text{new high-confidence junior job starts during 2014–2017}}{\text{firm worker-years during 2014–2017}}.$$

This measure captures how intensively the firm supplies junior-level job opportunities relative to its employment scale. It reflects the availability of these opportunities to any junior worker seeking a job, rather than to first-time entrants alone. Second, *pipeline strength* is the share of entrants beginning their career at that firm who reach a senior-level role within five years:

$$\text{Pipeline strength} \equiv \frac{\text{entrants reaching a senior-level role within five years}}{\text{eligible entrants at the origin firm}}.$$

We include both internal and external advancement events in this measure. Third, *talent trade balance* captures whether firms are net buyers or net sellers of trained workers:

$$\text{Talent trade balance} \equiv 100 \times \frac{\text{first-time senior arrivals} - \text{first-time senior departures}}{\text{firm worker-years}}.$$

A first-time senior arrival occurs when a worker begins their career elswhere and first reaches a senior role at the focal firm. A first-time senior departure occurs when an entrant first reaches a senior role at another firm. The measure thus captures whether a firm is a net provider or net recipient of trained workers who are newly eligible for senior-level work.

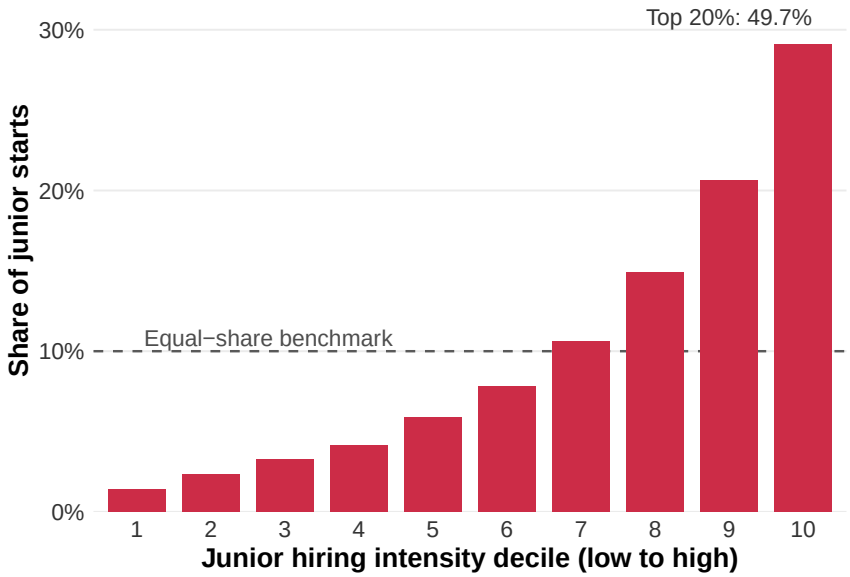
Throughout the analysis, we pool 2014–2017 entry cohorts within firms and restrict attention to

firms with at least 10 eligible entrants. Junior jobs are those with a junior probability of at least 0.8, while senior roles have a senior probability above 0.5. Junior hiring intensity uses all high-confidence junior starts at the firms in the analysis sample, rather than only the degree-dated entrant cohort. We weight cohort outcomes by the number of eligible entrants and firm-wide hiring and trade-flow comparisons by worker-years.

3.1. Junior job creation is concentrated

Firms differ in the rate at which they create new junior employment relationships, measured as new junior firm-spell matches per 100 worker-years. Figure 1 documents the concentration of junior jobs. We rank firms by junior hiring intensity, construct ten bins containing equal shares of worker mass, and plot each decile's share of all junior starts. The top 20 percent of most junior-intensive firms by worker mass account for approximately 49.7 percent of junior job starts. The top decile accounts for 29.1 percent of all junior starts, while the bottom decile accounts for just 1.4 percent. Because the junior hiring intensity measure normalizes by firm employment, this does not simply reflect differences in firm size. The figure highlights that creation of junior job opportunities is not uniform across firms.

FIGURE 1. Concentration of junior job creation

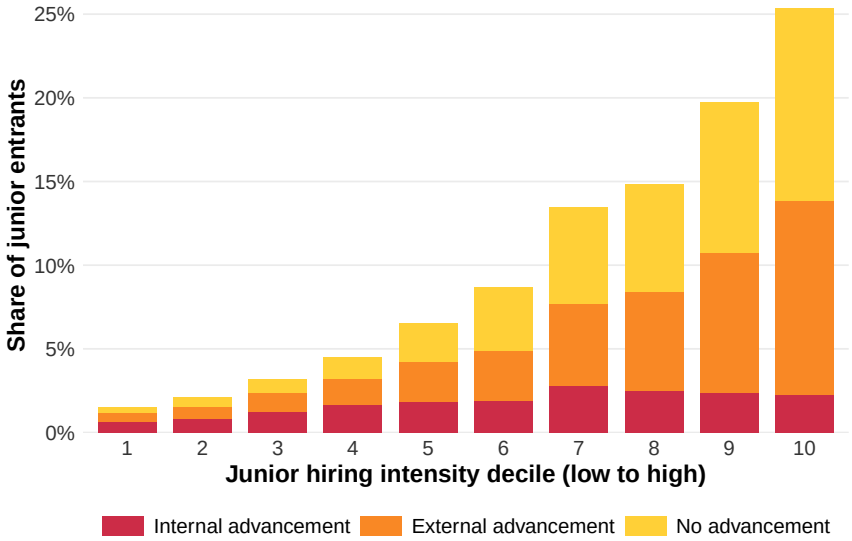


Notes: Junior hiring intensity is defined as 100 times the number of new high-confidence junior firm-spell starts during 2014–2017 divided by firm worker-years over the same period. The sample considers only high-confidence junior jobs satisfying $P(\text{junior}) \geq 0.8$. Firms are ranked by junior hiring intensity and divided into deciles containing approximately equal worker-year mass. Bars report each decile's share of all junior starts. The sample is restricted to firms with at least 10 eligible entrants in the five-year advancement sample.

3.2. Junior hiring and internal advancement are distinct margins

Firms can contribute to workers' advancement either by promoting them internally or by providing junior opportunities from which they advance elsewhere. Figure 2 documents that upward mobility frequently crosses firm boundaries. We rank firms again by junior hiring intensity, and plot the share of labor market entrants in each decile that advance to a senior job internally, externally, or not at all within five years.

FIGURE 2. Advancement pathways by junior hiring intensity



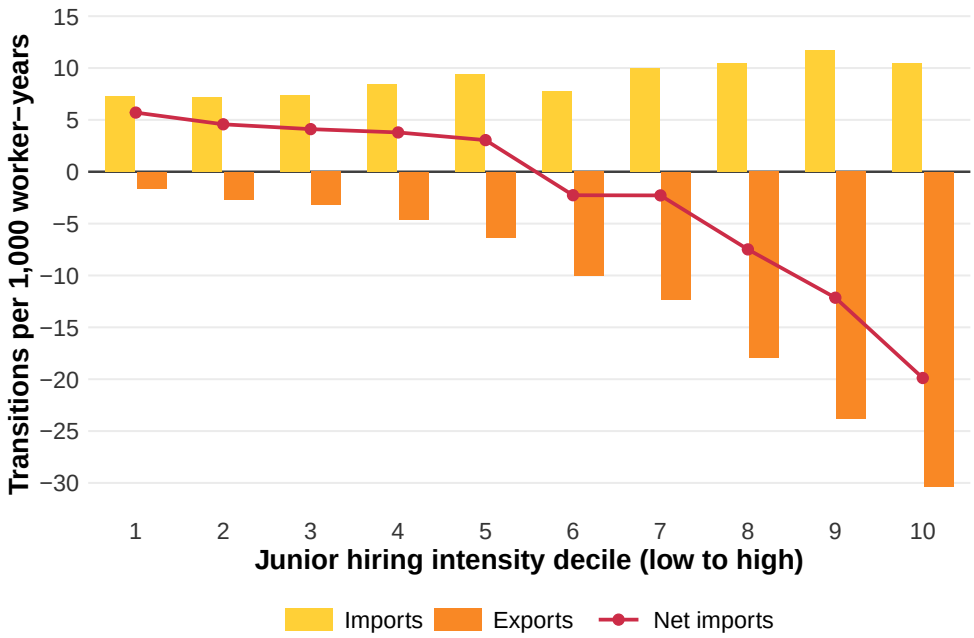
Notes: Firms are ranked by junior hiring intensity and divided into deciles containing approximately equal worker-year mass. Each bar reports the decile's share of all eligible degree-dated entrants who begin in a junior job, decomposed by their first qualifying senior job observed within five years: at the entrant's initial firm (internal advancement), at another firm (external advancement), or no qualifying senior job observed. Junior and senior jobs require an estimated classification probability of at least 0.80. Eligible entrants begin the job within one year of completing a bachelor's degree, have at most two years of prior experience, and remain in the initial job for at least 90 days. A qualifying senior job must begin at least 180 days after entry and last at least 90 days.

Firms in the top two junior-hiring-intensity deciles generate 42.4 percent of all junior jobs that lead to senior advancement within five years. Yet these firms retain relatively little of the resulting senior talent: 81.2 percent of their upwardly mobile juniors reach their first senior job at another employer, while only 18.8 percent advance internally. Internal advancement remains an important pathway overall—31.1 percent of advancing junior entrants reach their first senior job at the firm where they began—but it is substantially less common among the firms creating the greatest mass of upwardly mobile junior jobs. Thus, much of the economy's upward mobility originates at firms that create entry opportunities but ultimately supply senior talent to other employers.

3.3. Firms that create junior jobs are net exporters of senior talent

The firms that create junior opportunities and the firms that ultimately employ senior workers occupy distinct positions in the market for experienced talent. Figure B.1 asks whether firms with high rates of junior hiring are net buyers or net sellers of workers moving into senior jobs. Using the same junior-hiring-intensity deciles as Facts 1 and 2, we compare imports and exports of first-time senior workers. Imports are workers arriving at a firm for their first senior job after holding a junior job elsewhere, while exports are workers leaving a firm's junior jobs for their first senior job elsewhere. Workers who are internally promoted are excluded from both measures. The trade balance is imports minus exports.

FIGURE 3. Trained-worker trade by junior hiring intensity



Notes: Firms are ranked by junior hiring intensity and divided into deciles containing approximately equal worker-year mass. Imports are workers whose first observed senior job during 2014–2022 is at a firm in the indicated decile and whose most recent junior job was at another firm. Exports are workers whose recent junior job was at a firm in the indicated decile and whose first observed senior job is at another firm. A junior predecessor must end no more than one year before the senior job begins. Net imports equal imports minus exports; exports are plotted below zero for presentation. Flows are reported per 1,000 firm worker-years during 2014–2022. Junior and senior jobs require an estimated classification probability of at least 0.80.

The figure shows that firms that create the most junior opportunities are also strong net suppliers of senior talent. The least junior-intensive decile records 5.7 net imports per 1,000 worker-years, while the most junior-intensive records 19.9 net exports. This pattern is driven by differences in exports, which rise from 1.6 to 30.4 per 1,000 worker-years; imports rise only from 7.3 to 10.5. The top two deciles account for 49.5 percent of exports but only 25.9 percent of imports, showing that

other employers capture much of the senior talent emerging from these firms. As before, because we normalize outcomes per 1,000 worker-years, this does not reflect differences in firm size. This evidence establishes a cross-firm link between junior job creation and senior employment, and motivates the role of human capital externalities in our theoretical framework.

4. Conceptual Framework

Motivated by the descriptive facts in Section 3, we develop a stylized framework that adapts the model of Acemoglu and Pischke (1998). The framework serves two purposes. First, we characterize the incentives leading firms to develop training pipelines in the absence of AI. Second, we use the framework to develop a set of predictions on how AI will change those incentives, which will guide our empirical analysis in Section 5. The present framework abstracts from a full characterization of labor market equilibrium, which we defer to the full equilibrium model that will be developed as this project continues.

4.1. A simple framework of training, retention, and automation

Firms are infinitely lived and produce output using junior and senior tasks to provide distinct job-specific labor. Firms differ in the contribution of the junior-level job to output, denoted β, and firm-specific human capital, denoted κ_f. Workers live for two periods and differ in ability $\eta \sim F(\eta)$, which is initially unobserved and is portable across firms. A worker hired into a junior job in period t receives wage w_1, and the firm incurs an additional per-worker training cost c. Training reveals the ability of the worker to the training firm, but not to the rest of the market, at the end of period t. Workers also acquire firm-specific human capital κ_f, which can only be used at the incumbent firm. $\kappa_f \sim K$ is common to all workers trained at firm f. The J_t workers hired in period t become the firm's cohort of internal candidates for the senior job, C_{t+1}:

$$C_{t+1} = J_t.$$

We suppress time subscripts below. After observing ability, the firm chooses which workers in the candidate cohort to retain and which to lay off using a simple threshold rule. Workers with ability $\eta \geq \hat{\eta}$ are offered a retention wage w_2 to be promoted to the senior job in the following period. Workers with ability below $\eta < \hat{\eta}$ are laid off and enter the external market, where outside firms can observe that they have been trained but cannot observe their ability.

At the end of the first period, with probability λ, a worker also receive a preference shock $\theta \sim G(\theta)$ that exogenously changes her valuation of the outside option. θ captures reasons why a worker might leave her job that are unrelated to the job itself, such as a change in her spouse's job

location. She quits her job if she prefers the wage she could receive in the external market, v, to the equilibrium retention offer she receives from the incumbent firm: $v + \theta \geq w_2(\beta, \kappa, C)$. The resulting quit probability is

$$q(w_2(\beta, \kappa, C), v) = \lambda[1 - G(w_2(\beta, \kappa, C) - v)].$$

After realizing quits and layoffs, the firm chooses effective task inputs X_1 and X_2 for the junior and senior jobs, respectively. The firm's revenue function is given by

$$r(\beta, \kappa, X_1, X_2) = \mathcal{A}\left\{X_1^{\beta} X_2^{1-\beta}\right\}^{\rho}.$$

We assume decreasing returns to scale, $\rho < 1$.

Task inputs in the junior job can come from either junior workers J or AI technology A. AI can perform the junior task, but does not learn on-the-job or advance as human juniors do. The two inputs are linear substitutes in production of X_1:

$$X_1 = \alpha_1 J + \psi A.$$

The flow cost of employing one junior worker is $w_1 + c$, the wage plus the training cost. The flow cost of employing one unit of AI is

$$p_A + \frac{\phi}{2}\frac{A}{X_1}.$$

The cost includes a unit price p_A as well as a component that increases with the share of junior input X_1 that is automated. Intuitively, this captures a convex implementation cost of AI: the first marginal task is relatively easy and low-cost to automate, but reaching 100-percent automation is very costly. Convex implementation costs allow us to rationalize intermediate levels of automation in the junior task, $0 < \frac{A}{X_1} < 1$, while maintaining that junior workers and AI are perfect substitutes in production.

Task inputs in the senior job must come from trained workers, including those who are retained from the junior job in the previous period as well as external senior hires from the secondary market H. Expected senior input is given by

$$X_2 = \alpha_2 \Gamma(A)\left[\underbrace{C[1 - q(w_2, v)]\int_{\hat{\eta}}^{\bar{\eta}} (\eta + \kappa)\, dF(\eta)}_{\text{retained previous-period juniors}} + \underbrace{\mu^E H}_{\text{external hires}}\right],$$

where μ^E is the expected value of ability η for a worker chosen at random from the external market. The term $\Gamma(A)$ reflects the idea that AI could augment the productivity of experienced workers. It

takes the simple form

$$\Gamma(A) = 1 + (\overline{\Gamma} - 1)\left(1 - e^{-r_\Gamma A}\right), \qquad \overline{\Gamma} > 1,$$

which implies

$$\Gamma(0) = 1, \qquad \Gamma'(A) > 0, \qquad \lim_{A\to\infty} \Gamma(A) = \overline{\Gamma}.$$

where $\overline{\Gamma} > 0$ is a constant and thus does not depend on the organization-wide intensity of AI adoption.

We can summarize the firm's problem with a Bellman equation for the value a firm (β, κ) assigns to an inherited cohort C:

$$\begin{aligned} V_{\beta\kappa}(C) = \max_{J,H,w_2,\hat{\eta},A} \Big\{ & \mathcal{A}\left\{X_1^{\beta} X_2^{1-\beta}\right\}^{\rho} - (w_1 + c)J - p_A A - \frac{\phi}{2}\frac{A^2}{X_1} \\ & - w_2 C[1 - q(w_2, v)][1 - F(\hat{\eta})] - vH \\ & + \gamma V_{\beta\kappa}(J)\Big\}. \end{aligned} \tag{1}$$

We first characterize the firm's incentive to hire junior workers in the baseline without AI. Setting $A = 0$, the first-order condition for junior labor is

$$\underbrace{\frac{\rho\beta r}{X_1}\alpha_1}_{\substack{\text{flow marginal} \\ \text{revenue product}}} + \underbrace{\gamma V'_{\beta\kappa}(J)}_{\substack{\text{marginal} \\ \text{continuation} \\ \text{value}}} = \underbrace{w_1 + c}_{\substack{\text{marginal} \\ \text{flow cost}}}. \tag{2}$$

A junior worker provides value to the firm not only through current production, but also by becoming a candidate for senior work in the following period. Combining the envelope condition with the optimal retention and external-hiring conditions gives

$$V'_{\beta,\kappa_f}(C) = [1 - q(w_2, v)]\frac{v}{\mu^E}\int_{\widehat{\eta}}^{\overline{\eta}} (\eta - \widehat{\eta})\, dF(\eta) > 0. \tag{3}$$

The firm values additional internal candidates because it observes their ability and can selectively retain more productive workers. While the continuation value is not directly observed, we use the following lemma to establish an empirical analogue: the same firm characteristic that raises the continuation value also raises the share of junior candidates who advance internally.

LEMMA 1 (Internal advancement and pipeline value). *Holding external-market conditions fixed, greater firm-specific human capital increases both the marginal continuation value of the candidate*

cohort and the probability that a candidate remains with the firm:

$$\frac{dV_f'(J)}{d\kappa_f} = \frac{v}{\mu^E}[1 - q(w_2, v)][1 - F(\widehat{\eta})] > 0,$$
$$\frac{d}{d\kappa_f}\{[1 - q(w_2, v)][1 - F(\widehat{\eta})]\} > 0.$$

The proof is provided in Appendix A.2. Intuitively, the lemma establishes that ordering firms by the share of junior hires who advance internally also orders them by their unobservable continuation value. A firm's pre-AI internal share of junior-to-senior transitions serves as a revealed-preference measure of how much the firm values its internal pipeline. We therefore use this internal advancement measure to order firms by pipeline value.

4.2. Model predictions for firms' responses to AI

AI adoption and the composition of hiring. When AI adoption and junior hiring are both positive, $A, J > 0$, we obtain the following expression by combining their first order conditions:

$$\underbrace{p_A + \phi\frac{A}{X_1}}_{\substack{\text{marginal AI}\\ \text{deployment cost}}} = \underbrace{\frac{\psi}{\alpha_1}\left[w_1 + c - \gamma V'_{\beta,\kappa_f}(J)\right]}_{\substack{\text{net cost savings from replacing}\\ \text{human junior labor}}} + \underbrace{\rho(1-\beta)r\frac{\Gamma'(A)}{\Gamma(A)}}_{\substack{\text{augmentation of}\\ \text{senior workers}}}.$$

The optimality condition implies that cheaper AI increases AI intensity A/X_1, while a greater continuation value from junior hiring attenuates this response. Because junior workers and AI are substitutes in the junior task,

$$\frac{J}{X_1} = \frac{1}{\alpha_1}\left(1 - \psi\frac{A}{X_1}\right).$$

Greater AI intensity therefore reduces the number of junior workers required per unit of junior-task input. The task inputs X_1 and X_2 are not directly observed. To translate the firm's choice of AI intensity into an observable hiring prediction, decompose the ratio of junior to senior starts as

$$\frac{J}{R+H} = \frac{J}{X_1} \times \frac{X_1}{X_2} \times \left(\frac{R+H}{X_2}\right)^{-1}.$$

AI reduces J/X_1 by substituting for human labor in the junior task. The remaining terms capture the induced change in relative task demand and the number of senior workers required to provide senior input. Under Cobb–Douglas production and common senior augmentation, Appendix A.2

shows that these adjustments do not offset the direct reduction in junior labor. AI therefore lowers $J/(R+H)$, and the decline is smaller when the continuation value of junior hiring is greater.

PREDICTION 1. *All firms weakly reduce the ratio of junior starts to senior starts, $J/(R+H)$.*

PREDICTION 2. *Firms with greater continuation value, and hence greater pre-AI internal advancement, will reduce the junior:senior hiring ratio by less.*

Scale and junior hiring levels. The preceding result concerns the composition rather than the level of hiring. Junior starts can be decomposed as

$$J = (R+H)\frac{J}{R+H}.$$

AI reduces the junior share of worker starts, but its effect on optimal firm scale and the total number of starts is theoretically ambiguous. Greater productivity may expand firm output and labor demand, while automation and senior augmentation reduce the number of workers required to produce a given amount. Total junior hiring can therefore rise or fall depending on whether the scale response offsets the shift in hiring composition. This motivates examining the junior share of starts, total starts, and junior starts as separate outcomes.

5. Evidence from Copilot Adoption

5.1. Mapping model predictions to empirical hypotheses

In this section, we describe the empirical analogues of the theoretical predictions of Section 4.2 before bringing them to data. Prediction 1 hypothesizes that access to AI reduces junior starts relative to senior starts. We measure this response using the junior share of junior and senior starts,

$$\frac{J_{ft}}{J_{ft}+R_{ft}+H_{ft}},$$

which is a monotone transformation of the model ratio $J_{ft}/(R_{ft}+H_{ft})$ and is defined whenever the firm has at least one junior or senior start.

Prediction 2 hypothesizes that firms with greater continuation values of junior labor will reduce junior hiring by less. Lemma A2 shows that a firm's pre-AI internal share of senior starts,

$$\frac{R_f^{\text{pre}}}{R_f^{\text{pre}}+H_f^{\text{pre}}},$$

orders firms by the continuation value they place on junior hiring. Intuitively, firms that historically fill more senior positions internally reveal a preference for retention. Our conceptual framework microfounds these preferences in firm-specific human capital. These firms have more to lose when AI replaces junior labor and therefore preserve more junior hiring.

Empirical hypothesis 1. Copilot access reduces the junior share of junior and senior starts.

Empirical hypothesis 2. The reduction in the junior share is smaller at firms with a larger pre-AI internal share of senior starts.

Empirical hypotheses 1 and 2 map directly to the sharp theoretical predictions from Section 4.2 that admit clear signs. The model does not separately sign the response of junior hiring in levels because the decline in the junior share may be offset by a change in firm scale. To disentangle composition effects from scale effects, we also estimate effects on junior starts, internal senior starts, external senior starts, and total starts.

5.2. Copilot access and empirical design

Microsoft introduced Microsoft 365 Copilot to a limited group of enterprise customers in 2023 and made the product generally available to large enterprises in November 2023, initially subject to a 300-seat purchase minimum. Microsoft removed that minimum in January 2024, substantially broadening commercial availability. Firms nevertheless acquired access at different times.

We isolate variation in access generated by the renewal dates of firms' pre-existing Microsoft 365 enterprise contracts. These dates were scheduled before Copilot became commercially available. Because firms commonly add new products when their contracts renew, the occurrence of the first scheduled renewal after Copilot became available shifts the probability and timing of Copilot access. Our instrument is therefore an indicator that this scheduled post-launch renewal has occurred. The first stage is

$$\text{CopilotAccess}_{ft} = \pi\,\text{PostScheduledRenewal}_{ft} + \alpha_f + \delta_t + u_{ft},$$

where α_f are firm fixed effects and δ_t are calendar-period fixed effects. We first document that scheduled renewal strongly predicts Copilot access and use event-time specifications to test for differential trends before the renewal.

For outcome Y_{ft}, the second stage is

$$Y_{ft} = \beta\,\widehat{\text{CopilotAccess}}_{ft} + \alpha_f + \delta_t + \varepsilon_{ft}.$$

The primary outcome is the junior share of junior and senior starts. We then apply the same design

to the separate counts of junior starts, internal senior starts, external senior starts, and total starts. To test Hypothesis 2, we interact Copilot access with the pre-AI internal share of senior starts. In the corresponding first stage, we instrument the interaction with the scheduled-renewal instrument interacted with that same pre-AI firm characteristic.

The identifying assumption is that, conditional on firm and calendar-period fixed effects, scheduled renewal timing affects hiring outcomes only through its effect on Copilot access. We assess this assumption using balance tests, event-time pre-trends, placebo renewals before Copilot's launch, and tests for other contract changes coinciding with renewal. The IV estimates identify the response among firms whose timing of Copilot access is shifted by their scheduled renewal.

5.3. Effects on hiring composition

We first report the first-stage relationship between scheduled renewal and Copilot access. We then estimate the effect of access on the junior share of worker starts and test whether the effect is attenuated at firms with greater pre-AI internal senior sourcing. These estimates provide the direct tests of Hypotheses 1 and 2.

5.4. Hiring scale

We next separate the composition response from changes in scale. We report the effects of Copilot access on junior starts, internal senior starts, external senior starts, total senior starts, and total worker starts. This decomposition shows whether a decline in the junior share reflects fewer junior jobs, an expansion of senior hiring, or both.

5.5. Downstream career flows

Finally, we examine the early career consequences for workers entering after Copilot access. We study subsequent internal advancement, movement to senior positions at other employers, and time to the first senior transition. These outcomes are interpreted as downstream and partly exploratory: the model signs their levels only when Copilot changes the size of the entering junior cohort, and the relatively recent rollout limits the available follow-up period.

6. Conclusion

Entry-level jobs are both a source of current labor and a mechanism for producing future experienced workers. This paper studies how generative AI changes firms' incentives to create these jobs when the human capital they produce can ultimately benefit other employers. Using worker career histories,

we document that junior-job creation is highly concentrated across firms. The most junior-intensive firms, which account for one-fifth of employment, create nearly half of all junior starts. These firms also produce a disproportionate share of workers who advance into senior jobs, but most of that advancement occurs elsewhere. The market for experienced labor therefore connects firms that create entry opportunities to firms that subsequently employ the workers those opportunities produce.

Our framework shows why this connection matters for firms' responses to AI. When AI substitutes for junior tasks, firms weigh its immediate productive benefit against the continuation value of the workers they would otherwise hire and develop. The model predicts that AI reduces junior hiring relative to senior hiring, but that this shift is smaller at firms with more valuable internal pipelines. Because AI may also change firm scale, the response of junior hiring in levels is ambiguous. This distinction motivates our empirical focus on hiring composition, hiring scale, and firms' internal versus external sourcing of senior workers. We use predetermined Microsoft contract-renewal timing as an instrument for access to Microsoft 365 Copilot to estimate these responses.

The current draft leaves two central pieces to complete. First, we will test whether firms that historically fill more senior positions internally preserve more junior hiring, as the model predicts, and examine early downstream changes in internal advancement and trained-worker exports. Second, we will use the equilibrium model to aggregate these firm-level responses, quantify their implications for the future supply and allocation of experienced workers, and compare the decentralized response with the social optimum. This exercise will determine whether the private substitution toward AI produces an excessive decline in human-capital production and evaluate the scope for policies supporting junior employment or training.

The analysis also leaves open how AI will reshape the returns to skill that generate firms' pipeline incentives in the first place. In our framework, the continuation value of a junior hire depends on the value of identifying, retaining, and developing workers who will be productive in senior roles. We treat this value as a pre-existing source of heterogeneity across firms. AI may, however, change it. If AI disproportionately augments the most able workers or makes productivity in senior work more convex in ability, firms may place a greater value on training their own workers and learning who among them will excel. This force could partially offset the pipeline externality identified here. If AI instead substitutes for moderately skilled senior tasks, compresses productivity differences, or reduces the value of worker-specific knowledge, it may lower the continuation value of junior hiring and deepen the decline in entry opportunities. Understanding how AI changes the distribution of returns to skill—and how those changes feed back into training, mobility, and career advancement—is an important direction for future research.

References

Acemoglu, Daron and David Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 4B, Elsevier, 2011, pp. 1043–1171.

___ **and Jörn-Steffen Pischke**, "Why Do Firms Train? Theory and Evidence," *The Quarterly Journal of Economics*, 1998, *113* (1), 79–119.

___ **and** ___ , "The Structure of Wages and Investment in General Training," *Journal of Political Economy*, 1999, *107* (3), 539–572.

___ **and Pascual Restrepo**, "Tasks, Automation, and the Rise in U.S. Wage Inequality," *Econometrica*, 2022, *90* (5), 1973–2016.

___ **, Dingwen Kong, and Asuman Ozdaglar**, "AI, Human Cognition and Knowledge Collapse," NBER Working Paper 34910, National Bureau of Economic Research 2026.

Afrouzi, Hassan, Andres Blanco, Andres Drenik, and Erik Hurst, "Automation, Learning, and Career Dynamics," NBER Working Paper 35157, National Bureau of Economic Research 2026.

Agarwal, Nikhil, Alex Moehring, and Alexander Wolitzky, "Designing Human-AI Collaboration: A Sufficient-Statistic Approach," NBER Working Paper 33949, National Bureau of Economic Research 2025.

___ **,** ___ **, Pranav Rajpurkar, and Tobias Salz**, "Combining Human Expertise with Artificial Intelligence: Experimental Evidence from Radiology," NBER Working Paper 31422, National Bureau of Economic Research 2025. Originally issued in 2023 and revised in 2025.

Arellano-Bover, Jaime, "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size," *Journal of Labor Economics*, 2024, *42* (2), 549–589.

Arrow, Kenneth J., "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, 1962, *29* (3), 155–173.

Autor, David, "Applying AI to Rebuild Middle Class Jobs," NBER Working Paper 32140, National Bureau of Economic Research, Cambridge, MA February 2024.

___ **and Neil Thompson**, "Expertise," *Journal of the European Economic Association*, 2025, *23* (4), 1203–1271.

Autor, David H., Frank Levy, and Richard J. Murnane, "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 2003, *118* (4), 1279–1333.

Babina, Tania, Anastassia Fedyk, Alex Xi He, and James Hodson, "Artificial Intelligence, Firm Growth, and Product Innovation," *Journal of Financial Economics*, 2024, *151*, 103745.

Baker, George, Michael Gibbs, and Bengt Holmstrom, "The Internal Economics of the Firm: Evidence from Personnel Data," *The Quarterly Journal of Economics*, 1994, *109* (4), 881–919.

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

Bianchi, Nicola, Giulia Bovini, Jin Li, Matteo Paradisi, and Michael Powell, "Career Spillovers in Internal

Labour Markets," *The Review of Economic Studies*, 2023, *90* (4), 1800–1831.

Brynjolfsson, Erik, Bharat Chandar, and Ruyu Chen, "Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence," 2025. Stanford Digital Economy Lab and SIEPR working paper; revised versions have circulated.

___, Danielle Li, and Lindsey R. Raymond, "Generative AI at Work," *The Quarterly Journal of Economics*, 2025, *140* (2), 889–942.

Card, David and Thomas Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *The Quarterly Journal of Economics*, 2001, *116* (2), 705–746.

Crane, Leland D. and Paul E. Soto, "AI and Coder Employment: Compiling the Evidence," Finance and Economics Discussion Series 2026-018, Board of Governors of the Federal Reserve System 2026.

Cui, Kevin Zheyuan, Mert Demirer, Sonia Jaffe, Leon Musolff, Sida Peng, and Tobias Salz, "The Effects of Generative AI on High-Skilled Work: Evidence from Three Field Experiments with Software Developers," August 2025. Working paper.

Cullen, Zoë, Danielle Li, and Shengwu Li, "Labor as Capital: AI and the Ownership of Expertise," 2025. Preliminary and incomplete working paper; verify the current draft before citation.

Dell'Acqua, Fabrizio, Edward McFowland III, Ethan Mollick, Hila Lifshitz, Katherine C. Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R. Lakhani, "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality," *Organization Science*, 2026, *37* (2), 403–423.

Demirer, Mert, John J. Horton, Nicole Immorlica, Brendan Lucier, and Peyman Shahidi, "Chaining Tasks, Redefining Work: A Theory of AI Automation," February 2026. Working paper; stable public URL not yet verified.

Dillon, Eleanor Wiske, Sonia Jaffe, Nicole Immorlica, and Christopher T. Stanton, "Shifting Work Patterns with Generative AI," NBER Working Paper 33795, National Bureau of Economic Research 2025.

Doeringer, Peter B. and Michael J. Piore, *Internal Labor Markets and Manpower Analysis*, Lexington, MA: D.C. Heath, 1971.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock, "GPTs Are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models," *arXiv preprint arXiv:2303.10130*, 2023.

Friebel, Guido, Yao Huang, Jin Li, Soumitra Shukla, and Andrew Zhang, "Pyramids, Diamonds, and Oscillations: AI and the Structure of Internal Labor Markets," April 2026. Working paper dated April 10, 2026; title and SSRN metadata should be checked against the current draft.

Fudenberg, Drew and Luis Rayo, "Training and Effort Dynamics in Apprenticeship," *American Economic Review*, 2019, *109* (11), 3780–3812. Verify DOI against the final journal record before submission.

Gans, Joshua S. and Avi Goldfarb, "O-Ring Automation," NBER Working Paper 34639, National Bureau of Economic Research 2026.

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

___ **and Luis Rayo**, "Relational Knowledge Transfers," *American Economic Review*, 2017, *107* (9), 2695–2730.

___ **and** ___ , "Training in the Age of AI: A Theory of Career Viability," CEPR Discussion Paper 20634, Centre for Economic Policy Research 2026. The 2025 CEPR version circulated under the subtitle "A Theory of Apprenticeship Viability"; verify the title of the cited version.

Gibbons, Robert and Michael Waldman, "A Theory of Wage and Promotion Dynamics Inside Firms," *The Quarterly Journal of Economics*, 1999, *114* (4), 1321–1358.

Hall, Robert E., "New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy," NBER Working Paper 24574, National Bureau of Economic Research 2018.

Hui, Xiang, Oren Reshef, and Luofeng Zhou, "The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market," *Organization Science*, 2024, *35* (6), 1977–1989.

Huitfeldt, Ingrid, Andreas R. Kostøl, Jan Nimczik, and Andrea Weber, "Internal Labor Markets: A Worker Flow Approach," *Journal of Econometrics*, 2023, *233* (2), 661–688.

Humlum, Anders and Emilie Vestergaard, "Still Waters, Rapid Currents: Early Labor Market Transformation under Generative AI," NBER Working Paper 33777, National Bureau of Economic Research 2026. Previously circulated as "Large Language Models, Small Labor Market Effects".

Ide, Enrique, "Automation, AI, and the Intergenerational Transmission of Knowledge," CEPR Discussion Paper 20940, Centre for Economic Policy Research 2025. Also circulated as arXiv:2507.16078.

___ **and Eduard Talamàs**, "Artificial Intelligence in the Knowledge Economy," *Journal of Political Economy*, 2025, *133* (12).

Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg, "Learning from Coworkers," *Econometrica*, 2021, *89* (2), 647–676.

Konings, Jozef and Stijn Vanormelingen, "The Impact of Training on Productivity and Wages: Firm-Level Evidence," *The Review of Economics and Statistics*, 2015, *97* (2), 485–497.

Lichtinger, Guy and Seyed Mahdi Hosseini Maasoum, "Generative AI as Seniority-Biased Technological Change: Evidence from U.S. Résumé and Job Posting Data," 2025. Working paper; verify current draft metadata before submission.

Loecker, Jan De, Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *The Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

Pastorino, Elena, "Careers in Firms: The Role of Learning about Ability and Human Capital Acquisition," *Journal of Political Economy*, 2024, *132* (6), 1994–2073.

Shen, Judy Hanwen and Alex Tamkin, "How AI Impacts Skill Formation," arXiv preprint arXiv:2601.20245 2026. Verify the arXiv identifier and current version before submission.

Traina, James, "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements," Working Paper 272, George J. Stigler Center for the Study of the Economy and the State, University of Chicago Booth School of Business 2018.

Appendix A. Theory Appendix

A.1. Solution to the firm's problem and first order conditions

We can summarize the firm's problem with a Bellman equation for the value a firm (β, κ) assigns to an inherited cohort C:

$$\begin{aligned} V_{\beta\kappa}(C) = \max_{J,H,w_2,\hat{\eta},A} \{ & \mathcal{A}\left\{X_1^{\beta} X_2^{1-\beta}\right\}^{\rho} - w_1 J - p_A A - \frac{\phi}{2}\left(\frac{A}{X_1}\right)^2 \\ & - w_2 C[1 - q(w_2, v)][1 - F(\hat{\eta})] - vH \\ & + \gamma V_{\beta\kappa}(J)\Big\}, \end{aligned} \tag{A1}$$

where the junior and senior task inputs (X_1, X_2) are given by

$$\begin{aligned} X_1 &= \alpha_1 J +, \\ X_2 &= \alpha_2 \Gamma(A)\left[C[1 - q(w_2, v)] \int_{\hat{\eta}}^{\bar{\eta}} (\eta + \kappa)\, dF(\eta) + \mu^E H\right]. \end{aligned}$$

and the quit rate $q(w_2, v)$ is given by

$$q(w_2, v) = \lambda[1 - G(w_2 - v)].$$

The first order condition with respect to junior hiring J is given by

$$\boxed{\frac{\rho\beta r}{X_1}\alpha_1 + \gamma V'(J) = w_1 + c} \tag{J}$$

The first order condition with respect to AI adoption A is given by

$$\boxed{\frac{\rho\beta r}{X_1}\psi = p_A A + \phi\frac{A}{X_1}} \tag{A}$$

We can derive the first order condition with respect to the retention wage w_2:

$$\frac{\rho(1-\beta)r}{X_2}\frac{\partial X_2}{\partial w_2} = C[1 - F(\hat{\eta})]\left([1 - q(w_2, v)] - w_2\frac{\partial q}{\partial w_2}\right)$$

$$\frac{\rho(1-\beta)r}{X_2}\alpha_2\bar{\Gamma}C\lambda g(w_2 - v)\int_{\hat{\eta}}^{\bar{\eta}} (\eta + \kappa)\, dF(\eta) = C[1 - F(\hat{\eta})]\left([1 - q(w_2, v)] - w_2\lambda g(w_2 - v)\right)$$

which yields the final condition for $A > 0$:

$$\boxed{\frac{\rho(1-\beta)r}{X_2}\alpha_2\overline{\Gamma}\cdot\mathbb{E}[\eta+\kappa|\eta\geq\hat{\eta}] = \frac{1-q(w_2,v)}{g(w_2-v)} + w_2} \tag{w_2}$$

The first order condition with respect to external senior hiring H is given by

$$\boxed{\frac{\rho(1-\beta)r}{X_2}\alpha_2\Gamma(A) = \frac{v^E}{\mu^E}} \tag{H}$$

The first order condition with respect to the ability cutoff $\hat{\eta}$ is derived:

$$\frac{\rho(1-\beta)r}{X_2}\alpha_2\Gamma(A)\frac{\partial B}{\partial\hat{\eta}} = w_2\frac{\partial R}{\partial\hat{\eta}}$$
$$\frac{\rho(1-\beta)r}{X_2}\alpha_2\Gamma(A)C[1-q](\hat{\eta}+\kappa)f(\hat{\eta}) = w_2\cdot C[1-q]f(\hat{\eta}),$$

yielding the final condition

$$\boxed{\frac{\rho(1-\beta)r}{X_2}\alpha_2\Gamma(A)\cdot(\hat{\eta}+\kappa) = w_2} \tag{$\hat{\eta}$}$$

We can also combine the conditions for H and $\hat{\eta}$ to obtain the following useful result:

$$\boxed{\frac{w_2}{\hat{\eta}+\kappa} = \frac{v^E}{\mu^E}}$$

Finally, to complete the firm's equilibrium conditions, we derive the envelope condition to obtain a closed form for $V'(J)$, the marginal value of an additional worker in the inherited candidate cohort:

$$\begin{aligned}
V'_{\beta,\kappa_f}(C) &= [1-q(w_2,v)]\left[\alpha_2\Gamma(A)r_{X_2}\int_{\widehat{\eta}}^{\overline{\eta}}(\eta+\kappa_f)\,dF(\eta) - w_2[1-F(\widehat{\eta})]\right] \\
&= [1-q(w_2,v)]\alpha_2\Gamma(A)r_{X_2}\int_{\widehat{\eta}}^{\overline{\eta}}(\eta-\widehat{\eta})\,dF(\eta) \\
&= [1-q(w_2,v)]\frac{v}{\mu^E}\int_{\widehat{\eta}}^{\overline{\eta}}(\eta-\widehat{\eta})\,dF(\eta) > 0.
\end{aligned}$$

A.2. Intermediate results and proofs

LEMMA A1 (Adverse selection in the secondary market). *Suppose firms draw workers from a common ability distribution F, retain workers above a firm-specific cutoff, and face quit and exit*

shocks that are independent of ability. Then, for every firm type κ, the ability distribution of its retained workers dominates the ability distribution in the common secondary market. In particular,

$$\mu_\kappa^R > \mu^E \tag{A2}$$

whenever the firm uses an interior retention cutoff and retains a positive mass of workers.

PROOF. The mean ability of workers retained by a firm of type κ is

$$\mu_\kappa^R = \frac{\int_{\widehat{\eta}_\kappa}^{\bar{\eta}} \eta \, dF(\eta)}{1 - F(\widehat{\eta}_\kappa)}.$$

The mean ability of the workers that it contributes to the secondary market is

$$\mu_\kappa^E = \frac{\int_{\underline{\eta}}^{\widehat{\eta}_\kappa} \eta \, dF(\eta) + q_\kappa \int_{\widehat{\eta}_\kappa}^{\bar{\eta}} \eta \, dF(\eta)}{F(\widehat{\eta}_\kappa) + q_\kappa \left[1 - F(\widehat{\eta}_\kappa)\right]}.$$

Because $q_\kappa < 1$, the secondary-market flow places less weight on workers above the cutoff than the original ability distribution does. Hence, for an interior cutoff,

$$\mu_\kappa^E < \int \eta \, dF(\eta) < \mu_\kappa^R.$$

The common secondary market simply pools the outflows from every firm type:

$$\mu^E = \frac{\int E_\kappa \mu_\kappa^E \, d\Phi(\kappa)}{\int E_\kappa \, d\Phi(\kappa)}.$$

Since every μ_κ^E is below the population mean, their weighted average is also below it. Therefore, for every firm type with an interior retention cutoff,

$$\mu^E < \int \eta \, dF(\eta) < \mu_\kappa^R.$$

□

LEMMA A2 (Continuation value). *Consider firms that face the same labor-market conditions and ability distribution and have the same junior-task output share β, but differ in the value of firm-specific human capital κ_f. Suppose senior productivity of a retained worker is $h(\eta; \kappa_f) = \eta + \kappa_f$.*

Then the marginal continuation value generated by a junior hire is strictly increasing in κ_f:

$$\frac{\partial V_J(J_f;\kappa_f)}{\partial \kappa_f} > 0. \tag{A3}$$

PROOF. Combining the envelope condition with the condition for external senior hiring, the marginal continuation value $V'_{\beta\kappa}(J)$ is

$$V'_f(J) = [1 - q(w_2, v)]\left[\frac{v^E}{\mu^E}\int_{\hat{\eta}}^{\bar{\eta}}(\eta + \kappa_f)dF(\eta) - w_2[1 - F(\hat{\eta})]\right]$$

Differentiating with respect to κ_f gives

$$\frac{dV'_f(J)}{d\kappa_f} = \frac{v^E}{\mu^E}\lambda g(w_2 - v)w'_2(\kappa_f)\int_{\widehat{\eta}}^{\bar{\eta}}(\eta - \widehat{\eta})\,dF(\eta) - \frac{v^E}{\mu^E}[1 - q(w_2, v)][1 - F(\widehat{\eta})]\widehat{\eta}'(\kappa_f).$$

The retention-wage first-order condition is

$$\lambda g(w_2 - v)\frac{v^E}{\mu^E}\int_{\widehat{\eta}}^{\bar{\eta}}(\eta - \widehat{\eta})\,dF(\eta) = [1 - q(w_2, v)][1 - F(\widehat{\eta})].$$

Substituting this condition into the equation for $\partial V'(J)/\kappa$ yields

$$\frac{dV'_f(J)}{d\kappa_f} = [1 - q(w_2, v)][1 - F(\widehat{\eta})]\left[w'_2(\kappa_f) - \frac{v^E}{\mu^E}\widehat{\eta}'(\kappa_f)\right].$$

The cutoff condition is

$$w_2(\kappa_f) = \frac{v^E}{\mu^E}[\widehat{\eta}(\kappa_f) + \kappa_f].$$

Differentiating it gives

$$w'_2(\kappa_f) - \frac{v^E}{\mu^E}\widehat{\eta}'(\kappa_f) = \frac{v^E}{\mu^E}.$$

Therefore,

$$\frac{dV'_f(J)}{d\kappa_f} = \frac{v^E}{\mu^E}[1 - q(w_2, v)][1 - F(\widehat{\eta})] > 0.$$

□

LEMMA A3 (Retention share is increasing in continuation value). *Suppose a junior worker with portable ability* η *produces* $\eta + \kappa_f$ *as a senior at the firm where the worker was trained and* η *at any other firm. Consider firms facing the same external labor-market conditions* (v, μ^E)*, junior output share* β*, and ability distribution but differing in* κ_f*. Then the share of the junior candidate cohort retained by the firm,*

$$\pi_f^R(\kappa_f) = [1 - q(w_2(\kappa_f), v)][1 - F(\widehat{\eta}(\kappa_f))], \tag{A4}$$

is strictly increasing in the marginal continuation value of junior employment $V'_{\beta\kappa}(J)$.

PROOF. After substituting the cutoff condition, the retention-wage first-order condition is

$$\frac{v^E}{\mu^E} E[\eta - \widehat{\eta} \mid \eta \geq \widehat{\eta}] = \frac{1 - q(w_2, v)}{\lambda g(w_2 - v)}.$$

The left side is strictly decreasing in $\widehat{\eta}$, while the right side is strictly increasing in w_2. Differentiating this equation with respect to κ_f therefore implies that $w_2'(\kappa_f)$ and $\widehat{\eta}'(\kappa_f)$ have opposite signs.

The cutoff condition is

$$\widehat{\eta}(\kappa_f) + \kappa_f = \frac{\mu^E}{v^E} w_2(\kappa_f).$$

Differentiating gives

$$\widehat{\eta}'(\kappa_f) + 1 = \frac{\mu^E}{v^E} w_2'(\kappa_f).$$

If $\widehat{\eta}'(\kappa_f) \geq 0$, the retention-wage condition would imply $w_2'(\kappa_f) \leq 0$, contradicting the preceding equation. Hence

$$\widehat{\eta}'(\kappa_f) < 0, \qquad w_2'(\kappa_f) > 0.$$

Differentiating the retention share gives

$$\frac{d\pi_f^R}{d\kappa_f} = \lambda g(w_2 - v) w_2'(\kappa_f)[1 - F(\widehat{\eta})] - [1 - q(w_2, v)] f(\widehat{\eta}) \widehat{\eta}'(\kappa_f) > 0.$$

From Lemma A2,

$$\frac{dV_f'(J)}{d\kappa_f} = \frac{v^E}{\mu^E} \pi_f^R(\kappa_f) > 0.$$

Therefore,

$$\frac{d\pi_f^R}{dV_f'(J)} = \frac{d\pi_f^R/d\kappa_f}{dV_f'(J)/d\kappa_f} > 0.$$

Thus the retained share of junior workers is strictly increasing in the continuation value. □

A.3. Derivations of empirical predictions

This section derives the two predictions used in the empirical analysis. We compare a firm before and after access to AI, holding labor-market prices and external senior-market conditions fixed. We also hold fixed the firm's continuation value when comparing its pre- and post-access choices. Continuation values may differ across firms and generate the heterogeneous response studied below.

The derivation permits a common fixed augmentation of senior productivity after AI becomes available. Because this multiplier affects the unit cost of senior input and the number of senior workers required per unit of senior input in the same proportion, it cancels from the junior-to-senior start ratio. The argument does not allow senior augmentation to vary with the firm's chosen AI intensity or to alter the relative efficiency of internal and external senior workers.

Define the effective private cost of a human junior as

$$c_f^J = w_1 + c - \gamma(1-\delta)V_f'(J) > 0.$$

This cost equals the junior wage and training cost net of the continuation value generated by an additional junior hire. We maintain that the continuation value is locally linear in the junior cohort over the relevant range.

Junior-task input is

$$X_1 = \alpha_1 J + \psi(\beta_f)A.$$

Let

$$a_f \equiv \frac{A_f}{X_{1f}}$$

denote AI intensity. Human junior labor per unit of junior-task input is then

$$\frac{J_f}{X_{1f}} = \frac{1}{\alpha_1}\left[1 - \psi(\beta_f)a_f\right].$$

Dividing total junior-input cost by X_{1f}, the unit cost of junior-task input is

$$p_{1f}(a_f) = c_f^J \frac{1 - \psi(\beta_f)a_f}{\alpha_1} + p_A a_f + \frac{\phi}{2} a_f^2.$$

For an interior choice with $A_f > 0$ and $J_f > 0$, cost minimization gives

$$p_A + \phi a_f = \frac{\psi(\beta_f)}{\alpha_1} c_f^J,$$

so

$$a_f = \frac{1}{\phi}\left[\frac{\psi(\beta_f)}{\alpha_1} c_f^J - p_A\right].$$

The interior solution requires

$$\frac{\psi(\beta_f)}{\alpha_1} c_f^J > p_A, \qquad 0 < \psi(\beta_f)a_f < 1.$$

Substituting the first-order condition into the unit-cost expression gives

$$p_{1f}^A = \frac{c_f^J}{\alpha_1} - \frac{\phi}{2} a_f^2.$$

Before AI is available, $a_f = 0$, and therefore

$$p_{1f}^0 = \frac{c_f^J}{\alpha_1}, \qquad \left(\frac{J_f}{X_{1f}}\right)^0 = \frac{1}{\alpha_1}.$$

PROPOSITION A1 (AI and the composition of worker starts). *Under the conditions above:*

a. *AI weakly reduces the ratio of junior starts to senior starts,*

$$\frac{[J_f/(R_f + H_f)]^A}{[J_f/(R_f + H_f)]^0} \leq 1.$$

The inequality is strict when the firm chooses positive AI intensity.

b. *The proportional decline is smaller at firms with a larger marginal continuation value:*

$$\frac{\partial}{\partial V_f'(J)} \log\left\{\frac{[J_f/(R_f + H_f)]^A}{[J_f/(R_f + H_f)]^0}\right\} > 0$$

for firms with interior AI and human-junior choices.

PROOF. Under Cobb–Douglas production in junior and senior task inputs, the optimal task-input ratio satisfies

$$\frac{X_{1f}}{X_{2f}} = \frac{\beta_f}{1-\beta_f}\frac{p_{2f}}{p_{1f}}.$$

Consequently,

$$\frac{J_f}{R_f + H_f} = \frac{\beta_f}{1-\beta_f}\frac{p_{2f}}{(R_f + H_f)/X_{2f}}\frac{J_f/X_{1f}}{p_{1f}}.$$

Holding senior wages and sourcing efficiency fixed leaves

$$\frac{p_{2f}}{(R_f + H_f)/X_{2f}}$$

unchanged. A common senior-productivity multiplier also leaves this expression unchanged: it reduces both p_{2f} and $(R_f + H_f)/X_{2f}$ in the same proportion. The proportional change in the observable headcount ratio is therefore determined by

$$\frac{J_f/X_{1f}}{p_{1f}}.$$

Before AI,

$$\frac{(J_f/X_{1f})^0}{p_{1f}^0} = \frac{1}{c_f^J}.$$

After AI, equations (**??**) and (**??**) imply

$$\begin{aligned} p_{1f}^A - c_f^J\left(\frac{J_f}{X_{1f}}\right)^A &= \frac{c_f^J}{\alpha_1} - \frac{\phi}{2}a_f^2 - \frac{c_f^J}{\alpha_1}\left[1 - \psi(\beta_f)a_f\right] \\ &= a_f\left[\frac{\psi(\beta_f)}{\alpha_1}c_f^J - \frac{\phi}{2}a_f\right] \\ &= a_f\left[p_A + \frac{\phi}{2}a_f\right] > 0. \end{aligned}$$

It follows that

$$\frac{(J_f/X_{1f})^A}{p_{1f}^A} < \frac{1}{c_f^J} = \frac{(J_f/X_{1f})^0}{p_{1f}^0}.$$

Combining this inequality with equation (**??**) proves the first result. If the firm does not adopt AI, the two ratios are equal.

For the second result, the proportional response of the observable ratio is

$$\frac{[J_f/(R_f+H_f)]^A}{[J_f/(R_f+H_f)]^0} = c_f^J \frac{(J_f/X_{1f})^A}{p_{1f}^A}.$$

Using

$$c_f^J = \frac{\alpha_1}{\psi(\beta_f)}\left[p_A + \phi a_f\right],$$

this proportional response can be written as

$$\frac{[p_A + \phi a_f][1 - \psi(\beta_f)a_f]}{p_A + \phi a_f - \frac{1}{2}\psi(\beta_f)\phi a_f^2}.$$

Differentiating equation (**??**) with respect to a_f gives

$$-\psi(\beta_f)\frac{p_A^2 + p_A\phi a_f + \frac{1}{2}\phi^2 a_f^2 + \frac{1}{2}\psi(\beta_f)p_A\phi a_f^2}{\left[p_A + \phi a_f - \frac{1}{2}\psi(\beta_f)\phi a_f^2\right]^2} < 0.$$

Moreover, equation (**??**) implies

$$\frac{\partial a_f}{\partial c_f^J} = \frac{\psi(\beta_f)}{\alpha_1\phi} > 0,$$

while the definition of the effective private junior cost gives

$$\frac{\partial c_f^J}{\partial V_f'(J)} = -\gamma(1-\delta) < 0.$$

The proportional response is therefore increasing in $V_f'(J)$. Firms with a larger continuation value choose lower AI intensity and experience a smaller proportional decline in junior starts relative to senior starts. □

Mapping continuation value to the empirical proxy.. Lemma **??** establishes that the pre-AI internal share of senior starts is strictly increasing in the marginal continuation value. Therefore,

$$\frac{\partial}{\partial V_f'(J)} \log \left\{ \frac{[J_f/(R_f + H_f)]^A}{[J_f/(R_f + H_f)]^0} \right\} > 0,$$

$$\frac{\partial}{\partial V_f'(J)} \left[\frac{R_f^{\text{pre}}}{R_f^{\text{pre}} + H_f^{\text{pre}}} \right] > 0$$

jointly imply

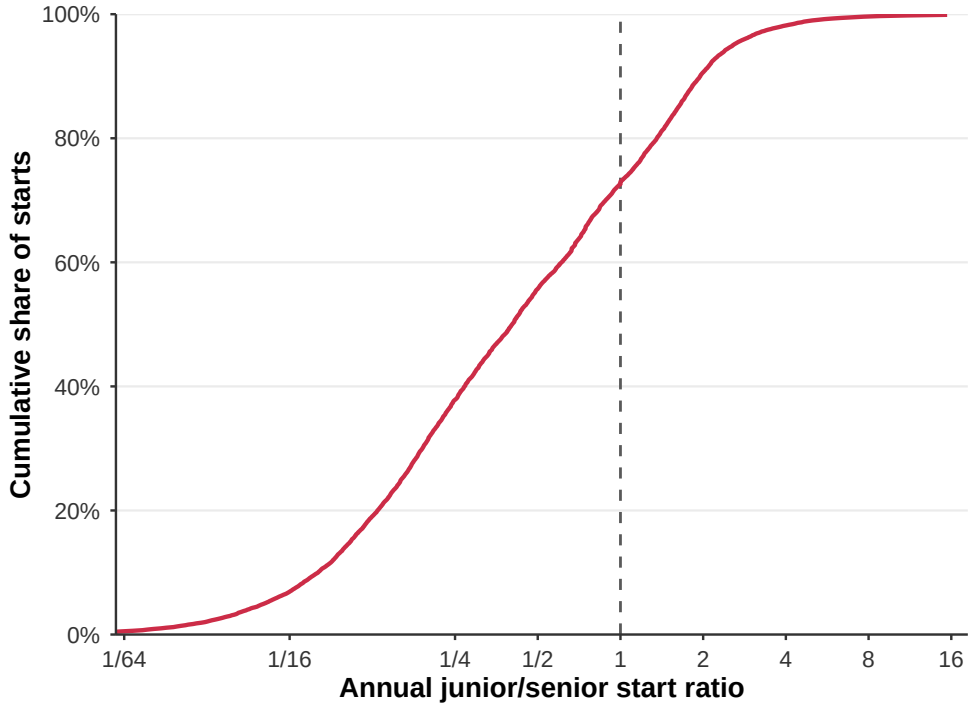
$$\frac{\partial}{\partial \left[R_f^{\text{pre}}/(R_f^{\text{pre}} + H_f^{\text{pre}}) \right]} \log \left\{ \frac{[J_f/(R_f + H_f)]^A}{[J_f/(R_f + H_f)]^0} \right\} > 0.$$

Thus the empirical counterpart of the second prediction is that Copilot causes a smaller decline in the junior-to-senior start ratio at firms with a larger pre-AI internal share of senior starts.

Appendix B. Supplemental Tables and Figures

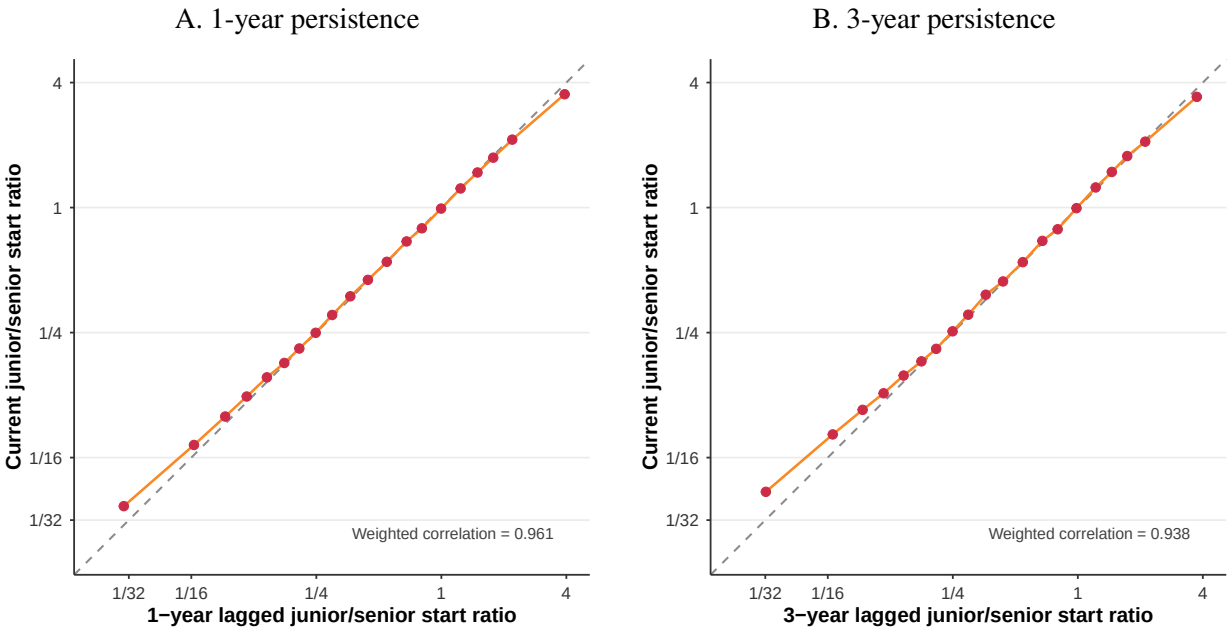
B.1. Supplement to empirical excercises in Section 5

FIGURE B.1. Distribution of pre-AI junior-senior hiring ratios



Notes: The figure presents the start-weighted empirical cumulative distribution of firms' annual junior-to-senior start ratios during 2014–2022. The sample contains 69,227 firms and 583,162 firm-year observations. Firms have at least 10 total junior plus senior starts over the sample period and are retained only if they have positive junior and senior starts in every active year. Each firm-year is weighted by its total number of junior and senior starts. Junior starts are new firm episodes beginning in a junior job. Senior starts include new firm episodes beginning in a senior job and the first junior-to-senior promotion within a continuous firm episode; a return after more than 90 days begins a new episode. Firm-title classifications use the 0.80 confidence threshold. The plotted ratio is $(J + 0.5)/(S + 0.5)$ and the horizontal axis is log-scaled but labeled in ratio levels. The dashed vertical line denotes equal numbers of junior and senior starts. Source: Lightcast Career Profiles.

FIGURE B.2. Pre-AI persistence in junior-senior hiring ratio



Notes: The figure plots firms' annual junior-to-senior start ratios against the corresponding ratio one year earlier (panel A) and three years earlier (panel B). The sample includes firms with at least 10 total junior plus senior starts over 2014-2022 and retains only firms with positive junior and senior starts in every active year. Junior starts are new firm episodes beginning in a junior job. Senior starts include new firm episodes beginning in a senior job and the first junior-to-senior promotion within a continuous firm episode; a return after more than 90 days begins a new episode. Firm-title classifications use the 0.80 confidence threshold. The ratio is transformed as $\log_2[(J+0.5)/(S+0.5)]$ for binning, while the axes report ratios in levels. Points show weighted means for 20 equal-weight bins of the lagged ratio. Pair weights equal the geometric mean of total starts in the two years. The dashed line is the 45-degree line, and the reported correlation uses the same weights.