AI, learning content, and the future supply of managers

Aroon Narayanan*
MIT

July 2026

Abstract

In a panel of roughly 9,000 U.S. public firms, AI-exposed entry occupations lost 3.6% of their headcount relative to unexposed occupations after ChatGPT's release, and 8.1% at the firms most able to deploy AI. These occupations supplied managers at 1.33 times the average rate. Why aren't firms preserving the work that produces their managers? Answering this question requires measuring which entry work builds managerial judgment, but standard output exposure indices instead measure how much of a job's output can be produced by AI. I measure this attribute – a job's *learning content* – using an index built from O*NET. Learning content predicts who reaches senior management, raising its probability by one third of the base rate per standard deviation, while output exposure does not once learning content is accounted for. At AI-capable firms that lose their trained juniors to rivals, the contraction is also deeper in learning-rich work, but at the firms that keep their trainees, learning-rich work shows no additional reduction beyond its output exposure. Firms face little incentive to shield learning-rich work, because the training it provides is general. Over half of the senior workers entering the panel were trained at another firm, so a firm that preserves such work ends up training its rivals' managers. Moreover, the cost of cutting the learning-rich occupations does not arrive quickly enough to incentivize immediate action, since a smaller cohort produces fewer future managers with an elasticity of only about two-thirds and a delay that spans a career. When the financial crisis cut finance entry hiring 23% below trend in 2009, the missing cohort produced 17% fewer seniors over the following decade, and firms did not offset the loss with outside hiring. The cohorts already missing due to the AI contraction imply a manager-stock deficit of 1–2% even if hiring recovers, and about 8.5% over two decades if it does not.

JEL codes: J24, O33, J23, M51, M53.

Keywords: generative AI; learning content; managers; human capital; internal labor markets; automation; training externality.

*I am grateful to Vishan Nigam, Roi Orzach, and Sukrit Puri for excellent comments and suggestions. Email: aroon@mit.edu.

1 Introduction

Entry-level cognitive work produces the output a firm needs and also trains the firm's future managers. A firm hires a junior analyst to perform the analyst's work, and the managerial skill that the worker builds in performing it is a by-product of her experience on the job. Much of this work can now be performed by generative AI. Since ChatGPT's release, entry hiring in the occupations most exposed to it has contracted, documented in payroll records (Brynjolfsson et al., 2025) and in résumé and vacancy data (Hosseini Maasoum and Lichtinger, 2025). I find the same contraction in a panel of roughly 9,000 U.S. public firms, where it begins in the quarter of ChatGPT's release. The occupations that are now contracting are also the ones that made managers. Among the workers who reach senior roles, 84% are first observed in an exposed occupation, and the exposed occupations, where the contraction is concentrated, supplied managers at 1.33 times the rate of average entry work.

Why aren't firms preserving the work that produces their managers? The firms that are cutting entry hiring evidently value managerial talent: they fill senior roles through long internal tournaments, promoting from within rather than hiring seniors from outside. But promotion draws on training the worker received earlier. Preserving the rung means investing in the training of new workers, and this investment does not pay off. The training that entry work provides is general human capital, so a firm that pays for it captures only part of the return (Becker, 1964; Acemoglu and Pischke, 1998, 1999). Over half of the workers who reach senior roles and whose entry is observed in the panel were trained at another firm, so a firm that maintains the manager-making rung ends up training its rivals' managers along with its own. What return the firm does keep arrives only decades later, because an entrant takes years to become a manager and the senior stock decays slowly even after a major contraction in entry hiring.

I take this explanation to the data by constructing an index of a job's *learning content* – the managerial judgment that a worker builds in doing the work. Standard indices capture a job's *output exposure*, that is, how much of its output can be produced by a machine (Webb, 2020; Felten et al., 2021; Eloundou et al., 2024). The two attributes need not coincide: a task that is easy for a machine to perform can still be rich in the judgment that is learned by doing it, and a task that is hard to automate can teach very little. I build the index from O*NET descriptions of what workers do on the job, using task content alone, with no employment or outcome input. The index correlates 0.47 with output exposure across occupations, so the two are not proxies for each other.

To understand how learning content moderates the effect of AI on the manager supply,

I build a model with entry work as a joint product: doing it yields output for the firm and managerial training that accrues to the worker. A role is staffed when its output value plus the share of the future management rent that the firm expects to keep is higher than the entry wage, and AI, by substituting for the trainee's output, pushes the marginal roles below the bar. A firm can choose to preserve the formative content of entry work by keeping juniors in decisions and reviews of AI-performed tasks, at an organizational cost, or it can choose to fully automate the job.

The model yields four predictions. First, the reduction in entry work should be proportional to output exposure times a firm's ability to deploy AI, and absent at firms unable to deploy it. Second, learning content, and not output exposure, should predict which entry workers become managers, and should predict it at both the entry firm and at other firms, because the judgment built by the work is general human capital. Third, because preserving training pays off only if the firm keeps a sufficient proportion of its trainees, there should be a higher reduction in learning-rich work at capable firms that lose their trainees to poaching, relative to those that retain them or cannot deploy AI. Fourth, a new hire becomes a manager only after many years, so a missing cohort should lead to a smaller manager stock with a delay, and with an elasticity less than one due to higher internal promotions.

I take the model's predictions to individual employment histories, assembled by Revelio Labs from public professional profiles and other sources, which I aggregate to a firm-level panel of roughly 9,000 U.S. public firms from 2017 to 2024, with each worker's employer, six-digit occupation, and seniority level. The levels run from an entry rung of junior analysts and associates up to senior management. Promotion through these levels is what I call the manager pipeline. I measure occupational output exposure using the task-based index of Eloundou et al. (2024), which captures how much of a job's work can be done by AI. A firm's AI capability is its pre-shock share of workers in AI roles (machine-learning engineers, data scientists, and workers with hands-on AI skills, among others). These workers are the in-house expertise that a firm needs in order to deploy AI.

By 2024, headcount in the exposed entry occupations is down 3.6% relative to the unexposed ones within the same firm, and the reduction comes through hiring rather than separation. The model's first prediction is that this decline should show up only where firms can deploy AI. I find that the top tercile of pre-shock AI capability cuts exposed entry headcount by 8.1%, while the bottom tercile, which employs the same occupations, does not cut at all. A decline that predated ChatGPT would move the low-capability firms too, which addresses a major concern in the literature (Frank et al., 2026). Nor is the decline just a reflection of the reorganization of remote-suitable work (Lambert and Schindler, 2026): the contraction

affects work that cannot be done from home as well. The remaining objection is that AI capability is itself chosen, as is the timing of adoption. I therefore instrument capability using the firm's pre-shock university recruiting. A firm that has long drawn its workers from the universities that produce AI talent is AI-capable for reasons that predate the release, and instrumented capability reproduces the contraction. I also measure adoption directly, using the spread of generative-AI skills through a firm's workforce, hiring and job postings for AI roles, and mentions of AI on earnings calls. Different firms adopt at different times, and the reduction coincides with the timing of each firm's own adoption.

Within the entry rung, learning content predicts who advances. Among the 16.1 million workers who entered these firms in 2017–2019, a standard-deviation increase in the learning content of an entry occupation predicts a 0.7-percentage-point higher probability of reaching senior management on a 2.1% base, about one third of the base rate, holding firm and entry cohort fixed and controlling for output exposure non-parametrically. Learning content predicts reaching a senior role at a different firm about as well as at the entry firm, which implies that the built skill is general human capital. Output exposure does not predict reaching management at all once learning content is included. Learning content, though, does predict the contraction net of output exposure: in line with the model's third prediction, among AI-capable firms learning-rich hiring falls by 1.4% per standard deviation of learning content at the firms whose trained juniors are routinely poached, and not at all at the firms that retain them.

Some 300,000 entrants are already missing from the 2023–24 cohorts. At the contracting occupations' manager-supply rate of 1.33 times the average, they imply about 20,000 fewer future managers – several thousand more than a cut of the same size falling on average entry work would have removed. Across the panel's cohorts, I estimate that a missing cohort shows up in the manager stock with an elasticity of only about two-thirds, because smaller cohorts promote a larger share of their survivors, and only as senior slots open. When the 2008–09 financial crisis left the 2009 finance entry cohort 23% below trend, the cohort's survivors were promoted at slightly higher rates and it still produced 17% fewer seniors – an implied elasticity of 0.73, close to the panel's two-thirds. The model's transition matrix, fit to the pre-crisis years, reproduces the realized path of senior production with a correlation of 0.97. Finance hiring recovered within two years, so the overall reduction in the senior stock was around 2.6%. In the current AI contraction, the shortfalls that have already been realized lock in a manager-stock deficit of 1–2%, and a contraction that persists takes it to about 8.5% over two decades.

The contraction is individually rational for each firm and excessive for the economy as a whole. Since not all of the rents generated by a manager-making job go to the employer that staffs it, the private benefit from staffing such a job falls short of the social benefit by exactly the

share of the rents that leaves with the worker. AI widens this wedge, and it does so most at the firms that can deploy it but cannot keep their trainees. Effectively, when a firm replaces entry-level workers with AI it records a cost savings, but the savings are overstated from a welfare perspective, because part of the reason for paying entry-level workers was to develop future managers. None of this is captured by current measures, which address what a job produces but not what it teaches. This paper's learning-content index measures what the work teaches using only public task descriptions, and similar indices can be built for any occupation-level descriptions. With such measurements, the correction to the staffing externality is to subsidize hiring into entry jobs rich in learning, with the size of the subsidy tied to the learning the job still provides. The delayed effect on the manager stock implies that there is still time to reduce the severity of the impact using such corrections.

Related literature. This paper contributes primarily to the literature on generative AI and the early career. In payroll data, Brynjolfsson et al. (2025) find that there was a decrease in employment for young workers in jobs exposed to AI; Hosseini Maasoum and Lichtinger (2025) document seniority-biased declines in résumé and vacancy data; Azar et al. (2025) report that younger workers' wages and their share of the workforce at exposed firms decline; Muehlemann (2025) reports that German training firms that adopted AI offer more apprenticeship slots while redirecting continuing training at their already-skilled workers. A parallel theoretical literature views the early career as a training technology that AI disrupts. Garicano and Rayo (2025) examine a single master-apprentice relationship and derive a technological threshold that determines whether such relationships stay viable when AI can perform the junior tasks; Ide (2025) examines how entry-level automation affects the transfer of tacit knowledge through generations; Afrouzi et al. (2026) study career dynamics when workers learn from the unautomated tasks remaining; Friebel et al. (2026) model changes in internal labor market structure from pyramid to diamond forms. The closest model, in Zhang (2026), is built around a poaching externality in a training game of identical firms, compounded by managerial myopia, and addresses whether universities can fill the gap created when firms pull back from training. This paper measures the learning content of entry work, demonstrates that two distinct characteristics govern the contraction of entry work and the formation of managers, and identifies the firms that strip training and shows that they do so because they do not retain the workers that they train.

The model builds on the economics of employer-provided training and general human capital: Becker (1964) showed that in a frictionless labor market it is workers who should pay for general skills, while Acemoglu and Pischke (1998, 1999) showed how information

frictions and wage compression can leave firms willing to finance training that they cannot fully appropriate; in practice much employer training is general, may double as screening, and hinges on turnover and other frictions (Autor, 2001; Loewenstein and Spletzer, 1999; Leuven, 2005), and skills that are general one by one may be specific in their weights, combinations, or task context (Lazear, 2009; Kambourov and Manovskii, 2009; Gathmann and Schönberg, 2010). This literature treated the poaching of trained workers as a constraint on investment, while I study the removal of the output value that funded the investment in the first place, finding that the same poaching now determines whether there is any opportunity for the firm to hire for those skills at all.

How firms organize managerial talent is the subject of the economics of internal labor markets and hierarchies: internal labor markets fix entry ports, promotion ladders, learning, and assignment (Doeringer and Piore, 1971; Baker et al., 1994; Gibbons and Waldman, 1999); workers are screened for the higher roles through tournaments and job matching (Lazear and Rosen, 1981; Jovanovic, 1979); in knowledge-based theories of hierarchy, managers earn their value as scarce problem solvers and coordinators (Aghion and Tirole, 1997; Garicano, 2000; Garicano and Rossi-Hansberg, 2006); and the quality of management matters for productivity and for keeping workers (Bloom and Van Reenen, 2007; Lazear et al., 2015). In management research on strategic human capital, that talent is a source of advantage: it builds up over time, folds into firm routines, is developed through selection and training, and mobility frictions protect it (Barney, 1991; Dierickx and Cool, 1989; Hatch and Dyer, 2004; Ployhart and Moliterno, 2011; Ployhart et al., 2014; Campbell et al., 2012). The talent-management literature poses it as a make-versus-buy problem: firms build internal pools, then lose part of the investment when employees leave (Cappelli, 2008; Collings and Mellahi, 2009). This paper ties that process to generative AI's substitution for the entry work where managerial judgment is formed.

The paper also builds on the task approach to automation, in which technologies substitute for labor in the tasks that they can perform and create new tasks around the frontier (Autor et al., 2003; Acemoglu and Autor, 2011; Autor, 2015; Acemoglu and Restrepo, 2019, 2022), and on the AI literature that brings this logic to large language models, which measures occupational exposure and documents productivity effects and early firm-level reorganization (Webb, 2020; Felten et al., 2021; Eloundou et al., 2024; Noy and Zhang, 2023; Brynjolfsson et al., 2023; Babina et al., 2024; Humlum and Vestergaard, 2025). These papers measure what AI can do or how workers use it; their indices are measures of output exposure. The nearest measure in spirit is the decision-making intensity of Deming (2021), built from related O*NET activities to track the rising return to decision-intensive work. Relative to these, this paper adds

a measured attribute – a task's learning content.

Section 2 sets out the model and its estimating equations, and Section 3 then covers the data, the two indices, and the within-firm design. The contraction, and its selection on output exposure, are presented in Section 4. Section 5 shows that learning content makes managers and that the training leaks, and the accounting exercise for the projected managerial deficit is in Section 6. Section 7 provides implications for measurement and policy. Section 8 concludes.

2 Model

Entry work at the junior analyst level produces both the end product the firm needs and a by-product: the experience that makes its future managers. I formalize this *joint product* using the general-training externality from Acemoglu and Pischke (1998, 1999) applied to technology that substitutes for output produced by trainees, and I let each firm choose how it deploys that technology: preserving the formative content of entry work at an organizational cost, or letting the machine perform it. The model delivers the hiring equation of Section 4, a deployment sorting prediction examined at the end of that section, the formation and portability prediction of Section 5, the cohort passthrough of Section 6, and, once the training is established to be general, the under-provision of manager-making work that is developed in Section 7. Proofs are in Appendix A.

2.1 The environment

Agents and timing. The economy contains a continuum of competitive firms and these firms have occupations representing their entry work. Time is discrete, each period carries a one-step discount, represented by β, and the probability of an individual exiting the labor force each period is δ.[1] In each period a firm-occupation cell, denoted (f, o), will open up a measure μ_{fo} of junior positions available to be filled in the cell's entry level (positions and projects that can be assigned to the entry-level employees of the firm). The firm fills a position when it is worth filling at the current junior wage (w_t^J); hiring in a given cell (f, o) at time t is the flow $H_{fo,t}$. The firm also holds a stock of managers (M) who manage the work of the juniors and are important due to their scarcity, similar to other models in the knowledge hierarchy literature (Garicano, 2000; Garicano and Rossi-Hansberg, 2006).

[1] The constant δ keeps the two-state model simple. The dynamic accounting of the pipeline in Section 6 tracks movement from one state to another using the state-dependent exit and promotion probabilities from the estimated chain, where juniors exit at 17% and seniors at 11% per year.

Firms differ along two dimensions: the pre-determined AI capability of the firm ($a_f \in [0,1]$), which allows firms to implement new AI technologies and represents the extent to which a firm possesses in-house expertise regarding AI, and the pre-determined poaching rate of the firm ($\phi_f \in (0,1)$), which represents the probability that an employee trained at the firm will leave before the firm can promote her.

The two arms of entry work. Roles produce different levels of output even within a single occupation: a role in occupation o carries an output value ω, drawn from a c.d.f. Γ over the cell's roles; write $S \equiv 1-\Gamma$ for its survival function, $\gamma > 0$ for its density, and $\eta \equiv \gamma/S$ for its hazard; I assume Γ has a log-concave survival function, so the hazard η is nondecreasing (Bagnoli and Bergstrom, 2005). Occupations, meanwhile, have attributes of their own in terms of what they provide to the junior employees working in those jobs. Specifically, an occupation has both output exposure (ν_o), which is the proportion of output that a competent machine can generate in the occupation, and learning content (ψ_o), which sets the probability $\pi(\psi_o)$, increasing in the content ($\pi' > 0$), that doing the job will turn a junior employee into a manager.[2] Output exposure starts to matter once AI of aggregate capability θ_t becomes available – θ_t is zero until the public release and $\bar{\theta}$ from then on – because a role's output value then drops by the displacement $\Delta_{fo,t} = \theta_t\, a_f\, \nu_o$: nothing before the release, nothing at firms unable to deploy the technology, and biggest where exposed occupations meet capable firms.

Trained workers walk away with general human capital, so a frictionless market would pay them their marginal product and leave no rents for the training firm (Becker, 1964). However, following Acemoglu and Pischke (1998, 1999), I assume that frictional costs compress the wage structure so that a retained manager is paid less than her marginal product ($G_M > w^M$), allowing the firm to earn quasi-rents per period ($\rho = G_M - w^M > 0$). Discounted over the expected length of her time as a manager (her tenure), a retained manager is worth $V^M = \frac{\rho}{1-\beta(1-\delta)}$ to the firm. A trained junior stays through to promotion only with probability $1-\phi_f$; with probability ϕ_f (the poaching rate) she is lost to another employer first. Therefore, a firm retains only the fraction $1-\phi_f$ of the management rent generated, and the remaining ϕ_f goes to the firms that poach.

Deployment. Once AI is available, a firm able to deploy it also chooses *how* to deploy it. After the release, each firm selects a deployment architecture $q_f \in \{A, S\}$, chosen once, before the period's role values are drawn, and common across its entry occupations. Architecture A

[2] Whether a job builds managerial skills or simply reveals existing skill does not affect the analysis, since both types of jobs lose workers to poaching regardless of type. Therefore, I use "training" to denote both processes.

preserves the formative content of entry work: juniors keep the decisions, analysis, and review that build judgment, with AI as a tool, so the formation probability remains $\pi_A(\psi_o) = \pi(\psi_o)$. Architecture S is substitution-intensive: AI absorbs the formative tasks along with the output, and the entry roles that remain teach less, $\pi_S(\psi_o) = (1-\lambda)\,\pi(\psi_o)$ with $\lambda \in (0,1]$. Preservation incurs an architectural cost at the firm level. Sustaining the review layers, mentoring, and human-in-the-loop assignments that keep juniors learning requires an organizational outlay $K(D_f)$, expressed in per-period units comparable to operating surplus, where $D_f \equiv \bar{\theta} a_f$ is the firm's post-release deployment scale, $K(0) = 0$, $K'(D) \geq 0$, and $K(D) > 0$ for $D > 0$.[3]

The firm's problem. The firm hires an employee when the net benefits to be realized by the new employee (post-AI) will exceed the cost of the employee's services to the firm. These benefits include the output still produced in the role once AI has done its part as well as the expected management rent generated by their employment; the cost is the going junior wage. The rent is represented as $\pi_{q_f}(\psi_o)(1-\phi_f)\beta V_t^M$ when the individual transitions directly to promotion after being hired.[4] An individual will be employed if:

$$\omega \;\geq\; \omega^*_{fo,t} \;\equiv\; w_t^J - \pi_{q_f}(\psi_o)(1-\phi_f)\,\beta V_t^M + \theta_t a_f v_o. \tag{1}$$

Thus, the hiring flow of employees in this occupation is given by $H_{fo,t} = \mu_{fo} S(\omega^*_{fo,t})$, which falls off according to a semi-elasticity $\eta(\omega^*)$ in response to a rise in the displacement (Appendix A). The presence of the rent reduces the amount needed for a role to meet or exceed the threshold; however, both the displacement and the rate of poaching increase that amount. The wage and the poaching rate are exogenous variables, as is the manager value, since I view this model as a partial equilibrium representation of the hiring margin behavior.[5]

[3] K collects the organizational inputs and the foregone operating gains of running AI around the trainee rather than in place of her, expressed per period; a firm that deploys no AI gives nothing up, which is why $K(0) = 0$. Before the release every firm trivially holds the preserving architecture.

[4] If there were multiple periods before promotion, the β would represent the present-discounted probability of surviving to promotion. Therefore, every equation that contains βV^M, including the wedge developed in Section 7, is subject to the deeper discounting implied by that longer time horizon. The time lag until promotion is restored in Section 2.4.

[5] A fuller discussion of how these results feed back into each other through changes in V^M can be found in Appendix A. That appendix shows that a thinner stock of managers leads to higher rents, which cushion but do not completely reverse the decline in employment.

2.2 The hiring equation and the two margins

The log transformation of the hiring flow and linearization around the cell's initial, pre-release thresholds turn the staffing constraint into the hiring equation estimated in the empirical sections, approximately and with the linearization's remainder in the error term (Appendix A):

$$\log H_{fo,t} \;=\; \nu_{fo} \;+\; \nu_{ft} \;-\; \eta_f\,\theta_t\,a_f\,\nu_o \;+\; \eta_f\,\beta(1-\phi_f)\,\pi_{q_f}(\psi_o)\,V_t^M \;+\; u_{fo,t}, \tag{2}$$

where ν_{fo} includes the cell-specific scale and level terms, ν_{ft} includes firm-specific changes in the entry wage, and η_f is the firm's local hazard rate – i.e., a hazard that is assumed common across the firm's entry occupations. If the hazard instead depends on output exposure, the design's exposure-decile-by-period effects capture the common factor, and estimation requires that the remaining hazard-weighted displacement within deciles be unrelated to learning content. This assumption is examined in detail using the cubic and common-support specifications in Section 4.

Prediction 1 (The contraction)**.** *After the release, holding learning content and the deployment architecture fixed, within-firm differences in hiring into exposed versus unexposed entry occupations decrease by $\eta_f\bar{\theta}a_f(\nu_e-\nu_u)$: in proportion to the amount of AI exposure and to the extent of the firm's ability to use AI, and not at all at a firm with no ability to deploy it.*

Before the release the displacement is zero; the release turns it positive, lowering the output value of exposed roles and pushing the marginal ones below the hiring threshold. The fall is steeper the higher the firm's local hazard η_f, and because the displacement is the product $\theta a \nu$, it is largest where exposed occupations meet capable firms and disappears at a firm that cannot deploy AI. Holding the deployment architecture fixed isolates this output channel.

Formation is the other margin. Learning content determines which entrants will become managers, conditional on being hired, via $\pi(\psi_o)$; how much output a machine produces relative to a job is uninformative about what doing the job teaches the worker. And because the judgment that entry work builds is general human capital, trained workers who leave carry the formation value to their next employer.

Prediction 2 (Formation and portability)**.** *Among workers hired before the release, the probability of becoming a manager rises with the learning content of the entry occupation and, conditional on learning content, does not vary with its output exposure. Because the human capital built is general, learning content predicts becoming a manager at the entry firm and at a different firm alike.*

The external gradient is the falsifiable implication of skill being general, since skill specific to the entry firm would predict advancement there and nowhere else. The prediction concerns pre-release entrants because their formation was completed while every firm still preserved the formative content of entry work.

2.3 The deployment margin

Preserving the formative content of entry work pays off only through the managers it later produces, and the firm keeps only the share $1-\phi_f$ of that value. The choice between the two architectures therefore depends on retention: the organizational cost of keeping juniors in the loop is worth paying only when enough of the trained juniors stay. Write $r_f \equiv (1-\phi_f)\,\beta V^M$ for the formation value the firm retains per unit of π, and write the staffing thresholds of condition (1) under the two architectures as

$$T^A_{fo} = w^J + D_f v_o - r_f\,\pi(\psi_o), \qquad T^S_{fo} = w^J + D_f v_o - r_f(1-\lambda)\,\pi(\psi_o) = T^A_{fo} + r_f\lambda\,\pi(\psi_o),$$

so moving to the substitution-intensive architecture raises every occupation's threshold by the retained formation value that the move destroys, $r_f\lambda\pi(\psi_o)$. The firm weighs the operating surplus that preservation protects, summed over its entry occupations, against the cost $K(D_f)$.

Proposition 1 (The deployment margin)**.** *Denote the private benefit of preserving training by*

$$B^F(r_f, D_f) \equiv \sum_o \mu_{fo} \int_{T^A_{fo}}^{T^S_{fo}} S(u)\,du.$$

The firm preserves training ($q_f = A$) if and only if $B^F(r_f, D_f) \geq K(D_f)$. Assume a positive measure of the firm's roles has $\pi(\psi_o) > 0$ and a positive measure of its exposed roles has $v_o\,\pi(\psi_o) > 0$. Then B^F is strictly increasing in r_f, and $B^F - K$ is strictly decreasing in D_f; hence the preserving set is an upper interval in r_f: there is a possibly corner retention cutoff $r^*(D)$ such that the firm preserves training if $r_f \geq r^*(D_f)$ and adopts the substitution-intensive architecture otherwise; wherever the cutoff is interior it is weakly increasing in D, and a firm with $D_f = 0$ preserves. Conditional on the architecture, the post-release change in log hiring of an occupation with pre-release threshold $x_o = w^J - r_f\,\pi(\psi_o)$, written $g_q(\psi_o) \equiv \log S(T^q_{fo}) - \log S(x_o)$, satisfies

$$\frac{\partial g_S}{\partial \psi_o} = r_f\,\pi'(\psi_o)\Big[(1-\lambda)\,\eta(T^S_{fo}) - \eta(x_o)\Big], \qquad \frac{\partial g_A}{\partial \psi_o} = r_f\,\pi'(\psi_o)\Big[\eta(T^A_{fo}) - \eta(x_o)\Big] \geq 0,$$

so hiring falls more steeply in learning content under S if and only if $\lambda > 1 - \eta(x_o)/\eta(T^S_{fo})$, while

under A learning-rich work is weakly protected, and exactly unaffected under a constant hazard.

Under the constant-hazard benchmark, a firm that switches will see its reduction deepen with learning content, while a firm that preserves will see no such deepening. A higher rate of poaching operates on this deepening in two ways. First, it reduces the rent left over and induces the firm to switch; however, conditional upon the firm switching (and therefore destroying some of the formation value generated by training) the poaching reduces the deepening, because a firm with fewer trainees has less to lose. This means that there is no single, monotone relationship between poaching and the deepening. Dividing the deepening by the retention rate $(1-\phi_f)$ strips out the second force and leaves a function that is zero for preserving firms and rises with learning content for those that switch.

Assumption 1 (Preservation is affordable at full retention)**.** *For deployment scales D in the observed support, $K(D) \leq B^F(\beta V^M, D)$.*

Assumption 1 leads to the deployment distortion being entirely in the appropriability of training. Absent poaching, no firm would strip the formative content of entry work, so a firm that does has been pushed across the cutoff by the trainees it loses.

Prediction 3 (Deployment sorting)**.** *Net of output exposure, and under the constant-hazard benchmark, the post-release deepening of the cut in learning content, divided by the pre-shock retention rate $1-\phi_f$, is zero at preserving firms and $\eta_f \beta V^M \lambda \pi'(\psi_o)$ at switching firms. Across firms it is weakly increasing in deployment capability and in the poaching rate, so the switching firms form an upper set in (D_f, ϕ_f) whose boundary slopes downward; it is zero at firms unable to deploy AI at any poaching rate; and under Assumption 1 the poaching cutoff among capable firms is interior: firms above it switch and firms below it preserve.*

The prediction is rejected by a deepening at capable firms that keep their trainees without one at capable firms that lose them, by any deepening at firms unable to deploy, or by a hole in the upper set. What the model does not pin down is the location of the cutoff within the poaching distribution, which depends on the unobserved cost K, so in Section 4 I impose a median split of the poaching distribution rather than estimating the cutoff, which the model does not identify.

2.4 The manager stock and its dynamics

A promoted junior does not get to management right away, so the firm has a "middle tier" between entry rung and senior roles – reflecting the large number of rungs on the career ladder.

The "stock" of employees at each of these levels is represented by x_t, where t represents time. The model of how employees move through this process is:

$$x_{t+1} = \mathbf{P}\, x_t + e\, H_t, \tag{3}$$

where $\mathbf{P}$ represents the matrix of probabilities of promotion or leaving at each rung, e routes new entrants into the junior state, and H_t represents the flow of new hires.

Promotions fill senior positions as they open up. I represent the "eventual" number of promotions from a given cohort based on its initial size (J_0) relative to the number of openings that are available in the senior group: the per-entrant yield is the entry-weighted average of the promotion probability scaled by congestion, $y(J_0) = \bar{\pi}\,(J_0/\kappa M)^{\eta_c}$, where κM is the per-period flow of senior openings, $\bar{\pi}$ is the entry-weighted average of $\pi_{q_f}(\psi_o)$ over the cohort's entry cells – equal to the average of $\pi(\psi_o)$ for cohorts hired before the release – and $\eta_c \in (-1, 0]$ is a congestion parameter – i.e., how responsive the yield of promotions is to an increase in the number of entrants – so a cohort of size J_0 will eventually produce $J_0\, y(J_0)$ managers. At $\eta_c = 0$ the chain has constant yield; in the limiting case $\eta_c = -1$ it is a pure vacancy chain, where a fixed number of openings is shared among whatever number of people are waiting.

Prediction 4 (Buffered passthrough)**.** *A shortfall in one entry cohort, with the rest of the pipeline at its steady state, passes through to the managers that cohort produces with elasticity* $1+\eta_c$*: one-to-one at constant yield and less than one-to-one under slot constraints, since a thinner cohort converts a larger percentage of its surviving members. The managers that are missing will appear only when the cohort would have reached management.*

Proposition 2 (Dynamics of the manager stock)**.** *Assume the chain* (3) *is stable with nonnegative hazard rates and use the yield from Prediction 4. A permanent proportional reduction in hiring reduces the long-run stock at each tier in the same proportion,* $\frac{\Delta M}{M} = \frac{\Delta H}{H}$*, and at every date of the transition the fixed-hazard version of the chain is an upper bound on the deficit.*

The chain is homogeneous of degree one, so the hiring reduction scales all stocks, keeps $J_0/\kappa M$ constant, and restores the yield to its baseline value. Along the transition, the earlier affected cohorts experience higher-than-baseline yields, which bounds the fixed-hazard path from above. The buffer provided by the congestion margin absorbs part of a missing cohort while not fully replacing what was lost. Deficits emerge in the cross section only as thinner cohorts age into management. Therefore, even though inflows have been reduced, the cross-sectional picture can remain healthy.

Mapping to the data. The output-exposure index v of Section 3 provides measurements of v_o and scales the displacement; the learning-content index ψ ranks occupations according to the formation probability π.[6] Prior to the shock, AI capability identifies those firms capable of implementing AI ($a_f > 0$, equivalently a positive post-release deployment scale D_f), and the pre-shock rate at which juniors leave to go to other firms measures the poaching rate ϕ. Hiring equation (2) and Prediction 1 are estimated in Section 4, which closes with the deployment sorting of Prediction 3: a learning-content deepening of the cut, divided by retention, that rises with capability and poaching (Proposition 1). The formation regressions of Section 5 test both parts of Prediction 2; the cohort regressions of Section 6 estimate the passthrough $1 + \eta_c$ of Prediction 4 and the hazards along the chain; and Section 7 develops the under-provision of manager-making work and the deployment distortion (Appendix A, Proposition A1).

3 Data and measurement

In this section, I describe the worker panel, the AI measures based on occupation and firm levels, and the within-firm research design.

Workforce panel. The worker information is from Revelio Labs. This is a dataset created by collecting employment history of individuals through their public professional profiles and online resumes and linking them to the public identifier of the employer. I create a worker-by-year panel of 2017–2024 using these employment histories. In addition, I retain all U.S. workers at firms linked to Compustat.[7] For each worker-year there is an employer field, a six-digit occupational code (SOC), and an ordinal seniority level ranging from 1 (entry) to 7 (executive), assigned by Revelio based on titles and roles. I limit the analysis to firms with at least fifty worker-years in 2021, leaving approximately 9,000 firms.[8] I classify seniority into entry (levels 1–2), middle (levels 3–4), and senior (levels 5–7). Level 1 is relatively low-paid work, only around half of it exposed to AI (average 2021 salary near $48,000), while level 2

[6] This index measures the *level* of learning content in an occupation, and every test uses it solely as a ranking of occupations by formation probability. Whether AI erodes what the surviving work teaches is the deployment margin of Section 2.3: the architecture itself is not observed, and λ is identified only through the hiring pattern that Proposition 1 assigns to it.

[7] Revelio has coverage after 2024 as well, however, due to an update lag in the profiles, the latest years are undercounted in levels. Therefore, I stop the yearly panel in 2024. The quarterly hire-date series (Figure 2) takes the difference between high- and low-capability firms, and the lag cancels out in that difference. So, it continues until 2025Q2.

[8] The minimum number of worker-years increases precision but does not affect the identification: the contraction is negative and significant under all panel definitions (Appendix C). Also, the contraction is greater on the threshold sample since that sample leans toward the larger, higher-capability firms where the cut concentrates.

is the analyst-and-associate rung – cognitive analysts and associates – exposed to AI at around 79% (average salary near $92,000). Table 2 gives some details on the sample. The panel contains around 192 million worker-years from 9,322 firms.

Managers and senior roles. The percentage of worker-years in management occupations (SOC major group 11) rises from around 10% at the bottom rung to 47% in the upper tier and 74% at the executive rung (per level in Table 16), and managerial titles rise along with it. The senior tier holds the largest single share of management-occupation worker-years (about 37%) and is itself about 47% management occupations, the rest senior individual contributors. Since I consider the senior tier as the empirical equivalent of the model's manager – the highest point on the promotion ladder where management occupations accumulate – I refer to both the seniority-based pipeline and stock as the "manager" pipeline and stock for convenience and apply the strict management-occupation definition (SOC major group 11) when necessary.

Output exposure. Occupation output exposure is the task-based measure of Eloundou et al. (2024) that indicates the proportion of an occupation's tasks whose completion time is materially reduced by a large language model. Thus, it reflects what part of a job's work a model can accomplish. I denote it as ν and say that an occupation is *exposed* if it has ν in the top tercile. A second output-exposure measure, namely the broad AI-occupational-exposure index of Felten et al. (2021), constructed before generative AI, is one that I utilize for comparison purposes; the two measures correlate 0.86 across occupations.

Learning content. I create a learning-content index ψ using the O*NET database version 29.1 as the average amount of importance of ten formative cognitive work activities such as decision-making, analyzing data, developing strategies and judging among others (see Appendix B for the list of all ten). It was created for 757 out of the 858 entry occupations that were in the panel, or roughly 90 percent of entry person-years; the remainder are new SOC codes that O*NET 29.1 does not have ratings for. I exclude supervisory activities, which entry-level workers do not yet perform, so ψ measures what entry work teaches rather than what management requires. It correlates at 0.47 with output exposure ν. The correlation between ψ and cognitive complexity, education and Job Zone is 0.6–0.8. Routine jobs have high levels of output exposure but low levels of learning content (see Table 1). These are jobs where AI can perform the work but doing it teaches little. Professional-analytic jobs are those with high levels of output exposure and high levels of learning content. There, AI can perform the job, and doing it teaches the worker how to manage. Using the language of the model, ν measures

Table 1: Representative entry occupations, by output exposure and learning content

	High output exposure ν	**Low output exposure** ν
High learning content ψ	data scientists, engineers, web administrators, recruitment specialists	electricians, chefs, aircraft mechanics, security guards
Low learning content ψ	customer-service representatives, administrative assistants, retail salespersons, shipping clerks	fast-food workers, waiters, stockers, delivery drivers

Entry occupations (seniority 1–2), split at the mean of standardized output exposure ν (Eloundou et al., 2024) and learning content ψ, excluding any management and supervisory titles.

an occupation's level of output exposure while ψ orders occupations based upon the formation probability π.

Firm AI capability. The share of workers employed by a firm in AI occupations prior to the shock (averaged over 2018–2021) is used to measure a firm's ability to implement AI. These employees are identified using skills, titles, and degrees: machine-learning engineers and data scientists, and more broadly workers whose profiles carry hands-on AI skills. They are the internal expertise a firm uses to implement generative AI for performing entry work. Since the correlation between the firm's employment-weighted task exposure and its AI capability is only 0.26, the capability measure supplies cross-firm variation that the collinear task-exposure proxies lack. Since I observe capability and not implementation, the estimates represent the impact of having the capacity to substitute.

AI usage. Exposure measures what AI *can* do. To evaluate the two-margin split against a behavioral record of what AI *actually* does, I include the task-specific AI-usage measure from the Anthropic Economic Index (AEI) (Anthropic, 2025), which is matched to O*NET tasks and then aggregated up to occupations. After orthogonalizing with respect to learning content, it is included in the regressions as a control for what AI actually does. The exposure and capability measures, along with their correlations, are shown in Panel C of Table 2.

Identification. The timing of the shock is when ChatGPT became publicly available on November 30, 2022. The main design involves comparing exposed vs. unexposed occupations within the same firm and year. For each outcome y, for each firm i, for each exposure group $g \in$

Table 2: Sample and measurement

Panel A. Sample			
Firms	9,322		
Worker-years, 2017–2024	≈192 million		
Match	Revelio Labs × Compustat, U.S.		
Panel B. Workforce by seniority (2021 cross-section)			
	Worker-years	AI-exposed share	Mean salary
Entry, level 1 (lower-wage)	7.8M	0.50	$48k
Entry, level 2 (cognitive analyst)	6.7M	0.79	$92k
Middle (levels 3–4)	6.1M	0.85	$116k
Senior (levels 5–7)	4.0M	0.83	$190k
Panel C. AI measures			
corr(output exposure v, broad AI exposure)	0.86		
corr(learning content ψ, output exposure v)	0.47		
corr(firm AI capability, firm task exposure)	0.26		

Output exposure is the share of an occupation's tasks a large language model can materially speed up (Eloundou et al., 2024); an occupation is *exposed* if it is in the top tercile. Learning content ψ is the mean O*NET importance of ten formative-cognitive activities (Appendix B). AI capability is the firm's pre-shock (2018–2021) AI workforce share. The first two correlations are across occupations; the third is across firms, with task exposure aggregated to the firm as the employment-weighted mean over its 2021 workforce.

$\{\text{exposed}, \text{unexposed}\}$, and for each year t,

$$y_{igt} = \sum_{s \neq 2021} \beta_s \, \mathbf{1}\{t = s\} \cdot \text{Exposed}_g + \nu_{it} + \nu_{ig} + \varepsilon_{igt}, \tag{4}$$

where ν_{it} represents firm-by-year fixed effects, ν_{ig} represents firm-by-group fixed effects, standard errors are clustered by firm, and 2021 is treated as the base year. Equation (4) represents the event-study version of the model's hiring equation (2). The firm-by-year fixed effects capture any shock common across a firm's exposed and unexposed occupations (e.g., a hiring freeze, a demand shock, the post-COVID recovery); therefore, identification arises solely from the differential changes in a firm's exposed versus unexposed occupations. When a series exhibits some pre-2022 trend behavior I fit a simple linear trend over the pure pre-period (2017–2019 and 2021) and estimate the deviation from that trend; hence, the design depends on there being little-to-no deviation in 2022 relative to that trend (Appendix B). The coefficients corresponding to pre-2022 periods are used to test for parallel trends, and the gradient representing AI capability across firms provides additional evidence consistent with an AI explanation. Section 4 takes up the leading alternatives to an AI reading head-on, and instruments capability using a shift-share predictor constructed from where the firm recruited

before the shock. Because the shock is one common date rather than staggered adoption, the heterogeneous-timing critiques from the recent difference-in-differences literature do not apply here; event-study deviations from the pre-period path are reported in the convention of Borusyak et al. (2024).

4 Output exposure and the contraction

4.1 Baseline results

The exposed entry rung shrinks, and the reduction does not begin until the shock hits. Across all firms, the within-firm differences in entry headcounts between exposed and unexposed occupations drop 2.1% in 2023 and 3.6% by 2024 (Table 3). Both the exposed and unexposed rungs were still drifting together as of the release date of 2022, and if I treat the shock as having occurred in 2021, the placebo shows no effect. Thus the 2023–24 decline is an actual break in the slow pre-existing drift of both rungs (Appendix C).[9] The decline is primarily due to less hiring of new employees into exposed occupations versus unexposed occupations. Entry hiring into exposed occupations declined by 3.6 percentage points compared to unexposed occupations, while entry separations did not rise (Table 3).[10] The contraction is concentrated in the cognitive entry position – the analyst-and-associate level (about 79% of whom were exposed), which fell by 7.5 percent, while lower-wage entry positions and the middle tiers remained unchanged and the senior and executive rungs declined a significant but far smaller 1.8 to 3.3 percent (Appendix C, Figure 6).

Finally, the cut is most extreme when firms are able to use AI (Prediction 1). I split firms into thirds based on their AI capability before the shock hit, and find that the high- vs. low-capability thirds move together through 2022 and separate starting only later. By 2024, the gap in entry headcounts between exposed and unexposed positions at high-capability firms had decreased by approximately 8%, whereas at low-capability firms it has decreased by approximately 2% (Figure 1). At a quarterly frequency, the difference in the time series for high- minus low-capability firms fluctuates around zero without a trend through the four years preceding

[9] I also test for violations of parallel trends in the way described by Rambachan and Roth (2023). On the raw within-firm event study for entry, the confidence set built on relative magnitudes excludes zero for violations of post-period trends up to 0.75 times the largest pre-period violation.

[10] Humlum and Vestergaard (2025) found that generative AI had only minor effects on incumbent workers' wages and hours worked. As such this represents a different margin of employment – the worker side – that is consistent with the firm-side reduction in entry hiring seen above, which operates on the number of junior positions that firms open rather than on existing employees' compensation.

Table 3: Within-firm entry headcount: exposed vs. unexposed

Entry rung (seniority 1–2)	Coef.	(s.e.)
Headcount, 2023 (dev. from pre-trend)	−0.021	(0.0041)
Headcount, 2024	−0.036	(0.0052)
Hiring rate, 2024	−0.036	(0.0025)
Exit rate, 2023	−0.005	(0.0021)

Within-firm specification (4) for the entry rung (seniority 1–2), detrended against the 2017–2019 and 2021 path; firm-by-year and firm-by-group fixed effects; standard errors clustered by firm. The two entry seniority levels enter as separate cells, so each coefficient is the within-level exposed-minus-unexposed gap, which the cognitive-analyst level dominates (level 2, −7.5%; Appendix C). The exit rate is shown for 2023, the last year in which a separation is identified (a 2024 separation requires 2025 continuation, which the panel does not yet observe). Observations 288,290 firm-by-exposure-by-level-by-year entry cells; 9,320 firm clusters.

the shock; it turns negative in Q4 of 2022 (the quarter of the release); reaches −0.28 by Q4 of 2024; and continues to be negative through Q2 of 2025 (the last quarter of data available), without evidence of reversal (Figure 2). Since the comparison is within the firm-year, a technology downturn or post-pandemic correction that affects every single worker at a firm would be differenced away.

4.2 Evidence for an AI interpretation

Frank et al. (2026) propose that the decline in occupations exposed to AI had already begun prior to ChatGPT; their position would suggest that a 2022 decline does not represent a new downturn, but merely a continued trend. What matters here, however, is not the absolute amount of exposure decline, but how differently AI-capable firms and less capable firms respond. A downward trend in exposed occupations would affect both types of firms – less capable firms hire into the same occupations as the high-capability ones – but only the high-capability firms move (Table 4, Panel A). At the high-capability firms the detrended contraction in exposed employment was flat through the release year (i.e., −0.7% in 2022, s.e. 0.61) and did not break until afterward, to −3.1% in 2023 and then further to −6.6% in 2024.[11] Conversely, occupations exposed to equal degrees at less capable firms experienced no such contraction (the 2024 deviation was +0.9%), and the difference between the high-capability and the low-capability firms was −7.5%.

[11]These detrended deviations subtract the fitted pre-2022 path, which was itself drifting down (the exposed-specific trend is −0.4% per year in the top capability tercile and −0.8% in the bottom); hence they fall short of the high-capability cut of −7.8% measured against the 2021 baseline (Table 20) and of the −8.1% detrended top-tercile deviation in Table 21, which is estimated on a differently-constructed capability tercile, and similarly fall short of the raw 2024 gap of about 8 percent seen in Figure 1.

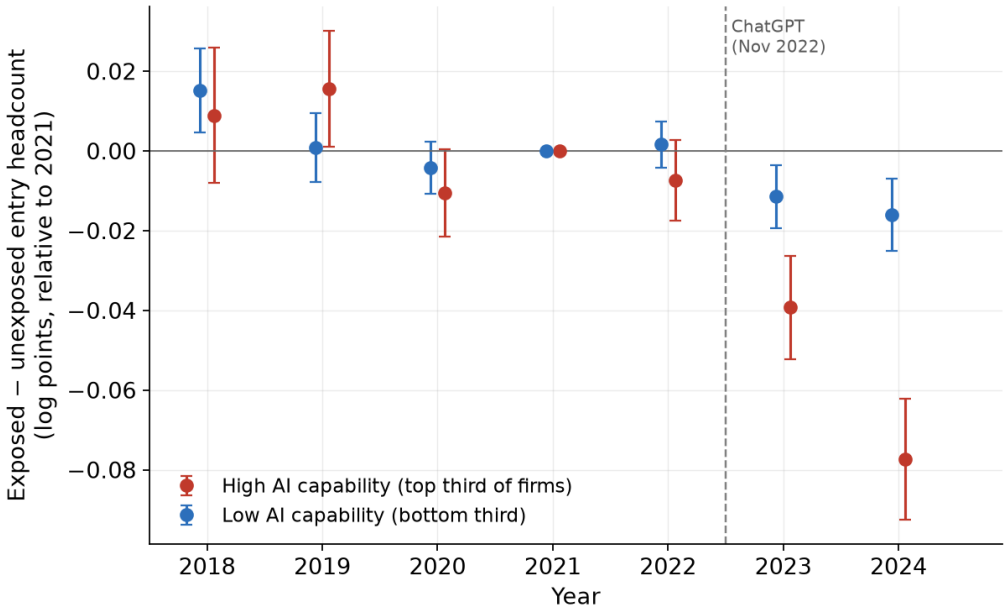


Figure 1: Within-firm exposed-minus-unexposed entry headcount, by pre-shock AI-capability tercile.

Specification (4) for entry (tier 1–2) headcount, estimated separately for firms in the top and bottom terciles of pre-shock AI capability. Points are exposed-minus-unexposed coefficients relative to 2021, with 95% intervals.

Lambert and Schindler (2026) interpreted the decline as remote-suitable work being reorganized, not as an AI effect. When splitting the exposed rung based on suitability for remote work, I find that AI-exposed work that was *not* suitable for remote work still contracted at high-capability firms (−3.9%, s.e. 1.3; Table 4, Panel B), and so an effect exists among jobs that cannot be worked remotely. However, remote-suitable exposed work declined about double that (−8.2%), and so remote work amplified the decline without creating it.

The capability gradient above is estimated off the firm's predetermined AI-workforce share, which is a firm characteristic rather than an instrument: high-capability firms may have differences unrelated to AI that impact their hiring into exposed entry work. Comparisons based on adoption timing suffer from the same issue (Hosseini Maasoum and Lichtinger (2025) found the junior contraction at firms that began employing generative-AI integrators, and the timing of adoption is itself a firm choice). Therefore, I instrument capability with a shift-share predictor built from where a firm recruits its workers. For each firm, I calculate its pre-shock mixture of recruiting sources and multiply it by each source's leave-firm-out propensity to produce workers destined for AI occupations. If a firm draws from AI-talent-rich universities, it is predicted to be more AI-capable. Whether the recruiting shares are computed from all workers or from the firm's non-AI workers alone, the first stage is strong ($F = 616$ for all workers;

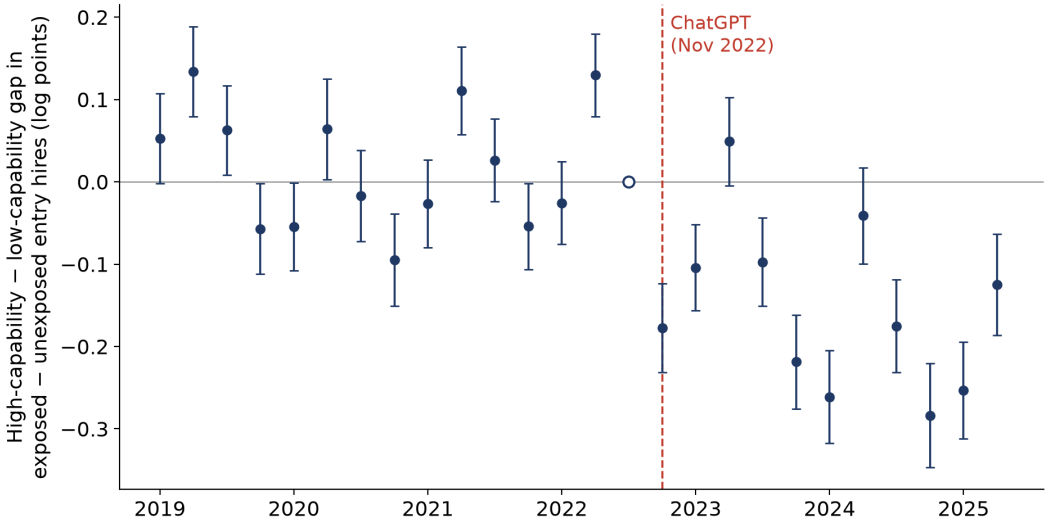


Figure 2: Quarterly high-minus-low-capability difference in the exposed-minus-unexposed entry-hiring gap.

Relative to 2022Q3 (hollow marker); points with 95% intervals, standard errors clustered by firm. Estimated from hire dates in the Revelio position file for public firms matched to the panel (top against bottom capability terciles), with firm-by-quarter and firm-by-exposure fixed effects in differences. Because the series is a high-minus-low-capability triple difference, the profile-update lag that caps the annual level panel at 2024 nets out across the comparison cells, so the quarterly hire-date series runs through 2025Q2 with stable precision.

$F = 361$, coefficient 0.58, for non-AI workers), and therefore the instrument reflects a firm's overall recruiting footprint and not some mechanical representation of its AI hiring (Table 5, Panel A).

The contraction reappears in the reduced form: predicted capability multiplied with the exposed-by-post design produces −2.6% by 2024 and places the 2020 placebo at zero (Panel B). Pooled over the post period, 2SLS puts the effect of a one-standard-deviation increase in capability on the exposed rung at −3.9% – slightly larger than the OLS estimate of −3.1% (Panel C). The recruiting-source predictor also addresses the geographic confound that where AI-capable firms choose to locate is itself correlated with the contraction, but the predictor correlates only weakly with the firm's headquarters-state AI share (0.27; the headquarters state is the modal state of the firm's pre-shock workforce) and the first stage continues to hold once that share is controlled for (Panel A). The predictor is also specific to AI talent: it survives a horse race against engineering-and-science and non-technical-professional recruiting predictors constructed the same way (Appendix C). Because this is a shift-share design, identification depends on the recruiting footprint and the leave-out shifters being exogenous. What the instrument identifies is the total effect of capability – the output displacement together with any capability-induced change in deployment architecture – rather than the displacement pa-

Table 4: Evidence for an AI interpretation: the two identification challenges

Panel A. Detrended exposed contraction by AI-capability tercile (Frank et al.)			
Dev. from pre-2022 path	2022	2023	2024
High-capability firms	−0.007	−0.031	−0.066
	(0.0061)	(0.0089)	(0.0115)
Low-capability firms	+0.012	+0.006	+0.009
	(0.0037)	(0.0058)	(0.0072)
High − low (2024)	−0.075		
Panel B. Exposed contraction by remote-suitability, high-capability firms (Lambert and Schindler)			
Dev. from pre-2022 path	2022	2023	2024
Exposed, *not* remote-suitable	−0.003	−0.020	−0.039
	(0.0066)	(0.0101)	(0.0126)
Exposed, remote-suitable	−0.008	−0.038	−0.082
	(0.0070)	(0.0099)	(0.0128)

Within-firm exposed-versus-unexposed entry log headcount, detrended against the 2017–2019 and 2021 path, firm-by-year and firm-by-group fixed effects, clustered by firm. *Panel A*: estimated separately within the top and bottom terciles of pre-shock AI capability; the AI-attributable contraction is the high-minus-low differential, flat through the release year and breaking at 2023–24. *Panel B*: among high-capability firms, the exposed rung split by occupation remote-suitability (Revelio's modeled position-level measure, averaged to occupations; correlated 0.46 with v); the contraction is present in non-remote exposed work and larger in remote-suitable work. Observations per estimate: Panel A 46,176 (high) and 47,584 (low); Panel B 45,956 and 46,138; about 2,980 firm clusters each.

rameter alone. Predicted capability also correlates moderately with pre-shock junior poaching (0.21–0.23 across the two share definitions). The first stage remains strong with the poaching rate controlled (F above 330 for both share definitions), and in the reduced form the 2024 coefficient keeps its sign and most of its size when a poaching triple is added (−1.7% to −1.4% per standard deviation of the instrument, on the subsample where poaching is measured).

To test whether the capable firms actually deploy AI, I use four indicators. These include how widely generative-AI skills exist among a firm's employees, the number of new hires for jobs that involve AI, as well as both the fraction of all job postings which mention AI roles and how often generative-AI technology is mentioned during earnings calls. Each of these indicators deepens the 2024 exposed cut off a flat 2020 placebo. The AI-role posting share also moves with the firm's own contemporaneous adoption, and not with its adoption one year ahead (Appendix C, Table 19) – and each indicator is strongly predicted by the shift-share instrument, which implies that the capability stock is being realized as actual use of AI.

Table 5: Shift-share instrument for firm AI capability

Panel A. First stage: observed AI capability on predicted (shift-share) capability	Coef.	(s.e.)
Predicted capability, all-worker recruiting shares	+0.596	(0.024)
first-stage F (all-worker shares)	616	
Predicted capability, non-AI-worker recruiting shares	+0.584	(0.031)
first-stage F (non-AI shares)	361	
adding HQ-state AI share	+0.526	(0.029)
corr(predicted capability, HQ-state AI share)	0.27	
Panel B. Reduced form: exposed × year × predicted capability (detrended)	Coef.	(s.e.)
2020 (placebo)	−0.000	(0.004)
2022 (release)	−0.008	(0.004)
2023	−0.014	(0.005)
2024	−0.026	(0.007)
Panel C. 2SLS: exposed × post × capability, instrumented by exposed × post ×Z		
2SLS	−0.039	(0.007)
OLS, for comparison	−0.031	(0.005)

Predicted capability $Z_i = \sum_u s_{i,u}\, a_u^{(-i)}$, where $s_{i,u}$ is firm i's pre-shock (2017–2021) share of workers recruited from university u and $a_u^{(-i)}$ is the leave-firm-out share of u's panel workers ever in an AI occupation; built from Revelio education records. Panel A: firm-level first stage (9,001 firms), standardized; the non-AI-worker shares exclude the firm's own AI workers from the recruiting mix. Panel B: the within-firm detrended design of Table 4 with predicted capability in place of observed capability, per standard deviation, firm-clustered. The reduced form reproduces the observed-capability contraction (Table 4) and is flat in the 2020 placebo year. Observations 9,001 firms (Panel A, heteroskedasticity-robust); 141,200 firm-by-exposure-by-year cells on 8,957 firm clusters (Panels B–C).

4.3 The deployment pattern

Under the local constant-hazard benchmark, the null of a fixed deployment architecture implies that learning content adds nothing to the cut once output exposure is absorbed flexibly, and under an increasing hazard it weakly protects learning-rich work (Appendix A). Prediction 3 makes a deepening of the cut with learning content the observable sign of an architecture switch. Therefore, I contrast occupational groups that share an output exposure but vary in learning content at the same firm and time, controlling for v with decile-by-year fixed effects and detrending against the pure pre-period path. Learning content deepens the cut. The coefficient on $\psi \times$ post is −0.38% (s.e. 0.08), the deviation grows from −0.30% in 2023 to −0.63% in 2024 with the 2022 coefficient flat at −0.02% (Table 7, Panel A), and cubic and common-support controls for v give the same answer (Table 6). Without the detrend the coefficient flips sign (+0.25%), but that specification fails the design's own pre-trend test

($\psi \times 2022 = +0.54\%$): high-learning-content work was gaining share within firms before the release, and the undetrended coefficient is that path continuing (Appendix D).

Prediction 3 also orders this deepening across firms. It is rising in capability and in poaching, zero without deployment capability, with an interior retention cutoff among the capable. Splitting firms at the medians of pre-shock AI capability and of the pre-shock poaching rate, the entire effect sits in the cell that is ranked highest by the ordering (Table 7): at capable firms whose juniors are routinely poached, the 2024 deviation is -1.4% per standard deviation of ψ (s.e. 0.30); at capable firms that retain their trainees it is -0.05%; and low-capability firms show nothing on either side of the poaching median. Within the capable-poached cell the deepening is largest in the top tercile of learning content (-2.5%, against -0.47% for the middle tercile, both relative to the bottom) – the profile a single firm-level switch produces when formation rises steeply in ψ, which it does: entrants from top-tercile occupations reach the senior tier at 9.1%, against 4.7% in the middle tercile and 2.8% at the bottom.

A possible concern here is that the deepening is related to where AI is actually used rather than what the work teaches, since the learning-content index is built from task descriptions and could proxy the occupations AI has penetrated. Adding the Anthropic Economic Index, a direct measure of actual AI usage, provides evidence against this. Net of output exposure, AEI AI usage barely moves entry hiring (-0.09% on its own, -0.11% alongside learning content), while learning content continues to have its full effect (-0.43%) in the same regression (Table 6, columns 4–5). This is precisely what the model implies: the structural parameter λ applies to formation content – the skills employees learn on the job – and not to production content – the output they create – so the reduction is larger with what the work teaches, not with where AI is used.

Table 6: Learning content deepens the contraction, net of output exposure

Within-firm entry headcount, × post	(1)	(2)	(3)	(4)	(5)
Learning content ψ	−0.0038	−0.0033	−0.0035		−0.0043
	(0.0008)	(0.0008)	(0.0010)		(0.0009)
AI usage (Anthropic Economic Index)				−0.0009	−0.0011
				(0.0006)	(0.0006)
Output-exposure control	decile	cubic	decile	decile	decile
Sample	full	full	common supp.	full	full

Within-firm entry log headcount, firm-by-occupation-by-year cells (seniority 1–2, 2017–2024), detrended against the pure pre-period path; firm-clustered standard errors in parentheses; per standard deviation of each measure. All columns absorb output exposure v non-parametrically (v-decile-by-year fixed effects) except column 2 (cubic in v); column 3 restricts to the common-support deciles. Column 5 enters AI usage orthogonalized to learning content. A negative ψ × post means high-learning-content work is cut *more* once v is flexibly controlled. The undetrended specification, which reverses the sign and fails the 2022 pre-trend test, is in Appendix D. Observations 6,474,945 firm-by-occupation-by-year entry cells and 24,747 firm clusters in the full-sample columns; the common-support and AEI columns are slightly smaller. These firm-by-occupation cells are the full Compustat-matched set, not subject to the fifty-worker-year screen of Table 2; restricting to that 9,321-firm panel leaves the deepening unchanged (−0.0040, s.e. 0.0009).

Table 7: The deployment pattern: the deepening by capability and retention

Learning content ψ, per SD	Coef.	(s.e.)
Panel A. ψ ×year deviations, full sample (ref. 2021)		
2022 (release year)	−0.0002	(0.0008)
2023	−0.0030	(0.0011)
2024	−0.0063	(0.0014)
Panel B. ψ × 2024, by pre-shock capability and poaching (median splits)		
Capable, poached	−0.0142	(0.0030)
Capable, retainers	−0.0005	(0.0029)
Low-capability, poached	−0.0036	(0.0035)
Low-capability, retainers	−0.0023	(0.0024)
Panel C. Within capable-poached firms: ψ terciles (ref. bottom tercile)		
Middle tercile × 2024	−0.0047	(0.0066)
Top tercile × 2024	−0.0252	(0.0068)

Detrended within-firm entry log headcount with v-decile-by-year fixed effects, firm-clustered standard errors. Capability is the pre-shock AI-workforce share; the poaching rate is the pre-shock (2018–2021) junior job-to-job rate to other firms (median 0.073, firms with at least fifty observable junior transitions); both split at their medians. Firm counts by cell: 2,826 / 1,675 / 1,675 / 2,825. The 2022 coefficient is flat in every cell (largest $|t|$ of 1.2). The four cells satisfy every inequality of the ordering in Prediction 3. Formation rates among 2017–2019 entrants by ψ tercile are 2.8%, 4.7%, and 9.1% for the senior tier, so the Panel C profile matches the convexity of formation in ψ.

5 Learning content and manager formation

The entry rung makes managers, primarily in the exposed occupations, but slowly and mostly for other firms. Most senior-tier workers start out on the exposed rung: 84 percent are first observed in an exposed occupation (83 percent when the count is restricted to workers first observed as juniors), and an exposed worker will reach the senior tier at approximately 2.4 times the rate of an unexposed worker (Figure 3; Table 8). Promotion is a high-stakes, selective tournament. Seven years after entering in 2017, 33 percent of the cohort have left the firm, 53 percent remain entry-level, and only 4 percent have advanced to the senior tier. It screens based upon ability – those who would eventually advance to a senior role entered the company at a 14 percent higher salary than those who would remain entry employees. The training received by workers is general and primarily captured by other companies. More than half of the senior-tier workers that were identified as having been entry-level employees at some point were entry-level employees at a different company, and junior employees are more than two times as likely to exit to another company versus being promoted internally (Table 8).

Within the entry rung, learning content predicts which entry-level employees will progress to more senior positions. From the 16.1 million workers who entered these companies in 2017–2019, entry-level employees in occupations that had higher levels of learning content had a greater likelihood of advancing to more senior positions. For example, while holding firm and entry cohort constant and controlling for output exposure using a non-parametric method, a one-standard-deviation increase in ψ increased the likelihood of achieving a senior position by 1.7 percentage points on a base rate of 5.0 percent and increased the likelihood of attaining senior management status by 0.7 percentage points on a base rate of 2.1 percent – about one-third of each base rate (Table 9, columns 1–2). By itself, output exposure was less effective at predicting advancement to the senior tier (1.2 percentage points), and it added no predictive value for reaching management once learning content was included. Since the comparison is made within the firm, the learning content effect does not reflect which firms promote more, and since the cohorts entered prior to ChatGPT, it also does not represent the impact of the shock.

Moreover, the formation effect persists at the level of an individual entrant, where first-job learning content raises senior job attainment by 0.9 percentage points, net of output exposure (column 4). Instrumenting first-job learning content with the supply of entrants in the occupational field during the entry year provides an estimate of the same sign and a larger magnitude (3.6 percentage points; column 5) – however, these results cannot be termed causal due to a weak first stage ($F \approx 7$).

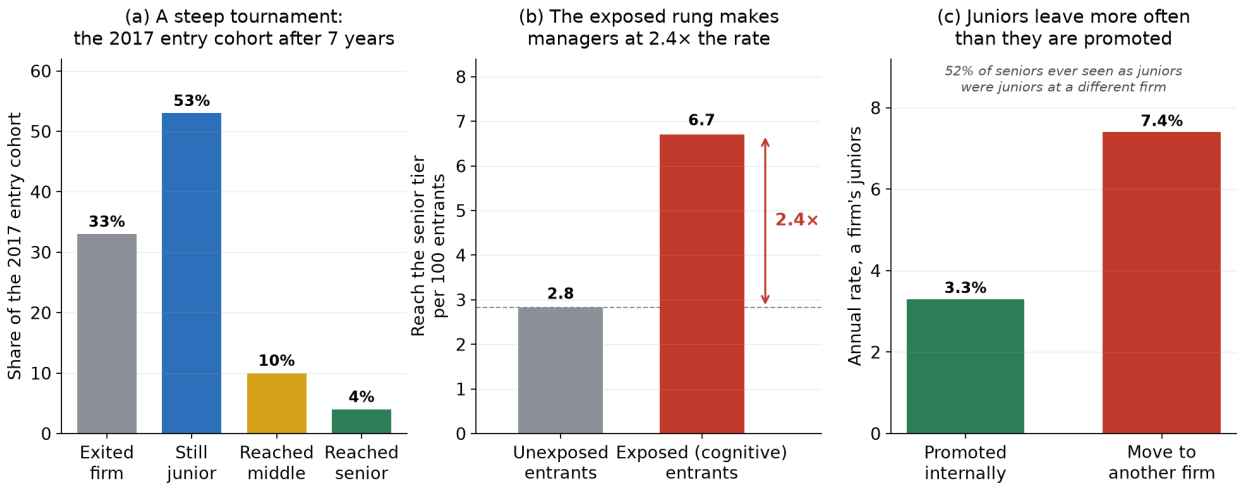


Figure 3: Career outcomes from worker histories, 2017–2024: cohort status, advancement by exposure, and mobility.

(a) Status of the 2017 entry cohort after seven years. (b) Share reaching the senior tier per 100 entrants, exposed vs. unexposed entry occupations. (c) Internal promotion versus outward moves for a firm's juniors, and the share of senior-tier workers who entered as juniors elsewhere.

The formative content created by entry jobs is transferable across employers. When examining the locus of employment at which entrants achieve a senior position, I find that learning content is similarly predictive of achieving a senior position both in the entry firm and in a firm that is separate from the entry firm, with the same pattern for senior management (Table 10). High-learning-content entrants do not simply obtain senior positions at their original employer – they obtain them regardless of where they began their career – therefore, what entry jobs build is general human capital rather than entry-firm-specific sorting. Finally, when controlling for individual attributes (i.e., entry salary and degree), I find that learning content still predicts a 0.4-percentage-point higher chance of attaining a senior position (Appendix D) – attenuated but not eliminated – with the entrant's own pay explaining about half of the cross-firm correlation.

The cutting behavior in Section 4 also moves with a firm's retention of its trainees. At high-capability firms, hiring into the most learning-rich work is reduced by a greater amount when the pre-shock rate at which junior employees depart from these firms for employment elsewhere is higher. For example, if the poaching rate is greater than the median, the 2024 cut is -1.4% per standard deviation of learning content (s.e. $= 0.30$), compared to -0.05% (s.e. $= 0.29$) when poaching rates are less than the median. That the two groups (firms retaining trainees and firms losing them) show different learning-rich hiring patterns is consistent with the deployment margin of Proposition 1 (preservation of learning-rich work requires retention),

Table 8: Manager-pipeline features: training origin, tournament, and retention

Training: exposed origins and advancement	
Share who ever held exposed cognitive work	0.89
Share first observed in an exposed occupation	0.84
Share reaching the senior tier, exposed vs. unexposed entrants	0.067 vs. 0.028
Tournament: 2017 entry cohort, status at +7 years	
Exited / still junior / reached middle / reached senior	0.33 / 0.53 / 0.10 / 0.04
Entry salary, ever-promoted vs. not	\$69.7k vs. \$61.3k
Generality: where senior workers were trained	
Senior workers who were junior at a different firm	0.52
Junior annual rate of moving to another firm	0.074
Junior annual rate of internal promotion	0.033

Career features of the manager pipeline from worker histories, 2017–2024 (also Figure 3). The retention shares are left-censored lower bounds, since the panel does not observe career spells before 2017. The origin shares use each worker's first observed year; restricting to workers first observed as juniors gives 0.83.

but there are two reasons why I find this suggestive: first, the architecture choice is a cutoff, so the model does not sign comparisons across intermediate poaching levels; and second, the rate of employee poaching is endogenously determined by a firm's location, which is entangled with where AI-capable firms cluster. Defense-sector exposure, the one shifter of employee mobility whose geography is uncorrelated with AI capability, shows no impact on learning-rich hiring, so the magnitude of the effect cannot be identified (see Appendix F).

There are four ways to show that ψ cannot mechanically be constructed to predict advancement. First, the two indices separate along the model's lines – ν predicts which work is cut but adds nothing to who reaches senior management once ψ is included, while ψ predicts who is advanced everywhere and deepens the cut only where the model puts the architecture switch, at capable high-poaching firms (Section 4). Second, the construction of ψ includes no mechanism whereby it can predict future outcomes – ψ is constructed based solely on O*NET descriptions of what entry-level jobs entail, with no information provided about either Revelio or actual outcomes, and does not include the supervisory tasks required to manage (Section 3). Third, the effect persists regardless of controlling for behavioral AI usage, occurs at the individual entrant level, and remains robust to an instrument for first-job content (see Table 9). Fourth, learning content is positively correlated with the skills needed for an entry-level job (0.6–0.8 with cognitive complexity, education, and Job Zone; see Section 3). One demanding test is whether entry work that is more learning-rich predicts more managers, holding skills constant. The answer is yes: conditional on cognitive-complexity deciles, a standard deviation in learning content still predicts 1.2 percentage points higher senior attainment. Similarly,

Table 9: Learning content predicts manager formation, net of output exposure

	Within firm × cohort			Entrant	Entrant
Per SD of the index	(1)	(2)	(3)	(4) OLS	(5) IV
Panel A. Reaches the senior tier (base 5.0%*)*					
Learning content ψ	+0.0172		+0.0173	+0.0093	+0.0361
	(0.0021)		(0.0021)	(0.0023)	(0.0136)
Output exposure v (alone)	+0.0120				
	(0.0036)				
AI usage (AEI)			−0.0001		
			(0.0012)		
Panel B. Reaches senior management, SOC-11 (base 2.1%*)*					
Learning content ψ		+0.0074	+0.0076		
		(0.0015)	(0.0015)		
Output exposure v (net of ψ)		+0.0018			
		(0.0015)			
Output-exposure control	decile	decile	decile	linear	none
Fixed effects		firm + cohort		cohort, SOC2	cohort, SOC3

Linear probability models for whether a 2017–2019 entrant reaches the senior tier (Panel A) or senior management (Panel B). Columns 1–3: 16.1 million workers, firm and entry-cohort fixed effects, occupation-clustered standard errors, output exposure absorbed non-parametrically (decile fixed effects); column 3 adds the AI-usage control (orthogonalized to ψ). Columns 4–5: the entrant-level specification on 2018–2021 entrants (3,635,923 workers, senior-tier base 3.5%), with output exposure controlled, clustered on entry SOC2 (column 4) or SOC3 (column 5); column 5 instruments first-job learning content with entry-year occupation-field entry volume (first-stage $F \approx 7$: the instrument is weak, so column 5 is exploratory). Exposures are standardized. Output exposure does not predict reaching management net of learning content. Columns 1–3 observations 16,119,935 entrants; 757 occupation clusters.

holding occupation families constant, a standard deviation in learning content still predicts 1.1 percentage points higher senior attainment. Finally, holding the full skill bundle constant within high-complexity occupations, a standard deviation in learning content still predicts 1.2 percentage points higher senior attainment (see Appendix D, Table 25).

Table 10: Learning content predicts advancement at a different firm: general human capital

Coefficient on learning content ψ (per SD)	Senior tier	Senior mgmt.
Reaches the outcome at the entry firm (internal)	+0.0092	+0.0042
	(0.0010)	(0.0007)
Reaches the outcome at a different firm (external)	+0.0084	+0.0035
	(0.0011)	(0.0008)

Linear probability models for whether a 2017–2019 entrant reaches the senior tier or senior management, by where it is reached; output exposure absorbed non-parametrically (ν-decile fixed effects), entry-firm and entry-cohort fixed effects, occupation-clustered standard errors in parentheses, per standard deviation of ψ. Learning content predicts reaching a senior role at a *different* firm about as strongly as at the entry firm, so the content is portable. Observations 16,123,812 entrants; 757 occupation clusters.

6 The deferred effect on the manager pipeline

The contraction's effect on the manager stock will depend upon the rate at which the cut work generated new managers and the elasticity with which a lost cohort propagates into the manager stock. This section evaluates these variables, and then performs an accounting exercise to project the lost managerial cohorts.

6.1 The reduction of manager-making work

Because the selection criteria of the reduction include output exposure (Section 4), the impact of the reduction on managers will be influenced by whether or not the work selected by the reduction creates managers at a rate greater than the average. There is a correlation between output exposure and the manager-formation rate by occupation of +0.55, and the senior-tier managers created in the contracting exposed rung are formed at a rate of 1.33 times the overall average for entry work, while senior management is formed at a rate of 1.28 times the average; see Table 11. This finding compounds Section 4: the work selected by the exposed tercile is learning-rich on average (the two indices correlate at 0.47), and at the AI-capable firms that lose their trainees the contraction removes learning-rich work directly, so the reduction is biased towards removing managers through composition, and where the architecture switches, through selection as well. Based on the approximately 300,000 missing entrants in the 2023–24 cohorts – the actual exposed entry shortfall (see Appendix C) – if I apply the manager-formation rate of the cut work, I would expect approximately 20,000 fewer future senior-tier managers (and 8,000 fewer using the stricter definition of SOC-11), where approximately 5,000 and 1,700 of those represent an excess relative to a comparable-size neutral cut. This directionality

Table 11: The cut work forms managers at an above-average rate

	Senior tier	Senior mgmt. (SOC-11)
Formation rate, all entry work (R_{avg})	0.050	0.021
Formation rate, exposed (AI-cut) work (R_{exp})	0.067	0.027
Disproportionality R_{exp}/R_{avg}	1.33	1.28
bootstrap 95% interval	[1.22, 1.50]	[1.14, 1.45]
corr(output exposure v, formation rate)	0.55	0.38
Missing future managers (exposed rate × 300k)	≈20,000	≈8,000
excess over a size-equivalent neutral cut	≈5,000	≈1,700

Occupation-level manager-formation rates among 2017–2019 entrants (share ever reaching the senior tier, and senior management) over the 666 occupations with at least 300 entrants; "exposed" is top-tercile output exposure. The realized exposed entry shortfall over 2023–24 is about 300,000 entrants (Appendix C). Disproportionality intervals are 2,000-draw occupation bootstraps. The raw occupation-level decline is not manager-biased ($R_{decl}/R_{avg} = 0.91$); the tilt operates through the AI-attributed exposed contraction.

is unique to the AI-attributed cut – the overall occupational decline, largely due to non-AI movement, is not manager-biased ($R_{decl}/R_{avg} = 0.91$). Thus, the bias exists specifically in the within-firm contraction that Section 4 isolates.

6.2 The delayed consequence

Missing cohorts do not immediately propagate into the manager stock. As shown by Proposition 2, a new entrant must first take years to develop enough experience before becoming a manager. I estimate the propagation directly and then project it forward. Cohorts that are thinner produce fewer future managers. Using data at the occupation-by-entry-cohort level, the elasticity is 0.67 for senior-tier managers (standard error = 0.08) and 0.50 for senior management (standard error = 0.05); see Table 12. The elasticity is below one because of congestion relief: when a cohort is thinner than normal, a larger portion of its surviving members are promoted (per-entrant senior yield falls by 0.33 log points per log point of cohort size, standard error = 0.08), thereby allowing some portion of a missing cohort to be absorbed into the system through this increased promotion rate instead of being passed through entirely. In the terms of the model, this implies a congestion elasticity of $\eta_c = -0.33$. The log-log elasticity omits the occupation-cohort cells that produce no senior manager; retaining them with an unrestricted Poisson count model, the estimand used for the firm-level passthrough of Section 6.3, raises the elasticity to 0.82 (senior tier) and 0.73 (SOC-11), still below one.

By feeding the cut through a stock-flow accounting framework that is calibrated to reflect

Table 12: Cohort-to-manager elasticity: a thinner cohort makes fewer managers

	Coef.	(s.e.)
Elasticity of future senior-tier managers to cohort size	+0.67	(0.08)
Elasticity of future senior management (SOC-11) to cohort size	+0.50	(0.05)
Congestion: per-entrant senior yield, per log cohort size	−0.33	(0.08)
Unrestricted Poisson, senior tier	+0.82	(0.04)
Unrestricted Poisson, senior management (SOC-11)	+0.73	(0.05)
2008–09 finance collapse: implied cohort-to-senior elasticity	0.73	

Occupation-by-entry-cohort cells (2018–2021 cohorts, fixed three-year attainment window), log future managers on log cohort size, with occupation and cohort-year fixed effects, occupation-clustered standard errors. The senior-tier elasticity is one plus the congestion term. The log-log rows drop the 14% of occupation-cohort cells that produce no senior-tier manager; the unrestricted-Poisson rows keep them and are robust to that selection (the count-model estimand of Table 15). Every convention puts the elasticity below one. The last row is the implied elasticity from the 2008–09 natural experiment (Section 6.3): a 23% entry shortfall produced a 17% senior shortfall. Observations 2,523 occupation-by-cohort cells (3,291 including the panel-start cohort); 684 occupation clusters (726).

the pre-AI transition hazards in the pipeline (Appendix E), I trace how a reduced entry rung eventually reaches the manager stock.[12] Figure 4 illustrates the shape, which is the substantive part of the finding. Initially, deficits in the entry stock grow rapidly while deficits in the manager stock remain relatively small for many years, before declining and ultimately reaching their proportionate cuts well beyond an individual's working life. The magnitudes are illustrative, since they are not direct estimates of future managerial supply. If I feed only the actually realized deficits from 2023–24 through the accounting framework and assume hiring returns to trend beginning in 2026, I obtain a peak value of 1.7% in 2037 under this scenario. Under an alternative permanent cut at the observed rate – the 2024 detrended hire deficit of 14.7%, rounded to 15% – the strict-management (SOC-11) manager stock experiences a reduction of approximately 13% over a twenty-year span, while the broader definition based on seniority is more conservative at 8.5%, with a longer half-life (Table 13; both definitions and their sensitivity to changes in the size of the cut are presented in Appendix E, Table 28).

Even though there is little ambiguity about how an eventual reduction will occur, there is wide uncertainty about how large it will be. This uncertainty largely stems from the reduced-form estimate of the contraction, as the transition hazard estimates have tight confidence in-

[12]The accounting holds the hazards fixed rather than having them change as a result of previous shocks. Because of the congestion margin (Proposition 2), part of every shortfall gets absorbed while in transit, so the fixed-hazard path is an upper bound on the transition deficit for any given period of time; the long-run passthrough is equivalent either way. The accounting also applies pre-shock formation rates to the entrants who survive the cut; if substituting firms erode formation ($\lambda > 0$, Section 2.3), the realized deficit gains a formation channel this headcount accounting omits, so the projection is conservative in that respect too.

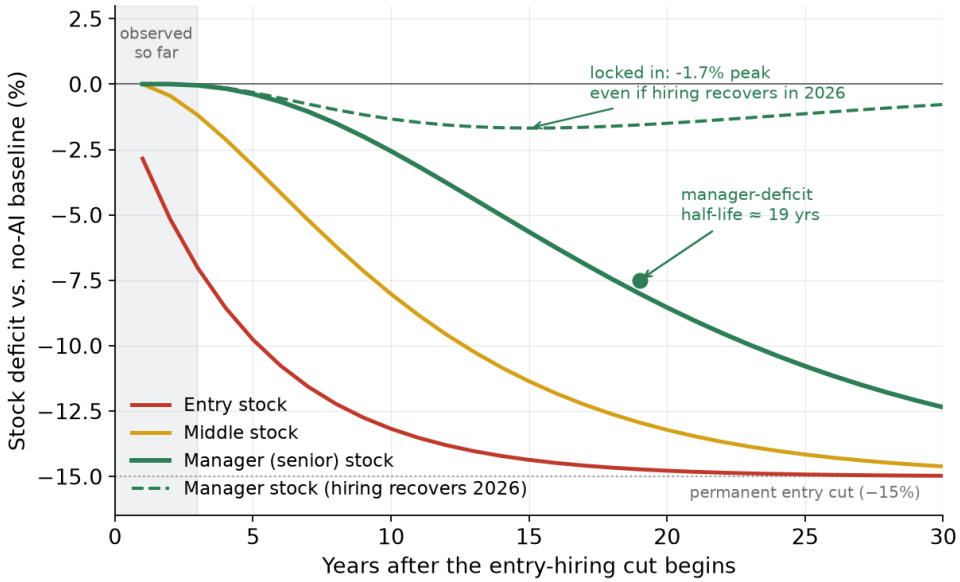


Figure 4: Projected manager-stock deficit over time, after a permanent 15% entry-hiring cut.

Stock-flow chain (3) calibrated to within-firm transition hazards (2018–2021), tracing the entry, middle, and manager (senior) stocks against a no-cut baseline. The shaded region marks the three post-shock years observed so far. The dashed path feeds the accounting only the realized 2023–24 entry shortfalls and returns hiring to trend in 2026; this locked-in deficit still peaks at 1.7% in 2037.

tervals – adding those sampling variances increases the width of the projected interval by 0.03 percentage points. The estimates of hazards are also made within-firm so that the projection can track the manager stock produced by the firm. When I recalculate this using the worker's career-wide survival (in which a worker is assigned to her tier at time t regardless of whether or not she has moved firms since then), seniors then turn over more slowly and the deficit is 3.8% relative to the baseline after twenty years; the long-run passthrough is 15% either way (see Table 28).

The contraction has not yet made its way to the senior tier. Within firms, the promotion rate into senior positions remained flat through 2024 while the senior headcount changes very little (Appendix E), as firms sustain the same number of seniors by promoting from within, drawing on a declining base of entry-level employees. If firms were reducing the base of entry-level employees due to expectations that AI will erase the demand for managers then the senior tier should also be shrinking, but this is not currently the case.

This projection represents the public-firm sector that the panel tracks. The missing entrants may be being trained outside of the panel's coverage – either at privately held firms, in the government, in self-employment or in classrooms – in which case the public-firm shortage would be an overstatement of the overall economy-wide shortage. Within the panel, however,

Table 13: Projected manager-stock deficit after a permanent 15% cut to entry hiring

Years after the cut	5	10	15	20
Entry stock	−9.8%	−13.2%	−14.4%	−14.8%
Middle stock	−3.1%	−8.0%	−11.4%	−13.2%
Manager (senior) stock	−0.4%	−2.6%	−5.6%	−8.5%
Manager stock, locked-in component	−0.3%	−1.3%	−1.7%	−1.5%

Stock-flow chain (3) calibrated to within-firm transition hazards (2018–2021): junior promotion 1.8%/yr, junior exit 17%/yr, manager separation 11%/yr. Manager-deficit half-life (time to half the eventual permanent effect) ≈ 19 years. The locked-in row feeds the accounting only the realized 2023–24 entry shortfalls (with 2025 held at the 2024 rate) and returns hiring to trend in 2026.

the missing entrants are not being reabsorbed: gross exposed-junior employment decreased in each of the three capability groups, and none of the contraction reappeared at the low-capability firms (Appendix E).

6.3 The 2008–09 precedent

I test the structural premise of the projection – that the hazards of the pipeline can only respond to a missing cohort on the congestion margin of Prediction 4 and not sufficiently to offset it – using the 2008–09 financial crisis, whose entry cohorts have already passed through the promotion window this paper needs to observe.[13] I track every individual whose first career position was a U.S. entry-level job as a financial analyst, management analyst, or accountant (578,000 total entries over 1998–2014) and find that the financial crisis reduced entry hiring in finance sharply – creating a 2009 cohort that was 23% below the trend set by its non-crisis neighbors. The promotion rates of this cohort were slightly higher than those of neighboring cohorts: the cohort entered the mid-tier within seven years at a rate of 53.1% compared to 50.9% (trend), and entered the senior tier within ten years at a rate of 33.7% compared to 31.2% (Table 14). Therefore, the seniors produced by this cohort were 17% less than what would have otherwise existed, approximately three-fourths of the shortfall in its entrants, with an implied cohort-to-senior elasticity of 0.73 – close to my estimate within the panel – with evidence of congestion relief in the higher promotion rates. Firms did not change the promotion bar, and accepted fewer managers instead of lowering quality. A transition matrix estimated based only on transitions prior to the crisis reproduced the actual path of senior production

[13]The recession-cohort literature indicates significant permanent earnings effects of the recession on those workers (Kahn, 2010; Oreopoulos et al., 2012). However, my interest lies in how many seniors were created from the workers who would have been entering during that time.

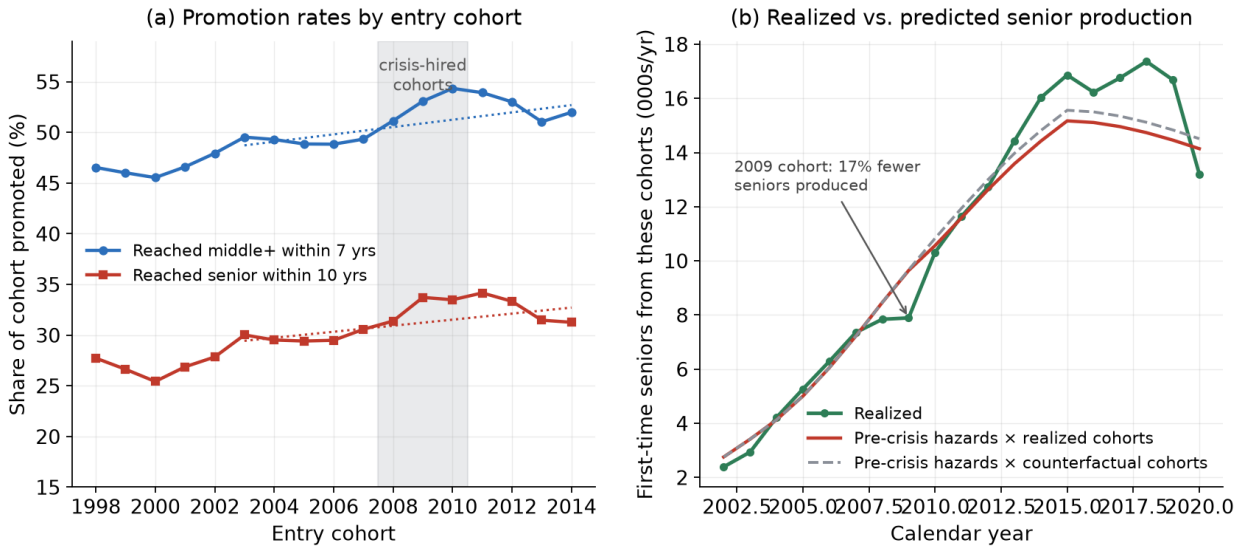


Figure 5: Promotion rates and senior-tier production around the 2008–09 entry collapse.

First-career U.S. entrants in finance and consulting occupations, cohorts 1998–2014, followed through 2024 (577,719 workers; Appendix E). Left: the share of each cohort reaching the middle tier within seven years and the senior tier within ten; dotted lines fit the non-crisis cohorts; the 2009 cohort entered 23% below trend. Right: first-time seniors produced per calendar year, against a pre-crisis transition matrix applied to realized and to counterfactual cohort sizes; prediction and realization correlate 0.97 over 2008–2020.

given the realized cohort sizes, with a correlation of 0.97 over 2008–2020 (Figure 5). The eventual recovery limited the damage, since finance entry returned to trend levels within two years. So the missed seniors of 2010–19 represent only 2.6% of what would have otherwise existed. In contrast, the AI contraction is in its third year with no sign of such a recovery yet (Figure 2).

In addition to implementing this exercise for three types of finance professionals, I extend it to all publicly traded companies. Using the raw position records, I build a firm-based panel of cohorts (6.31 million U.S. entrants, 2005–2016), and then determine if the firms that reduced the number of workers who entered during the Great Recession produced fewer of their *own* managerial personnel ten years later. Within a given firm and occupation, reducing an entry cohort reduces the number of internal managers. When I instrument the cohort with the "leave-one-out" industry-by-cohort entry shock, which is the part of an entry cohort that was influenced by the recession in the firm's industry rather than its own decisions, the quasi-causal passthrough is 0.40 (first-stage $F \approx 441$, Table 15). The estimated effect at the firm level is 0.83, about twice the instrumented effect, and the descriptive effect within firm and occupation is 0.93, both inflated by firm-specific cutting, as a significant endogeneity test confirms. Thus, when a firm reduces its entry cohort by 30 percent, it has approximately a 12

Table 14: Validation-sample cohorts around the 2008–09 entry collapse

Cohort	Size	Middle+ by +7	Senior by +10	Salary gap, promoted − never
2007	46,451	49.3%	30.6%	$4,065
2008	43,915	51.1%	31.4%	$3,245
2009 (crisis trough)	33,397	53.1%	33.7%	$2,407
2010	39,109	54.3%	33.5%	$2,981
2011	40,621	53.9%	34.2%	$3,218

First-career finance and consulting entrants by cohort year (full table in Appendix E). The crisis cohort (2009) entered 23% below the neighbor trend but was promoted at *higher* rates, so the seniors it produced fell about 17%. The salary gap is the mean entry salary of workers reaching middle+ within ten years minus that of workers never reaching it.

percent shortfall in internal managers, and the cut hurts most in the occupations that make managers.

Firms absorb the rest of the reduction on the congestion margin already built into the projection, allowing a larger portion of the survivors from a thinner cohort to advance. But they do not do so any sooner: under the same instrument, a 30 percent reduction in an entry cohort increases a worker's time to seniority inside the company by approximately two months (−0.54 years per log cohort size, s.e. = 0.12, Table 15). A smaller share of a cohort's eventual managers reach that status within five years, so the promotions that replace a missing cohort come later. Nor do firms substitute for the lost cohort by purchasing additional experience from outside the company: a larger cut produces no within-firm shift toward hiring senior workers from outside the firm (+0.0005, s.e. = 0.003). The deferred deficit of the projection is thus already visible in the firms that experienced a cohort cut – determined by the industry's shock rather than the firm's own decision-making – and neither the survivors' promotions nor outside hiring undid it.

The results at the firm level are distinct from the occupation-level elasticity of 0.67 that drives the projection: mobility across firms, together with the higher promotion rates of the survivors, makes a firm's actual realized passthrough less than the economy-wide figure, in which the movement of workers between firms washes out. Therefore, the firm-level estimates identify how firms produce their own managers, whereas the occupation-level elasticity measures the overall aggregate impact.

Table 15: The realized firm-level consequence: thinner entry cohorts make fewer internal managers

	Coef.	(s.e.)
Make-side: realized internal managers (10 years) on entry-cohort size, within firm		
Within firm×occupation, Poisson (descriptive)	+0.93	(0.011)
Within firm×cohort, Poisson (the uninstrumented benchmark)	+0.83	(0.020)
Shift-share IV-Poisson (quasi-causal)	**+0.40**	**(0.088)**
first stage (leave-one-out NAICS3×cohort), F	441	
endogeneity test (first-stage residual)	+0.48	(0.093)
stronger in manager-making occupations ($\times R_{\text{sen}}$)	+0.032	(0.009)
Timing: a survivor's years to internal seniority on entry-cohort size, within firm		
shift-share IV (same instrument)	−0.54	(0.12)
Buy-side: do firms with bigger cuts hire more external seniors later?		
within-firm Δ(external share of senior arrivals) on 2009 cut	+0.0005	(0.003)

Firm-resolved cohort panel from raw Revelio (6.31M U.S. entrants; the 2008–09 window covers ∼13–15k firms). The make-side rows are estimated on the 2005–2014 entry cohorts, whose ten-year manager outcomes the panel observes in full; the instrumented rows additionally retain only industry-by-cohort cells with at least five other firms. Counts are Poisson, since a log-count specification attenuates spuriously (Santos Silva and Tenreyro, 2006). The first two rows are different estimands: the 0.93 holds firm and occupation fixed, the 0.83 holds firm and cohort fixed, and the instrument deflates the latter. The make-side instruments the firm's entry-cohort size with the leave-one-out NAICS3-by-cohort entry total; the control-function IV-Poisson deflates the uninstrumented firm-level passthrough about twofold, with a significant endogeneity test, so a ∼30% cohort cut implies a ∼12% shortfall in a firm's internal managers. Because the control function enters an estimated first-stage residual, the IV-Poisson and endogeneity-test standard errors are bootstrapped over firms, re-estimating both steps in each replication. The exclusion restriction (that the industry-by-cohort shock reaches a firm's later managers only through its own cohort) makes this quasi-causal. The buy-side row is the within-firm change in the external share of senior arrivals on the firm's 2009 manager-making cut: no make-to-buy substitution. The timing row shows the buffer's form: a thinner cohort's survivors are promoted at higher rates but later (a 30% cut adds about two months to the path into internal seniority, estimated among workers reaching it within ten years; first-stage $F \approx 171$ in this subsample), so the congestion relief operates through more promotions, not quicker ones. Observations 243,666 firm-by-occupation-by-cohort cells (descriptive row) and 114,995 firm-by-cohort cells (instrumented rows); 4,733 and 22,345 firm clusters.

7 The market failure and its implications

Why don't firms protect the work that generates their managers? Private and social values of this manager-making work diverge because the training is general – and AI is widening the gap. The wedge exists only if what entry work teaches is general rather than firm-specific. This is because a firm receives all of the management rents generated by skills that its workers cannot take elsewhere. Section 5 demonstrated that the training provided by entry work is general. As a result, the share ϕ of trainees that a firm loses before collecting the management rent sits well above zero.

With $\phi > 0$ there is divergence between the private and social returns to doing the work – the classic under-investment in general training (Becker, 1964; Acemoglu and Pischke, 1998, 1999). A firm hires to fill positions for the rents expected to remain, $\pi_q(1-\phi)\beta V^M$, while society values the full rent $\pi_q \beta V^M$. Thus, conditional on the deployment architecture, private hiring standards differ from social ones by the marginal wedge $\phi\beta V^M \pi_q \geq 0$, strictly positive wherever $\pi_q > 0$, and a firm that has switched to the substitution-intensive architecture while the planner would preserve falls short by more. Therefore, equilibrium entry hiring and the steady-state manager stock lie weakly below the planner's optimum, strictly under the conditions of Proposition A1 (Appendix A). Firms under-produce manager-making entry work, and under-preserve its formative content, precisely because part of the value created by this type of work leaves with the employee.

A rise in the displacement Δ reduces the output buffer, $\omega - \Delta$, that carries a role over the hiring bar, and pushes the marginal roles below it. These marginal roles are the learning-rich, rent-bearing ones that the firm already discounts by ϕ; therefore these roles sit closer to the firm's bar than to the planner's. Within each exposed occupation, and over the range in which preserving training remains socially efficient, the percentage decline in equilibrium hires relative to the planner's increases with the displacement (Proposition A1). The same shock that reduces entry hiring and disproportionately cuts the manager-making rung (Sections 4 and 6) therefore removes the most manager-making work precisely where the private return understates the social return by the most. The pipeline corrects only part of the reduction on its own (Section 6), the missing entrants are not reabsorbed by low-capability firms (Appendix E), and a role that is left unstaffed trains nobody. Thus, the deficit is a loss, rather than a redistribution.

The wedge also defines the form of corrective action needed: a subsidy of $s_q^* = \phi\beta V^M \pi_q$ per entry hire moves both hiring and the deployment choice towards the planner's choice (Proposition A1). Because the wedge scales with the training content that a firm cannot capture, the subsidy must apply to learning-rich entry roles whose formative content survives, and

not to entry hiring at large. A planner could instead close the wedge on the retention side, by limiting the mobility that drains the rents earned by firms providing training; but that mobility is also how trained managers reach the firms that value them, so the hiring margin is the safer instrument. Apprenticeship levies and training credits would fund formal instruction, but the wedge here relates to the hiring of work where the *doing* is itself the training.

8 Conclusion

This paper has presented evidence that the impact of AI on the future availability of managers operates through an attribute of entry-level work which the current indices of exposure to automation do not capture. Entry-level work produces both output and the development of judgment from which future managers are produced. Generative AI directly displaces the output component of entry work, and its deployment architecture decides whether it erodes the formative component as well. The two effects are governed by two distinct attributes of the same work. The first attribute is the amount of a job's output that is exposed to automation, or output exposure, which governs the displacement component of job loss. The second attribute is the content in terms of learning involved in performing the tasks, or learning content. I build an index of the latter attribute using O*NET descriptions of work activities.

I show how learning content moderates the effect of AI on the supply of managers in a sample of approximately 9,000 publicly traded companies in the U.S. For entry jobs that are highly exposed to automation, firms reduce headcounts by 3.6 percent relative to non-exposed jobs during the period immediately following the launch of ChatGPT. Firms with higher capabilities to implement AI experience reductions of 8.1 percent, whereas firms with lower capabilities experience no such reduction. Within the rung of entry-level positions, learning content is a significant predictor of upward mobility. Increasing learning content by one standard deviation increases the probability of being promoted to senior management by about a third of the baseline rate, regardless of whether the worker is employed at the entry firm or at some other firm. Conversely, output exposure does not predict reaching senior management once learning content is controlled.

The exposed occupations, where the contraction is concentrated, produce managers at 1.33 times the average rate. A contraction concentrated in this work therefore thins the future supply of managers, not only today's labor input. But this thinning occurs with a delay, since a missing cohort reaches the manager stock only as senior slots open, and with an elasticity of about two-thirds, because thinner cohorts promote a larger share of their surviving members.

To understand how these dynamics play out, I study the effect of the 2008–09 financial crisis on finance entry jobs. Firms reduced hiring 23 percent below the trend set by neighboring non-crisis cohorts, the 2009 cohort had 17 percent fewer seniors over the following decade, its survivors were promoted at slightly higher rates, and firms did not replace the missing seniors by hiring externally. Applying a similar accounting exercise to the AI contraction, the entrants already missing from the 2023–24 cohorts lock in a manager-stock deficit of 1 to 2 percent, and a contraction that persists at its current rate takes the deficit to roughly 8.5 percent over two decades.

These losses are privately rational and, at the same time, socially inefficient. Training received via entry-level positions is general: more than half of the senior employees whose entry position I observe were trained by another employer. Therefore, a portion of the management rent generated by hiring an employee leaves with them. Thus, firms under-provide the training that creates managers. Generative AI exacerbates this problem by reducing the output component that previously funded the training. Among AI-capable firms, hiring into learning-rich work is lower by 1.4 percent per standard deviation of learning content where trained juniors are routinely poached, and not at all where firms retain their trainees. Thus, the firms that strip the formative content of entry work are the ones that cannot keep the workers that they train. I identify a subsidy to hiring entry-level employees engaged in learning-rich jobs as the correct remedy for this under-provision.

There are several possible avenues for further research. First, I restrict attention to public firms. Missing entrants may receive training elsewhere – at private firms, in government, or in the classroom. This prompts the question whether this contraction is an economy-wide phenomenon or merely one limited to public firms. Answering this question requires access to career-long employment histories extending beyond public-firm employment. Second, identifying the price of under-providing training, and hence determining the rate at which a subsidy should be provided, requires some form of variation in worker mobility that is orthogonal to geographic differences in AI use. Third, while I measure what entry-level jobs teach using the learning-content index, the deployment architecture itself is never observed, and direct evidence on what the surviving work now teaches, at preserving and at substituting firms alike, remains open.

References

Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*, volume 4B, pages 1043–1171. Elsevier, 2011.

Daron Acemoglu and Jörn-Steffen Pischke. Why do firms train? theory and evidence. *Quarterly Journal of Economics*, 113(1):79–119, 1998.

Daron Acemoglu and Jörn-Steffen Pischke. The structure of wages and investment in general training. *Journal of Political Economy*, 107(3):539–572, 1999.

Daron Acemoglu and Pascual Restrepo. Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30, 2019.

Daron Acemoglu and Pascual Restrepo. Tasks, automation, and the rise in U.S. wage inequality. *Econometrica*, 90(5):1973–2016, 2022.

Hassan Afrouzi, Andres Blanco, Andrés Drenik, and Erik Hurst. Automation, learning, and career dynamics. Technical report, Becker Friedman Institute Working Paper 2026-61, 2026.

Philippe Aghion and Jean Tirole. Formal and real authority in organizations. *Journal of Political Economy*, 105(1):1–29, 1997.

Anthropic. The Anthropic Economic Index. Anthropic, 2025. Task-level measures of AI usage matched to O*NET tasks.

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

David H. Autor. Why are there still so many jobs? the history and future of workplace automation. *Journal of Economic Perspectives*, 29(3):3–30, 2015.

David H. Autor, Frank Levy, and Richard J. Murnane. The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333, 2003.

José Azar, Mireia Giné, and Javier Sanz-Espín. AI is already eroding wages: Quasi-experimental evidence from occupational exposure. Technical report, SSRN Working Paper 5842084, 2025.

Tania Babina, Anastassia Fedyk, Alex Xi He, and James Hodson. Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151:103745, 2024.

Mark Bagnoli and Ted Bergstrom. Log-concave probability and its applications. *Economic Theory*, 26(2):445–469, 2005.

George Baker, Michael Gibbs, and Bengt Holmström. The internal economics of the firm: Evidence from personnel data. *Quarterly Journal of Economics*, 109(4):881–919, 1994.

Jay B. Barney. Firm resources and sustained competitive advantage. *Journal of Management*, 17(1):99–120, 1991.

Gary S. Becker. *Human Capital: A Theoretical and Empirical Analysis*. Columbia University Press, 1964.

Nicholas Bloom and John Van Reenen. Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics*, 122(4):1351–1408, 2007.

Kirill Borusyak, Peter Hull, and Xavier Jaravel. Quasi-experimental shift-share research designs. *Review of Economic Studies*, 89(1):181–213, 2022.

Kirill Borusyak, Xavier Jaravel, and Jann Spiess. Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6):3253–3285, 2024.

Erik Brynjolfsson, Danielle Li, and Lindsey R. Raymond. Generative AI at work. Working Paper 31161, National Bureau of Economic Research, 2023.

Erik Brynjolfsson, Bharat Chandar, and Ruyu Chen. Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence. Technical report, Stanford Institute for Economic Policy Research Working Paper, 2025.

Benjamin A. Campbell, Russell Coff, and David Kryscynski. Rethinking sustained competitive advantage from human capital. *Academy of Management Review*, 37(3):376–395, 2012.

Peter Cappelli. *Talent on Demand: Managing Talent in an Age of Uncertainty*. Harvard Business Press, 2008.

David G. Collings and Kamel Mellahi. Strategic talent management: A review and research agenda. *Human Resource Management Review*, 19(4):304–313, 2009.

David J. Deming. The growing importance of decision-making on the job. NBER Working Paper 28733, National Bureau of Economic Research, 2021.

Ingemar Dierickx and Karel Cool. Asset stock accumulation and sustainability of competitive advantage. *Management Science*, 35(12):1504–1511, 1989.

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

Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. GPTs are GPTs: Labor market impact potential of large language models. *Science*, 384(6702):1306–1308, 2024.

Edward Felten, Manav Raj, and Robert Seamans. Occupational, industry, and geographic exposure to artificial intelligence. *Strategic Management Journal*, 42(12):2195–2217, 2021.

Morgan R. Frank, Alireza Javadian Sabet, Lisa Simon, Sarah H. Bana, and Renzhe Yu. AI-exposed jobs deteriorated before ChatGPT. arXiv preprint arXiv:2601.02554, 2026.

Guido Friebel, Yao Huang, Jin Li, Soumitra Shukla, and Andrew Zhang. Pyramids, diamonds, and oscillations: AI and the structure of internal labor markets. Technical report, Working Paper, 2026.

Luis Garicano. Hierarchies and the organization of knowledge in production. *Journal of Political Economy*, 108(5):874–904, 2000.

Luis Garicano and Luis Rayo. Training in the age of AI: A theory of apprenticeship viability. Discussion Paper 20634, Centre for Economic Policy Research, 2025.

Luis Garicano and Esteban Rossi-Hansberg. Organization and inequality in a knowledge economy. *Quarterly Journal of Economics*, 121(4):1383–1435, 2006.

Christina Gathmann and Uta Schönberg. How general is human capital? a task-based approach. *Journal of Labor Economics*, 28(1):1–49, 2010.

Robert Gibbons and Michael Waldman. A theory of wage and promotion dynamics inside firms. *Quarterly Journal of Economics*, 114(4):1321–1358, 1999.

Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift. Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624, 2020.

Nile W. Hatch and Jeffrey H. Dyer. Human capital and learning as a source of sustainable competitive advantage. *Strategic Management Journal*, 25(12):1155–1178, 2004.

Seyed Mahdi Hosseini Maasoum and Guy Lichtinger. Generative AI as seniority-biased technological change: Evidence from U.S. résumé and job posting data. Technical report, SSRN Working Paper 5425555, 2025.

Anders Humlum and Emilie Vestergaard. Still waters, rapid currents: Early labor market transformation under generative AI. Working Paper 33777, National Bureau of Economic Research, 2025.

Enrique Ide. Automation, AI, and the intergenerational transmission of knowledge. Technical report, IESE Business School Working Paper, 2025.

Boyan Jovanovic. Job matching and the theory of turnover. *Journal of Political Economy*, 87 (5):972–990, 1979.

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

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

Peter John Lambert and Yannick Schindler. The broken ladder: AI, remote work, and early-career hiring. Technical report, SSRN Working Paper No. 6787638, 2026.

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

Edward P. Lazear and Sherwin Rosen. Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, 89(5):841–864, 1981.

Edward P. Lazear, Kathryn L. Shaw, and Christopher T. Stanton. The value of bosses. *Journal of Labor Economics*, 33(4):823–862, 2015.

Edwin Leuven. The economics of private sector training: A survey of the literature. *Journal of Economic Surveys*, 19(1):91–111, 2005.

Mark A. Loewenstein and James R. Spletzer. General and specific training: Evidence and implications. *Journal of Human Resources*, 34(4):710–733, 1999.

Jonathan Meer and Jeremy West. Effects of the minimum wage on employment dynamics. *Journal of Human Resources*, 51(2):500–522, 2016.

Samuel Muehlemann. Artificial intelligence adoption and workplace training. *Journal of Economic Behavior & Organization*, 238:107206, 2025.

Shakked Noy and Whitney Zhang. Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192, 2023.

Philip Oreopoulos, Till von Wachter, and Andrew Heisz. The short- and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, 4(1):1–29, 2012.

Robert E. Ployhart and Thomas P. Moliterno. Emergence of the human capital resource: A multilevel model. *Academy of Management Review*, 36(1):127–150, 2011.

Robert E. Ployhart, Anthony J. Nyberg, Greg Reilly, and Mark A. Maltarich. Human capital is dead; long live human capital resources! *Journal of Management*, 40(2):371–398, 2014.

Ashesh Rambachan and Jonathan Roth. A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–2591, 2023.

Jonathan Roth. Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–322, 2022.

J. M. C. Santos Silva and Silvana Tenreyro. The log of gravity. *Review of Economics and Statistics*, 88(4):641–658, 2006.

Michael Webb. The impact of artificial intelligence on the labor market. Working paper, Stanford University, 2020.

Miao Ben Zhang. The broken job ladder: AI, vanishing apprenticeships, and the role of universities. Working paper, University of Southern California, Marshall School of Business, 2026.

A Derivations

Value of a manager. A retained manager yields the per-period quasi-rent $\rho = G_M - w^M$ and survives to the next period with probability $1-\delta$. Recursively, $V^M = \rho + \beta(1-\delta)V^M$, hence $V^M = \rho/[1-\beta(1-\delta)]$.

The hiring equation. A role in cell (f,o) is staffed iff its output value clears the threshold (1), holding the deployment architecture q_f fixed, so the hiring flow $H_{fo,t} = \mu_{fo} S(\omega^*_{fo,t})$ is strictly decreasing in the displacement, with $-\partial \ln H/\partial \Delta = \eta(\omega^*)$, strictly increasing in the formation probability, and decreasing in poaching. In general equilibrium V^M is endogenous: a smaller manager stock raises managers' marginal product and the rent, which lowers the threshold and partially offsets the contraction, but does not reverse it as long as the direct effect of the displacement dominates the induced rise in V^M. For the equation, write $\log H_{fo,t} = \log \mu_{fo} + \log S(\omega^*_{fo,t})$ and expand around the cell's pre-release threshold $\bar{\omega}_{fo}$: $\log S(\omega^*) \approx \log S(\bar{\omega}_{fo}) - \eta_{fo}(\omega^* - \bar{\omega}_{fo})$ with $\eta_{fo} \equiv \eta(\bar{\omega}_{fo})$, the higher-order remainder entering $u_{fo,t}$. Substituting the threshold (1), $\omega^*_{fo,t} - \bar{\omega}_{fo} = (w^J_t - \bar{w}^J) + \theta_t a_f \nu_o - \beta(1-\phi_f)\big[\pi_{q_f}(\psi_o) V^M_t - \pi(\psi_o)\bar{V}^M\big]$, where the pre-release baseline carries $\pi_{q_f} = \pi$; the bracketed term collects both movement in the manager value and the architecture-induced rent loss. Under a common local hazard across the firm's entry occupations ($\eta_{fo} = \eta_f$), the wage movement loads on the firm-by-year effect, the baseline levels load on the cell effect, the manager-value movement is carried by the rent term with its occupation-varying weight $\pi(\psi_o)$, and equation (2) follows. Without the common hazard, the coefficient on the displacement is an η_{fo}-weighted average, the exposure-decile-by-period effects absorb the dependence of η on ν, and identification of the learning-content coefficient requires the within-decile residual hazard-weighted displacement $\eta_{fo}\Delta_{fo,t}$ to be uncorrelated with ψ.

The fixed-architecture benchmark. Fix a cell, hold the architecture at A, let $x \equiv w^J - \beta V^M(1-\phi_f)\pi(\psi_o)$ and $d \equiv \bar{\theta} a_f \nu_o > 0$, so the post-minus-pre change in log hiring is $\log S(x+d) - \log S(x)$. Differentiating in ψ_o,

$$\frac{\partial}{\partial \psi_o}\big[\log S(x+d) - \log S(x)\big] = \beta V^M(1-\phi_f)\,\pi'(\psi_o)\big[\eta(x+d) - \eta(x)\big] \geq 0$$

under log-concavity of S (Bagnoli and Bergstrom, 2005), strictly positive where the hazard strictly rises on $[x, x+d]$, and vanishing in the limit as ϕ_f approaches one. Under a fixed architecture, then, the release weakly *protects* learning-rich work, exactly zero under a constant

hazard, and the protection requires retention. This benchmark is what the deployment margin overturns: a switch to the substitution-intensive architecture adds the threshold shift $r_f \lambda \pi(\psi_o)$, whose implications are collected in Proposition 1.

Proof of Proposition 1. Since $T^S_{fo} = T^A_{fo} + r_f \lambda \pi(\psi_o)$ with $\partial T^A_{fo}/\partial r_f = -\pi(\psi_o)$ and $\partial T^S_{fo}/\partial r_f = -(1-\lambda)\pi(\psi_o)$,

$$\frac{\partial B^F}{\partial r_f} = \sum_o \mu_{fo}\, \pi(\psi_o)\big[S(T^A_{fo}) - (1-\lambda) S(T^S_{fo})\big] > 0,$$

because $S(T^A_{fo}) > S(T^S_{fo})$ and $1-\lambda \leq 1$. Both thresholds rise by v_o per unit of D_f, so

$$\frac{\partial B^F}{\partial D_f} = \sum_o \mu_{fo}\, v_o \big[S(T^S_{fo}) - S(T^A_{fo})\big] < 0$$

wherever $v_o > 0$ on a positive measure of roles, and $\partial(B^F - K)/\partial D_f < 0$ follows from $K' \geq 0$. $B^F - K$ is continuous and strictly increasing in r_f, so for each D the preserving set $\{r_f : B^F(r_f, D) \geq K(D)\}$ is an upper interval, possibly empty or the whole admissible range; wherever the cutoff is interior it solves $B^F(r^*, D) = K(D)$ and is weakly increasing, $dr^*/dD = (K' - B^F_D)/B^F_r \geq 0$. At $D_f = 0$, $K = 0$ and $B^F \geq 0$, so every firm preserves. For the slopes, $g_q(\psi_o) = \log S(T^q_{fo}) - \log S(x_o)$ with $x_o = w^J - r_f \pi(\psi_o)$; differentiating, $\partial g_S/\partial \psi_o = r_f \pi'(\psi_o)\big[(1-\lambda)\eta(T^S_{fo}) - \eta(x_o)\big]$ and $\partial g_A/\partial \psi_o = r_f \pi'(\psi_o)\big[\eta(T^A_{fo}) - \eta(x_o)\big]$, the latter nonnegative under the increasing hazard since $T^A_{fo} = x_o + D_f v_o \geq x_o$. The sign condition on λ follows by rearrangement, and under a constant hazard η the two slopes reduce to $-\eta\, r_f \lambda\, \pi'(\psi_o) < 0$ and 0. □

Prediction 3. Under a constant hazard η_f, the slopes above reduce to $\partial g_S/\partial \psi_o = -\eta_f r_f \lambda \pi'(\psi_o)$ and $\partial g_A/\partial \psi_o = 0$; dividing the switcher's slope by the retention rate gives $-(1-\phi_f)^{-1}\, \partial g_S/\partial \psi_o = \eta_f \beta V^M \lambda \pi'(\psi_o)$, free of ϕ_f, against zero at a preserving firm. For the sorting, the switching set is $\{(r_f, D_f) : B^F(r_f, D_f) < K(D_f)\}$; B^F is strictly increasing in r_f and $B^F - K$ strictly decreasing in D_f (Proposition 1), and $r_f = (1-\phi_f)\beta V^M$ is decreasing in ϕ_f, so the set is an upper set in (D_f, ϕ_f) with boundary $\phi^*(D) = 1 - r^*(D)/\beta V^M$ weakly decreasing in D. For interiority, fix $D > 0$. As $\phi_f \to 1$, $r_f \to 0$ and $B^F \to 0 < K(D)$, so high-poaching firms switch with no assumption; at $\phi_f = 0$, Assumption 1 gives $B^F(\beta V^M, D) \geq K(D)$, so a full-retention firm preserves; continuity and strict monotonicity of B^F in r_f then deliver a unique cutoff $\phi^*(D)$, interior in $(0,1)$ wherever the assumption holds strictly, with switching above it and preservation below. Along the poaching dimension, the unscaled deepening at a switching firm is $\eta_f(1-\phi_f)\beta V^M \lambda \pi'(\psi_o)$: zero below $\phi^*(D)$, discrete at the cutoff, and declining toward zero

as $\phi_f \to 1$, which is the non-monotonicity recorded in the prediction's final claim. □

Prediction 4 and Proposition 2. For the prediction, consider one cohort of entering size J_0 in an otherwise steady pipeline; over the cohort's transit the senior stock, and with it the flow of openings, is fixed to first order, since seniors turn over at 11% a year. The managers the cohort eventually produces are $Y(J_0) = J_0\, y(J_0) = \bar{\pi}\,(\kappa M)^{-\eta_c} J_0^{1+\eta_c}$, so $\partial \ln Y / \partial \ln J_0 = 1 + \eta_c$ exactly, and the per-entrant yield has elasticity η_c; the shortfall reaches the senior state only through the chain's transit. For the proposition's long-run claim, the chain is linear in the stocks at given hazards and the yield depends only on the ratio $J_0/\kappa M$, so the system is homogeneous of degree one: under a permanent proportional cut all stocks scale by the cut, the ratio and hence the yield are unchanged in the new steady state, and $x^* = (\mathbf{I}-\mathbf{P})^{-1} eH$ scales one-for-one, $\Delta x^*/x^* = \Delta H/H$. For the transition bound, every tier loses workers to exit or promotion, so the spectral radius of $\mathbf{P}$ is below one and approaches one as the hazards shrink. Along the transition to a permanent cut the early cohorts enter while M is still near its old level, so their $J_0/\kappa M$ lies below baseline and their yields exceed $\bar{\pi}$; manager production therefore exceeds the fixed-hazard path at every date before the new steady state, and the fixed-hazard chain bounds the transition deficit from above. □

Under-provision and the deployment distortion. Assume the survival function S is log-concave with a strictly increasing, differentiable hazard η (Bagnoli and Bergstrom, 2005), and let the planner value a trained manager at the same capitalized rent V^M wherever she works, so that wages are transfers; more generally, with a social value V^S distinct from the employer's V^M, the common-architecture staffing wedge below carries $\beta[V^S - (1-\phi)V^M]\pi_q$ in place of $\phi\beta V^M \pi_q$, and the form used in the text is its $V^S = V^M$ case; the architecture comparison itself is stated for $V^S = V^M$. The planner comparison holds the congestion factor of Section 2.4 fixed, so aggregate manager formation rises with hiring and a hiring shortfall passes into a manager shortfall. The planner also chooses each firm's deployment architecture, facing the same preservation cost $K(D_f)$; its thresholds are $T_{fo}^{P,q} = w^J + D_f v_o - \beta V^M \pi_q(\psi_o)$, and its benefit from preserving training, B^P, is defined as B^F with βV^M in place of r_f.

Proposition A1 (Under-provision and the deployment distortion)**.** *(i) $B^P > B^F$, so there is a nonempty region $B^F < K \le B^P$ in which the firm adopts the substitution-intensive architecture while the planner preserves training. (ii) Whatever the pair of architectures, the planner staffs weakly more roles: conditional on a common architecture q the marginal wedge between the social and private value of staffing a role is $\phi\,\beta V^M \pi_q(\psi_o) \ge 0$, strictly positive wherever $\pi_q(\psi_o) > 0$, and where the firm alone has switched it is $\phi\,\beta V^M \pi(\psi_o) + \lambda\, r_f\, \pi(\psi_o)$, strictly positive wherever*

$\pi(\psi_o) > 0$. *(iii) Equilibrium junior hiring and the steady-state manager stock lie weakly below the planner's on both margins,* $H^{\text{eq}} \le H^{\text{opt}}$ *and* $M^{\text{eq}} \le M^{\text{opt}}$*, strictly wherever a positive measure of roles at the staffing margin carries* $\pi_q(\psi_o) > 0$ *or the firm alone has switched. (iv) Provided preservation remains socially efficient over the relevant range of deployment,* $K(D) \le B^P(D)$*, then for any exposed occupation (*$\nu_o > 0$*) with* $\pi(\psi_o) > 0$ *the proportional hiring shortfall* $L_{fo}(D) = \ln H^{\text{opt}}_{fo} - \ln H^{\text{eq}}_{fo}$ *is strictly increasing in the displacement and jumps upward where the firm crosses its cutoff* $r^*(D)$*; for* $\lambda < 1$ *the shortfall in manager formation widens by more, since formation per surviving hire falls by the factor* $1-\lambda$ *at switching firms, and at* $\lambda = 1$ *the formation comparison holds in levels. A subsidy to junior hiring of* $s^*_q = \phi\,\beta V^M \pi_q(\psi_o)$*, paid on the formation content the architecture realizes, restores both the planner's staffing and the planner's architecture.*

Proof. (i) B is strictly increasing in the valuation placed on formation, by the argument in the proof of Proposition 1, and $\beta V^M > (1-\phi)\beta V^M = r_f$; hence $B^P > B^F$, and the region is nonempty for K in the stated interval. (ii) If both preserve, $T^{P,A}_{fo} - T^A_{fo} = -\phi\beta V^M \pi(\psi_o) < 0$; if both substitute, the same difference carries the factor $1-\lambda$; if the firm substitutes while the planner preserves, $T^S_{fo} - T^{P,A}_{fo} = \beta V^M \pi - r_f(1-\lambda)\pi = \phi\beta V^M \pi + \lambda r_f \pi > 0$. In every case the planner's threshold is weakly lower, strictly so where the relevant $\pi_q(\psi_o) > 0$, so it staffs weakly more roles, and it forms weakly more managers per hire since $\pi_A \ge \pi_S$; (iii) follows with the congestion factor fixed. (iv) The claim is per cell. While the two agree on the architecture, $L'_{fo}(D) = \nu_o\big[\eta(T^q_{fo}) - \eta(T^{P,q}_{fo})\big] > 0$ under the strictly increasing hazard, since $T^q_{fo} > T^{P,q}_{fo}$ when $\pi_q(\psi_o) > 0$; after the firm alone has switched, $L'_{fo}(D) = \nu_o\big[\eta(T^S_{fo}) - \eta(T^{P,A}_{fo})\big] > 0$, since $T^S_{fo} > T^{P,A}_{fo}$ by (ii); at the firm's cutoff the private threshold jumps up by $r_f\lambda\pi(\psi_o)$ while the planner's is unchanged, so L_{fo} jumps upward. Formation additionally falls by the factor $1-\lambda$ at the switching firm when $\lambda < 1$; at $\lambda = 1$ the switching firm's formation is zero and the comparison holds in levels. If deployment grows powerful enough that the planner itself would abandon preservation, $K > B^P$, the comparison collapses to the common-S case and the monotonicity claim applies within, not across, that boundary. For the subsidy, a payment of $\phi\beta V^M \pi_q$ per hire shifts every private staffing threshold to the planner's, and it raises the private benefit of preservation from B^F, computed at the retained value r_f, to B^P, computed at βV^M, so the private architecture choice coincides with the planner's. □

B Data construction and measurement

Workforce panel. Revelio Labs parses each employment spell into an employer, a six-digit SOC occupation, and an ordinal seniority level 1–7, and resolves employers to public-firm

(gvkey) identifiers. I keep U.S. worker-years at Compustat-matched firms over 2017–2024 and require at least fifty worker-years per firm in 2021, leaving 9,322 firms and about 192 million worker-years. The within-firm specifications collapse this to firm-by-exposure-group-by-year (or firm-by-occupation-by-year) cells; the formation and projection results use the worker-level transitions. Seniority groups into entry (1–2), middle (3–4), and senior (5–7); the two entry levels differ, with level 1 lower-wage and about half exposed and level 2 the cognitive analyst rung about four-fifths exposed (Table 16).

Table 16: Workforce structure by seniority level (2021 cross-section)

Level	Group	Worker-years	Exposed share	Mean salary	Mgmt. occ.	Mgr. title
1	Entry (lower-wage)	7.8M	0.50	$48k	0.07	0.09
2	Entry (cognitive analyst)	6.7M	0.79	$92k	0.15	0.14
3	Middle	3.2M	0.85	$110k	0.21	0.22
4	Middle	2.9M	0.85	$124k	0.34	0.30
5	Senior	3.0M	0.84	$180k	0.45	0.39
6	Senior	0.9M	0.83	$218k	0.52	0.45
7	Executive	0.1M	0.70	$235k	0.74	0.65

Revelio worker-years in 2021. Exposed share is the fraction in top-tercile output-exposure occupations (Eloundou et al., 2024). Mgmt. occ. is the share in SOC major group 11; Mgr. title the share with a people-management title keyword. Salaries are Revelio's modeled compensation.

Learning content ψ. The index is the mean O*NET (version 29.1) importance rating of ten work activities through which entry work builds managerial judgment: *making decisions and solving problems*, *analyzing data or information*, *developing objectives and strategies*, *judging the qualities of objects, services, or people*, *thinking creatively*, *evaluating information to determine compliance with standards*, *interpreting the meaning of information for others*, *providing consultation and advice to others*, *organizing, planning, and prioritizing work*, and *updating and using relevant knowledge*.[14] I average eight-digit detail occupations within each six-digit SOC code and standardize across the matched occupations. The index excludes supervisory activities entry workers do not yet perform, so it measures what doing the entry work teaches. It is rated for 757 of the 858 entry occupations that also carry an output-exposure score (90% of entry person-years), and the formation results are unchanged when the unrated codes are imputed from O*NET 30.3, which rates Software Developers and the other new codes: both coefficients move by less than half a standard error.

[14] O*NET element IDs 4.A.2.a.1, 4.A.2.a.3, 4.A.2.a.4, 4.A.2.b.1, 4.A.2.b.2, 4.A.2.b.3, 4.A.2.b.4, 4.A.2.b.6, 4.A.4.a.1, and 4.A.4.b.6; eight in Mental Processes (4.A.2) and two in Interacting with Others (4.A.4).

AI usage. The Anthropic Economic Index reports the share of usage on each task category, matched to O*NET tasks and aggregated to occupations as a usage-penetration measure. It is predetermined relative to the formation outcomes (which depend on pre-2019 entrants) and is measured after the contraction begins, so it enters the contraction regression as a contemporaneous behavioral measure of deployment and the formation regression as a control for "what AI does." Orthogonalized to ψ, it isolates usage variation independent of learning content.

Estimating equations. The dynamic within-firm specification is (4). Where a series has a pre-2022 trend, I estimate the detrended difference-in-differences

$$y_{igt} = \gamma\,(\text{Exposed}_g \cdot t) + \textstyle\sum_{s\in\{2020,2022,2023,2024\}} \delta_s\,(\text{Exposed}_g \cdot \mathbf{1}\{t=s\}) + \nu_{it} + \nu_{ig} + \varepsilon_{igt},$$

with the exposed-specific trend γ identified off the pure pre-period (2017–2019, 2021) and δ_s the deviation from it; the design requires $\delta_{2022} \approx 0$. Firm-level capability and the continuous interactions take the triple form $y = \beta\,(\text{Exposed}_g \cdot \text{post}_t \cdot M_i) + \cdots + \nu_{it} + \nu_{ig}$ with M_i a pre-shock firm measure. Headcount regressions use the log cell count.

Coverage. Revelio headcount tracks Compustat reported employees throughout, with a cross-sectional rank correlation of 0.88–0.89 and a median ratio near three-quarters in every year (Table 17), so coverage does not contract through the shock. The panel is built from professional profiles, which are densest in exactly the white-collar occupations these firms report, so the level of the ratio is unsurprising; what the design needs from it is that it does not move. The within-firm design is the backstop: firm-by-year effects absorb any coverage change common to a firm's workforce, and Brynjolfsson et al. (2025) find the same early-career contraction in administrative payroll data, which has no profile-reporting lag.

Table 17: Revelio coverage over time

Year	Revelio / Compustat ratio	Rank corr. with Compustat	Junior share
2017	0.73	0.88	0.60
2019	0.73	0.88	0.60
2021	0.75	0.89	0.59
2022	0.72	0.89	0.58
2023	0.73	0.89	0.57
2024	0.74	0.88	0.56

Median ratio of Revelio U.S. headcount to Compustat reported employees across firms with both; cross-sectional Spearman rank correlation; junior (seniority 1–2) share of panel person-years.

Measurement quality. Three variables are modeled by Revelio rather than reported: seniority, occupation, and pay. I use pay only in within-firm comparisons and ratios, where the modeling error differences out, and I corroborate any pay-based claim with a headcount- or rate-based measure, so that no result rests on a modeled salary alone. The headline contraction is a headcount and hiring-flow result that does not rely on the pay model. The seniority assignment that defines the entry rung tracks salary and management-occupation share monotonically across levels (Table 16), and Revelio headcount tracks Compustat employment throughout the sample (Table 17), so neither the levels nor the within-firm comparison depends on a single imputed field. As an external check, Brynjolfsson et al. (2025) find the same early-career contraction in administrative payroll data, which involves no imputation.

C Robustness of the entry-rung contraction

Reading parallel trends. The primary evidence on parallel trends is the within-firm event study itself, which imposes no trend model and lets the pre-period be read directly: the exposed-minus-unexposed gap is flat through 2021 and breaks only after the release (Table 3, Figure 1). The linear detrend used in the tables below is a supplementary stress test, and it is conservative against a *cumulating* treatment effect: an exposed-specific trend fitted through the post period absorbs part of the 2022–24 AI ramp and biases the estimate toward zero (Meer and West, 2016), while conditioning on passing a pre-test would itself distort inference (Roth, 2022). As a quantitative summary I report, following Rambachan and Roth (2023), the largest relative magnitude of post-period parallel-trends violation (as a multiple of the worst pre-period violation) under which the contraction's confidence set still excludes zero: the headline within-firm contraction survives to $\bar{M} = 0.75$. The detrend therefore understates rather than manufactures the effect. Where it is decisive – the learning-content deepening, whose undetrended version carries the pre-2022 drift of high-ψ work (Table 23) – the flat 2022 coefficient of the detrended event study confirms the trend is removed rather than manufactured.

Placebo and the pandemic. A placebo dating the event to 2021 yields post-event deviations indistinguishable from zero in every group (all $|t| < 1.1$), so the pre-period shows no curvature the detrending would mistake for an effect. The within-firm design also nets out the 2020–21 remote-work hiring surge: the exposed-versus-unexposed gap shows no positive pandemic surge (the 2020 coefficient is −0.007) and diverges only after the release; fitting the pre-trend on pre-pandemic years alone gives a 2024 deviation of −0.030 and dropping 2020–21 gives −0.041, and the capability triple is unchanged at −0.033 with or without the pandemic years.

The level of the shortfall. Exposed junior hiring across the sample fell from 1.83 million in 2021 to 1.08 million in 2024; net of the pre-2022 trend the 2024 flow is 14.7% below its extrapolated path. Applying the detrended deficit share to realized hires gives about 159,000 missing entrants in 2024 and 141,000 in 2023, roughly 300,000 across the two post-shock cohorts (Table 18). Counting against the extrapolated path instead of realized hires gives about 343,000; the text uses the conservative convention. The firm-wide hiring freeze tracks prior over-hiring, not AI capability: in a firm-level regression prior over-hiring predicts the total-hiring decline (−0.061, s.e. 0.0033) while AI capability does not (−0.005, s.e. 0.0033).

Table 18: Additional results cited in the text

	Coefficient/value	(s.e.)
Exposed junior hires, levels (millions):		
2021 / 2023 / 2024	1.83 / 1.30 / 1.08	
detrended hire deficit, 2023 / 2024	−10.8% / −14.7%	
⇒ missing entrants, 2023 / 2024	≈141,000 / 159,000	
Promotion ladder intact (homegrown senior share, dev. 2024):		
All AI-capable firms	+0.022	(0.013)
High-capability firms	+0.035	(0.016)
Firm-wide hiring is not the driver (firm-level):		
Total firm hiring on prior over-hiring	−0.061	(0.0033)
Total firm hiring on AI capability (net)	−0.005	(0.0033)
Earnings-call capability validation (within-firm triple, dev. 2024):		
generative-AI call mentions × exposed × 2024	−0.029	(0.012)
jointly with AI capability: mentions / capability	−0.026 / −0.031	(0.016) / (0.017)

Within-firm detrended estimates unless noted. The detrended deficit share applied to realized hires implies about 141,000 missing entrants in 2023 and 159,000 in 2024. Earnings-call generative-AI mentions are measured over 2023–24, per million transcript words, winsorized and standardized over the firms with calls. Observations: Panel 1 9,322 firms; Panel 2 6,663 firms; Panel 3 1,830 firm-years; Panel 4 2,402 firms (2,249 in the joint columns); clustered by firm.

Earnings-call validation of capability. The capability measure is a pre-shock workforce share, not an instrument, so I corroborate it with an independent, behavior-based proxy for AI deployment: the firm's generative-AI mentions on its 2023–24 earnings calls. These reproduce the within-firm contraction on their own (−2.9% per standard deviation; Table 18), and entered jointly with the AI-workforce capability measure the two split the gradient (−0.026 and −0.031 per standard deviation). Measured after the release, the call mentions corroborate the pre-shock capability measure rather than add identification, but they confirm that the firms cutting exposed entry work are the ones actually talking about deploying AI.

The adoption battery. The capability gradient puts the contraction at the firms able to deploy AI. Whether those firms in fact deploy it can be checked directly, and on four independent measures they do: the spread of generative-AI skills through a firm's workforce, its hiring into AI occupations, the AI-role share of its job postings, and its earnings-call generative-AI mentions. The exposed contraction is deepest at the deploying firms on every measure: a standard deviation of each deepens the 2024 exposed cut by 2.8 to 3.9 percent, off a flat 2020 placebo, and a standard deviation of the combined index deepens it by about five percent (Table 19, Panel A). The pattern holds within firms over time: the exposed rung contracts in step with a firm's own contemporaneous adoption, while its adoption a year later does not predict the current cut (Panel B), so the timing follows deployment rather than a fixed firm type. The shift-share instrument of Table 5 predicts each adoption measure (Panel C): the variation it isolates reflects actual deployment. These measures are correlated with the pre-shock capability stock and do not separate from it – capability is a firm's readiness to deploy, adoption its realization – so they show that the capability gradient is realized AI deployment rather than a separate channel.

The capability gradient by tercile and tier. The high- and low-capability series are coincident through 2022 and diverge only afterward, reaching −7.8% against −1.6% by 2024 (Table 20); across tiers, the cognitive entry rung contracts most (tier 2, −7.5%), the lower-wage and middle tiers are flat, and the senior and executive tiers decline a significant but far smaller 1.8 to 3.3 percent. The gradient is monotone in capability, holds net of either task-exposure index, is present and stronger in non-technology firms, and is driven by exposure rather than cost: at high-capability firms the exposed rung falls 8.1% (the same high-capability cut the event study above puts at −7.8%) while equally paid unexposed occupations do not move (Table 21). The gradient also survives stricter definitions of the AI workforce: requiring at least two of the three classifier signals – role, skill, or degree – rather than any one, the high-capability exposed contraction is −7.3% (s.e. 0.011) and the high-minus-low differential is −6.1%; the all-three variant is positive for only about a third of firms, too sparse for a tercile split to separate treated from untreated firms (differential −0.5%).

Shift-share instrument diagnostics. The shift-share instrument of Section 4 (Table 5) passes the standard checks for a Bartik design (Table 22). It is built from 18,840 university shocks, and the Rotemberg weights that measure each shock's contribution to the estimate are diffuse: the effective number of shocks is 49, the five highest-weight universities carry 21% of the weight and the top twenty carry half, and the highest-weight sources are the large producers

Table 19: Direct evidence that the contracting firms are deploying AI

Per standard deviation of the measure	Coef.	(s.e.)
Panel A. Exposed contraction by direct adoption measure (detrended exposed × 2024 × adoption)		
Generative-AI skill diffusion (workforce share)	−0.032	(0.0085)
AI-occupation hiring	−0.039	(0.0078)
AI-role job-posting intensity	−0.036	(0.0063)
Earnings-call generative-AI mentions	−0.028	(0.0117)
Combined adoption index	−0.050	(0.0079)
combined index, 2020 placebo	−0.003	(0.0041)
Panel B. Within-firm timing: exposed × firm's own AI adoption, per percentage point		
Contemporaneous adoption	−0.00119	(0.00025)
One-year-ahead adoption (placebo)	−0.00010	(0.00027)
Panel C. Shift-share instrument predicts adoption (first-stage F, $Z \rightarrow$ measure)		
Postings / earnings calls / combined index	$F = 357 / 86 / 64$	
GenAI skills / AI hiring	$F = 44 / 17$	

Each adoption measure is firm-level and standardized; firm sources are Revelio worker skills and positions, Revelio job postings, and Capital IQ earnings-call transcripts. Panel A: the detrended within-firm triple difference of Table 4 (firm-by-year and firm-by-exposure fixed effects, firm-clustered) with the adoption measure in place of capability; the entry is the 2024 exposed-by-adoption deviation from the pre-2022 path, with the 2020 placebo for the index. Panel B: the exposed gap interacted with the firm's own contemporaneous, and one-year-ahead, AI-role posting share (same fixed effects). Panel C: first-stage F from regressing each adoption measure on the shift-share predicted capability Z of Table 5. The adoption measures are correlated with the pre-shock capability stock and do not separate from it, so they confirm the capability gradient is realized deployment rather than identifying a separate channel. Observations per measure: 6,235 / 5,705 / 9,798 / 2,435 firms (skills / hiring / postings / calls), clustered by firm.

of AI talent (Northeastern, Berkeley, the Massachusetts Institute of Technology, the University of California San Diego, and San José State), as a valid instrument should be. As in most Bartik designs the weights are not all positive: 70% of the universities carry negative weight, summing to −0.80 against the positive mass. The first stage survives dropping the highest-weight universities (the F falls only to 371 dropping the top five and to 332 dropping the top twenty, from a baseline of 401), so it is not driven by a handful of sources. It also survives dropping the firms headquartered in the technology-hub states: on the firms outside California, Washington, New York, Massachusetts, and Texas the first-stage F is still 112, so the instrument is not merely the geography of where AI-capable firms cluster. The instrument is balanced on firm size (correlation −0.14) and survives controlling for firm size and the headquarters-state AI share together (0.543, s.e. 0.029). It is also specific to AI talent rather than to elite or technical recruiting in general (Panel D): in a horse race against parallel predictors built from each university's engineering-and-science placement and its non-technical-professional placement, the AI-talent predictor still drives the contraction (−2.2%, s.e. 0.0045) while the

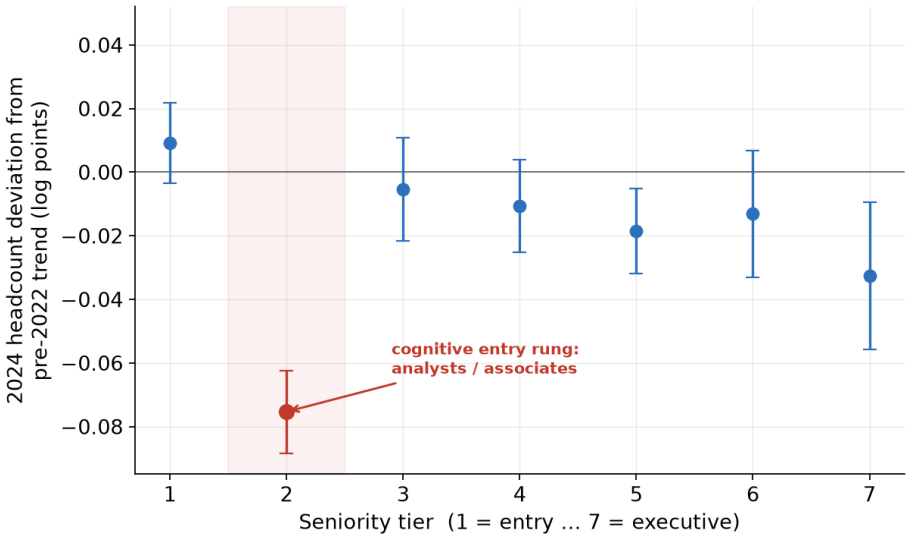


Figure 6: 2024 deviation from the within-firm exposed pre-trend, by seniority tier.

Specification (4): the 2024 deviation from the pre-2022 exposed path, by seniority tier. Coefficients (s.e.) by tier: +0.9 (0.65), −7.5 (0.66), −0.6 (0.83), −1.1 (0.75), −1.8 (0.68), −1.3 (1.01), −3.3 (1.18), in log points ×100. The cognitive entry rung (tier 2; analysts and associates; ≈79% exposed) falls 7.5%; lower-wage entry work (tier 1) and the middle tiers do not move; the senior and executive tiers (5 and 7) decline a significant but far smaller 1.8 and 3.3%.

engineering-and-science predictor is null, and it survives the analogous first-stage horse race (0.29, s.e. 0.017). Non-technical professional recruiting predicts a separate contraction of its own, but on an axis nearly uncorrelated with AI-talent recruiting (correlation 0.04): the AI-talent predictor survives net of it (−2.2%, s.e. 0.0045), so this second correlate does not dilute the AI-talent channel – it means only that the contraction has more than one recruiting correlate. The standard share-exogeneity caveat remains – the recruiting footprint and the leave-out shifters must be exogenous – but these diagnostics bound the concern rather than leaving it open.

Table 20: Entry event study by AI-capability tercile

	High capability		Low capability	
Year (ref 2021)	Coef.	(s.e.)	Coef.	(s.e.)
2020	−0.011	(0.0056)	−0.004	(0.0033)
2022 (release)	−0.008	(0.0052)	+0.002	(0.0029)
2023	−0.039	(0.0066)	−0.011	(0.0040)
2024	−0.078	(0.0077)	−0.016	(0.0046)

Within-firm exposed-minus-unexposed entry log headcount, ref 2021, estimated separately for the top and bottom terciles of pre-shock AI capability. Observations 39,666 (high) and 40,174 (low) firm-by-exposure-by-year cells; 2,910 and 2,880 firm clusters.

Table 21: AI capability versus generic technology intensity

	Coefficient	(s.e.)
Panel A. Entry headcount by AI-capability tercile (dev. 2024):		
Low / Middle / High	+0.009 / −0.017 / −0.081	
Panel B. Capability effect net of task exposure (within-firm triple):		
AI capability × exposed × post	−0.029	(0.0051)
task exposure × exposed × post (same regression)	−0.014	(0.0032)
Panel C. By sector (within-firm triple):		
Non-technology firms	−0.045	(0.0079)
Technology firms	−0.021	(0.011)
Panel D. Exposure vs. cost (high-capability firms, dev. 2024):		
Exposed entry rung	−0.081	(0.012)
High-wage unexposed entry occupations	+0.002	(0.020)

AI capability is the firm's pre-shock AI workforce share. Panel B nets out either task-exposure index; Panel C shows the gradient is not generic technology intensity; Panel D sets the exposed contraction beside equally paid unexposed entry occupations at the same high-capability firms. Observations about 45,900 firm-by-exposure-by-year cells per capability tercile (Panel A) and 138,208 (Panel B); about 2,884 and 8,727 firm clusters.

Table 22: Shift-share instrument diagnostics

Panel A. Rotemberg weights (which university shocks drive identification)	
University shocks (with ≥ 50 workers)	18,840
Effective number of shocks ($1/\sum \alpha_u^2$)	49
Top-5 / top-20 weight share	0.21 / 0.50
Negative-weight share (of universities)	70%
Highest-weight sources	Northeastern, Berkeley, MIT, UC San Diego, San José State
Panel B. First-stage robustness (coefficient; s.e.; F)	
Baseline	0.596 (0.030; F=401)
Drop top-5 Rotemberg universities	0.575 (0.030; F=371)
Drop top-20 Rotemberg universities	0.548 (0.030; F=332)
Drop firms HQ in CA/WA/NY/MA/TX	0.397 (0.038; F=112)
Drop California-HQ firms	0.490 (0.033; F=228)
Net of firm size and HQ-state AI share	0.543 (s.e. 0.029)
Panel C. Balance	
corr(instrument, log firm size)	−0.14
corr(instrument, HQ-state AI share)	+0.28
Panel D. Recruiting horse race (exposed × post × predictor)	
AI/computer-talent recruiting	−0.022 (s.e. 0.0045)
engineering and science recruiting	−0.000 (s.e. 0.0044)
non-technical professional recruiting	−0.017 (s.e. 0.0037)

Diagnostics for the shift-share instrument Z_i (Table 5), following Goldsmith-Pinkham et al. (2020) and Borusyak et al. (2022). Rotemberg weights $\alpha_u \propto a_u \sum_i s_{iu}\tilde{X}_i$ decompose the instrument into the contribution of each university shock; they sum to one, with positive weight concentrated on major AI-talent producers and an effective number of shocks of 49. Panel B re-estimates the firm-level first stage dropping the highest-weight universities and the technology-hub-state firms; the instrument survives both. Its baseline (0.596; F=401) reproduces the point estimate of Table 5's first stage; the standard error and F differ because the diagnostics are computed on a separate build of Z. Panel C shows it is not a proxy for firm size or headquarters-state geography. Panel D horse-races it (reduced form) against parallel predictors built from each university's engineering-and-science and non-technical-professional placement; the AI-talent predictor survives and the engineering-and-science predictor is null. Observations 9,001 firms (Panels A–C, heteroskedasticity-robust); 141,200 cells on 8,957 firm clusters (Panel D).

D Additional results on the two indices

The pre-trend artifact. Learning-rich work was gaining share within firms before the release, so a specification without the detrend attributes that pre-2022 path to the post period and flips the sign of $\psi \times$ post to +0.0025, with a 2022 coefficient of +0.0054 that fails the design's pre-trend test. Detrended, every control for ν – linear, decile, cubic, common support – agrees on the deepening (Table 23). The earlier reading that learning content is spared once ν is flexibly controlled was this pre-trend, not a fact about the cut.

Table 23: Learning content in the contraction, by specification

$\psi \times$ post, within-firm entry headcount	Coef.	(s.e.)
Decile-by-year control for ν, detrended	−0.0038	(0.0008)
Linear control for ν, detrended	−0.0022	(0.0008)
Cubic control for ν, detrended	−0.0033	(0.0008)
Decile control, common support, detrended	−0.0035	(0.0010)
Decile control, *not* detrended (significant 2022 coefficient)	+0.0025	(0.0007)

Per standard deviation of ψ; firm-by-occupation-by-year entry cells, firm-clustered standard errors. A negative coefficient means high-learning-content work contracts more. The sign is set by the detrend, not by the form of the ν control: every detrended specification agrees, and the undetrended row carries the pre-2022 drift of learning-rich work into the post period. Observations 6,474,945 cells and 24,747 firm clusters (the full Compustat-matched set; Table 6) in the full-sample rows; the common-support row is smaller.

Formation, full specifications. Table 24 reports the learning-content and output-exposure coefficients in the formation regressions under linear, decile, and cubic controls and entered alone, for both outcomes. Learning content survives a flexible control for output exposure throughout; output exposure entered alone is far weaker and falls further once learning content is included. Entering AI usage alongside learning content leaves it unchanged (senior +0.0173, s.e. 0.0021) while usage is null on formation (−0.0001, s.e. 0.0012).

Learning content versus skill level. Learning content correlates 0.76 with cognitive complexity, 0.71 with Job Zone, 0.68 with required education, and 0.62 with the occupation's wage, so a fair question is whether the formation effect is learning content or simply the skill the job requires. Entered together, the two are hard to separate by linear controls, and the occupation's wage, itself partly the return to the human capital the work builds, absorbs ψ (Table 25, lower block). The sharper test holds the skill level fixed nonparametrically and asks whether, within it, the more learning-rich work predicts more managers. It does: ψ predicts senior attainment within cognitive-complexity deciles, within occupation families, and within

Table 24: Formation: learning content and output exposure, full specifications

Per SD	linear v	decile v	cubic v	alone
Reach senior tier (base 5.0%*)*				
Learning content ψ	+0.0199	+0.0172	+0.0161	+0.0218
	(0.0024)	(0.0021)	(0.0021)	(0.0020)
Output exposure v	+0.0060	—	+0.0221	+0.0120
	(0.0030)		(0.0040)	(0.0036)
Reach senior management, SOC-11 (base 2.1%*)*				
Learning content ψ	+0.0096	+0.0074	+0.0074	+0.0102
	(0.0020)	(0.0015)	(0.0018)	(0.0017)
Output exposure v	+0.0018	—	+0.0105	+0.0047
	(0.0015)		(0.0026)	(0.0016)

Linear probability models, 2017–2019 entrants, firm and entry-cohort fixed effects, occupation-clustered standard errors, per standard deviation. "Decile v" absorbs output exposure with decile fixed effects, so its own coefficient is not separately identified. Learning content is robust to every control; output exposure is not. Observations 16,119,935 entrants; 757 occupation clusters.

the high-complexity occupations where skill is uniformly high, and the part of ψ orthogonal to the entire skill bundle still predicts formation (Table 25). ψ thus adds incremental content within skill bands, the formative composition of the work, without being cleanly separable from the level of skill it accompanies.

Robustness of the ψ construction. The formation effect does not depend on how learning content is built (Table 26). Reconstructing ψ as the first principal component of the ten activities, as the eight Mental-Processes items alone, dropping the strategy or the consultation item, or using the Level rather than the Importance scale all leave the senior-tier coefficient between 0.016 and 0.019; dropping any single one of the ten activities leaves it in $[0.0163, 0.0181]$. A placebo built from generic cognitive complexity also predicts formation – co-equal with ψ for the senior tier and weaker than ψ for the management core – which is the collinearity with skill discussed above (Table 25), not a construction artifact.

Formation net of the entrant's own characteristics. The within-firm formation result controls for occupation traits; a complementary entrant-level test controls for the worker. Among the same entrants, regressing senior attainment on learning content with entry-cohort and SOC-major-group fixed effects and controls for the individual's own entry salary and degree, a standard deviation of learning content predicts a 0.4-percentage-point higher probability of reaching the senior tier (s.e. 0.19), against a 1.0-percentage-point cross-firm association before

Table 25: Learning content within skill bands: the formation effect is composition, not level

Coefficient on learning content ψ (per SD)	Senior tier	Senior mgmt.
Baseline (ψ and output exposure v)	+0.0199 (0.0024)	+0.0096 (0.0020)
Holding skill level fixed:		
within cognitive-complexity decile	+0.0117 (0.0031)	+0.0072 (0.0024)
within occupation family (SOC major group)	+0.0107 (0.0021)	+0.0040 (0.0012)
high-complexity occupations, full skill bundle	+0.0117 (0.0051)	+0.0103 (0.0059)
ψ residualized on the full skill bundle	+0.0050 (0.0025)	+0.0036 (0.0019)
Most aggressive control:		
full skill bundle including occupation wage	+0.0033 (0.0029)	+0.0038 (0.0029)

Linear probability models for whether a 2017–2019 entrant reaches the senior tier or senior management; firm and entry-cohort fixed effects, occupation-clustered standard errors, per standard deviation of ψ. The skill bundle is cognitive complexity (O*NET's 21 cognitive abilities), Job Zone, required education, and the occupation's 2021 log wage. "Holding skill level fixed" adds cognitive-complexity-decile or SOC-major-group fixed effects, or restricts to above-median-complexity occupations with the full bundle controlled. ψ correlates 0.62–0.76 with these traits, so it is collinear with skill; it survives every test that holds skill fixed and is absorbed only by occupation wage, which is itself partly the return to the human capital the work builds. Observations 16,119,935 entrants (15,855,123 with the full skill bundle); 757 occupation clusters (745 with the bundle).

the worker controls. The entrant's own entry salary absorbs about half (+1.8 percentage points per standard deviation, s.e. 0.14). Learning content therefore retains incremental predictive content net of the entrant's own pay and schooling, even though much of the raw cross-firm association reflects worker quality – the same collinearity-with-quality that the within-skill-band tests above bound.

Table 26: Robustness of the learning-content construction (formation effect)

Construction of ψ (per SD)	Senior tier	Senior mgmt.
Equal-weighted ten activities (baseline)	+0.0172 (0.0021)	+0.0074 (0.0015)
First principal component	+0.0173 (0.0021)	+0.0077 (0.0015)
Mental-Processes eight only	+0.0172 (0.0022)	+0.0078 (0.0016)
Excluding strategy	+0.0164 (0.0021)	+0.0065 (0.0014)
Excluding consultation/advice	+0.0165 (0.0022)	+0.0071 (0.0015)
Level scale (not Importance)	+0.0191 (0.0022)	+0.0069 (0.0014)
Leave-one-out range (senior tier)	[+0.0163, +0.0181]	
Placebo: cognitive complexity	+0.0188 (0.0026)	+0.0066 (0.0019)

Coefficient on each construction of learning content in the formation regression, per standard deviation, occupation-clustered standard errors in parentheses, with output exposure absorbed non-parametrically (v-decile fixed effects), firm and entry-cohort fixed effects; estimated on firm-by-occupation-by-cohort cells weighted by cell size. The effect is stable across constructions. The cognitive-complexity placebo also predicts formation (collinearity with skill, Table 25); ψ exceeds it for the management core. Observations 1,340,761 firm-by-occupation-by-cohort cells (16,123,812 entrants); 757 occupation clusters.

E The manager pipeline and the projection

Promotion into the senior tier. The promotion inflow into senior roles shows no contemporaneous break: in the raw within-firm event study the 2024 coefficient is −0.0007 (s.e. 0.0011) for the senior tier and +0.0001 (s.e. 0.0013) for the management-occupation core, and senior headcount shows only a small detrended deviation (−1.8%). External hiring into senior roles does not rise either. This is what the dynamics predict (Proposition 2): the thinned entry cohorts have not yet reached the tier.

Transition hazards and calibration. The stock-flow model is the ladder of equation (3) – entry into the junior state, promotion up one tier at a time, and exit from each – so it is calibrated to the promotion and exit hazards of the within-firm transition matrix over 2018–2021 (Table 27), and not to that matrix's off-ladder cells, for which the ladder has no transition. Of juniors who leave the junior state in a year, 90% exit the firm and 10% are promoted; the senior stock turns over at 11% per year. The proportional manager-deficit trajectory and half-life are nearly identical whether the hazards are estimated on the broad, the exposed, or the unexposed pipeline, because every pipeline is leaky enough that exits dominate promotions; the exposed (cognitive) pipeline sustains about three times the senior stock per unit of entry flow (0.20 vs. 0.06), against the 2.4-fold gap in the share of entrants who ever reach the senior tier (0.067 vs. 0.028; Table 8). Throughout, "exposed" is the top tercile of the output-

exposure index; the calibration computes the tercile over all rated occupations rather than the ψ-matched set, and the two sets coincide for 254 of the 256 exposed occupations.

Table 27: Annual transition matrix (within-firm, 2018–2021)

From \ To	Exit	Junior	Middle	Senior
Junior	0.172	0.805	0.018	0.005
Middle	0.131	0.019	0.824	0.026
Senior	0.108	0.006	0.017	0.869

Rows sum to one. Promotion hazards are JR→MID 1.8% and MID→SR 2.6%.

The cohort-to-manager elasticity. The elasticity of Section 6.2 is estimated at the occupation-by-entry-cohort level (2018–2021 cohorts, ≥ 30 entrants, fixed three-year attainment window), regressing log future managers produced on log cohort size with occupation and cohort-year fixed effects, clustered by occupation. The senior-tier elasticity is 0.67 (s.e. 0.08) and the management elasticity 0.50 (s.e. 0.05); both are below one because per-entrant yield rises as cohorts thin (the congestion term is −0.33 log points, s.e. 0.08). Including the panel-start cohort raises the estimates (0.77 and 0.61). The 2008–09 natural experiment gives an out-of-sample counterpart of 0.73.

Projection sensitivity. The deferred deficit's level scales nearly linearly with the size of the entry cut while its timing does not, and the half-life is set by the senior-exit hazard rather than the cut (Table 28). The twenty-year deficit runs from −2.8% at a 5% permanent cut to −11.4% at 20%, with a half-life near nineteen years throughout; defining the manager state strictly as a management occupation (SOC major group 11) rather than by seniority makes the deficit larger and faster (−12.6% at twenty years, half-life eleven years), so the seniority-based figures in the text are the conservative ones. They are conservative in a second respect. The estimated transition matrix also records small flows out of the senior tier that are not exits – 0.6% per year back to the junior tier and 1.7% to the middle – for which the ladder has no transition and which it therefore omits. Treating those flows as genuine departures from the senior state, so that the senior tier drains at 13.1% per year rather than 10.8%, raises the twenty-year deficit to −11.4% and shortens the half-life to thirteen years. The proportional trajectory is also nearly identical whether the transition hazards are calibrated on the broad, the exposed, or the unexposed pipeline, because every pipeline loses far more juniors to exit than to promotion. What the projection cannot bound from within the panel is reallocation

to private firms, government, or self-employment, which it does not observe (the within-panel reallocation check is above).

Table 28: Projection sensitivity: manager-stock deficit by entry cut and definition

Permanent entry cut	Deficit, 10 yr	Deficit, 20 yr	Half-life
−5%	−0.9%	−2.8%	19 yr
−10%	−1.7%	−5.7%	19 yr
−15% (baseline)	−2.6%	−8.5%	19 yr
−20%	−3.4%	−11.4%	19 yr
Manager = management occupation (SOC-11), −15% cut:			
management-occupation stock	−7.2%	−12.6%	11 yr
Career-wide hazards (tier survives a move across firms), −15% cut:			
senior (manager) stock	−0.8%	−3.8%	31 yr

Senior (manager) stock deficit from the stock-flow chain (3). The deficit scales nearly linearly with the cut while the deferred profile (half-life) is governed by the senior-exit hazard, not the cut size. The SOC-11 row redefines the manager state as a management occupation (SOC major group 11), which is less sticky, so the deficit is larger and arrives faster. The career-wide row recalibrates the hazards so that a worker's tier survives a move to another firm; seniors then turn over more slowly, and the within-firm baseline is the deficit in a firm's internally grown manager stock rather than in the economy-wide one.

The 2008–09 validation sample. The out-of-sample test follows every worker whose first career position was a U.S. entry job as a financial analyst (SOC 13-2051), management analyst (13-1111), or accountant (13-2011), cohorts 1998–2014, in the full Revelio position file (577,719 workers). Attainment is the first year a worker reaches the middle (3–4) or senior (5–7) tier; counterfactual sizes for the crisis cohorts interpolate log cohort size between the non-crisis neighbors. The prediction in Figure 5 applies a transition matrix estimated only on 2002–06 transitions to the realized cohort sizes; it tracks the realized series at correlation 0.97 over 2008–2020 and matches its cumulative 2008–19 level within 6%. Table 29 reports the cohorts.

The firm-resolved downstream and its shift-share instrument. The firm-level result of Section 6.2 (Table 15) is built from a separate, larger extraction of the raw Revelio position file. For each U.S. worker the entry position is the first career spell (earliest start date), kept if its seniority is in the entry range and its year is 2005–2016; the first senior position is the earliest spell at seniority ≥ 5; and a manager is *internal* if that senior spell is at the entry firm. Firm identifiers come from the Revelio–Compustat crosswalk. This yields 6.31 million entrants and, in the 2008–09 window, some 13–15k firms, more than the modern panel. Be-

Table 29: Validation-sample cohorts around the two entry collapses

Cohort	Size	Middle+ by +7	Senior by +10	Salary gap, promoted − never
2002 (dot-com)	22,809	47.9%	27.8%	$4,585
2004	32,260	49.3%	29.5%	$4,363
2006	43,726	48.8%	29.5%	$3,747
2007	46,451	49.3%	30.6%	$4,065
2008	43,915	51.1%	31.4%	$3,245
2009 (crisis trough)	33,397	53.1%	33.7%	$2,407
2010	39,109	54.3%	33.5%	$2,981
2011	40,621	53.9%	34.2%	$3,218

First-career finance and consulting entrants by cohort year. The crisis cohort (2009) entered 23% below the neighbor trend but was promoted at higher rates. Counterfactual sizes for 2008–10: 44.9k, 43.4k, 42.0k.

cause only about 5% of firm-by-occupation-by-cohort cells produce any internal manager, the make-side is a Poisson count model; a log-count specification attenuates the passthrough spuriously toward zero (Santos Silva and Tenreyro, 2006). Within firm and occupation the Poisson passthrough is 0.93 (s.e. 0.011); at the firm level it is 0.83 (s.e. 0.020). The instrument is the leave-one-out NAICS3-by-cohort entry total (the firm's own cohort removed): its first stage is +0.46 (s.e. 0.022, $F \approx 441$), and a control-function IV-Poisson gives 0.40 (s.e. 0.088), with the coefficient on the first-stage residual +0.48 (s.e. 0.093) confirming that firm-specific cutting inflates the uninstrumented firm-level estimate about twofold. Because the control function enters an estimated first-stage residual, both standard errors are bootstrapped over firms, re-estimating the first stage and the count model within each replication; the analytic errors that treat the residual as data are slightly wider (0.097 and 0.10). The passthrough is stronger in manager-making occupations (interaction with the occupation's senior-formation rate +0.032, s.e. 0.009). The shock is observed retrospectively through later-created profiles, which within-firm differencing mitigates but does not eliminate.

The make-versus-buy margin. A second extraction splits each firm's senior arrivals into internal promotions and external hires by year. Firms that cut their 2009 manager-making cohort more do not later buy more seniors externally: in levels the external share of senior arrivals is if anything lower (−0.024, s.e. 0.006) and the external senior count lower (−0.043, s.e. 0.011), and the within-firm change in the external share on the 2009 cut is a precise zero (+0.0005, s.e. 0.003). The deficit is realized as fewer managers, not bought back.

Within-panel reallocation. The missing entrants do not reappear at low-capability firms within the panel: exposed-junior worker-years fell across all three capability terciles (−4.2% overall, 2021–2024), −2.3% at low-capability firms against −4.7% at high-capability ones. Nor is AI creating an offsetting entry rung: entry hiring across all computer and mathematical occupations fell 34% from 2021 to 2024. The panel does not observe private firms, government, or self-employment, where some of the missing entrants may be reallocated.

F The appropriability identification boundary

The poaching rate that governs the externality is endogenous to a firm's location, and the candidate shifters of it (rival labor demand, non-compete enforceability, the H-1B lottery) all vary with where AI-capable firms cluster, so none separates a learning-content-specific poaching effect from the geography of AI capability. The exception is defense-sector exposure, built from a firm's pre-shock government and defense customer segments, whose geography is uncorrelated with AI capability (correlation 0.01) and which should lower worker mobility – though in these data it barely moves the measured poaching rate (correlation 0.001), so the shifter lacks a first stage. If defense exposure moved firms across the preservation cutoff, the learning-content deepening would be absent at these low-mobility firms; absent a crossing, the model does not sign the comparison. The learning-content–defense interaction is +0.001 (s.e. 0.0087), a null in every continuous and binary variant, before and after absorbing AI capability and occupation-by-year shocks. Where poaching variation can be separated from AI-capability geography, then, the estimate is a null too imprecise to pin the poaching elasticity's magnitude; its existence and sign are established in Section 7 from the model and the measured generality of the training.