AI and the Structure of Internal Labor Markets *

Guido Friebel Yao Huang Jin Li Yunchou Zhang

January 15, 2026

Abstract

AI is changing knowledge work: it leads to higher individual productivity and faster learning. We model the effects of these changes on internal labor markets and show that the classical pyramidal structure may or may not survive in steady state. We also show that large shocks will lead to oscillations in the span of control between senior and junior workers before reaching a new steady state. The model is robust against embedding the organization into a labor market and allowing for firm-specific human capital.

Keywords: Span of control, knowledge work, labor markets, firm-specific human capital

*Contact information: Guido Friebel (gfriebel@wiwi.uni-frankfurt.de), Department of Management and Microeconomics, Goethe University Frankfurt, 60323 Frankfurt/Main, Germany; Yao Huang (yaohuang@connect.hku.hk), Jin Li (jli1@hku.hk), Yunchou Zhang (zyc616@connect.hku.hk), Faculty of Business and Economics, The University of Hong Kong, Pokfulam Road, Hong Kong SAR.

1 Introduction

Many firms' internal labor markets have pyramidal shape. "Managing the pyramid" is tantamount to staffing junior layers in the hierarchy that feed more senior layers. Juniors acquire human capital on the job, get promoted and rewarded by higher-level jobs. Firms also hire from outside labor markets, but a healthy pyramidal structure of the internal labor market (ILM) is seen as quintessential for sustainable corporate growth. The financial firm in Baker, Gibbs and Holmstrom (1994*a*) is a classic example of such a pyramidal ILM.

In times in which AI is changing workplaces fast and already has massive repercussions on the labor market[1], many ask themselves how AI will shape the ILM of the future?

Duncan, Anderson and Saviano (2025) argue that consulting firms, the archetypical pyramidal firms, are changing their structure fast in response to working with AI. The emerging consensus seems to be that pyramids will be transformed into diamonds, in which the size of junior layers will be reduced, or into monoliths, obelisks or other exotically named rectangle structures. A good example of this view is provided in this quote from PwC's chief AI officer Dan Priest: "The more intrinsic AI agents become to how work is done, the more value can be unlocked. Our analysis — informed by our work with clients and our own AI adoption — shows this human-AI collaboration could boost productivity and speed by 50 percent. What emerges will be a shift of the traditional labor pyramid to a diamond shape" (Priest, 2025).

Data about the hiring and firing of consulting and other firms are hard to come by, and there are other views, for instance, that junior hiring may even increase (at the expense of seniors) if, by the use of AI, juniors learn faster. Amazon Web Services CEO Matt Garman is quoted to have said that firing junior workers because AI can do their jobs is "the dumbest thing I've ever heard... how's that going to work when ten years in the future you have no one that has learned anything"(Sharwood, 2025).

We suggest a simple model to provide predictions on what the steady state of firms will be and what the transition paths may look like, depending on what AI really does to the work and learning or people in organizations. In order to do so, we look at firms in which a CEO distributes tasks to senior workers that can do a certain number of tasks; junior workers take care of the remaining tasks. The productivity of senior workers is reasonably assumed to be higher than the one of junior workers, because the former have accumulated human capital.

[1]According to the last Bentley Gallup Business in Society Poll, 73 percent of the population in the USA believe that AI will reduce the total number of jobs (Bentley University and Gallup, 2025). A McKinsey survey reports that technology and professional service firms are most affected by AI; the functions most affected are marketing and sales, product and services development and IT (Singla et al., 2025).

Consider a law or consulting firm, for instance, in which each project manager is supported by many juniors in collecting and preparing data, doing background research, preparing board presentations, etc. These tasks can, by now, be readily farmed out to an AI. To save costs, firms may consequently decide to hire fewer junior workers or let junior workers go.[2] Firms may, however, be reluctant to layoff junior workers because today's junior workers will be future managers. Junior workers learn on the job and some of them will be qualified to join the ranks of senior workers. Put differently, in deciding how many juniors to have at any point in time, firms will take into consideration the constraints imposed by maintaining a steady flow of workers from the junior to the senior layer.

We investigate to what extent AI can be expected to reshape the talent management systems of firms that employ numerous junior workers to a) get tasks done, b) ensure human capital acquisition and fill the talent pipeline. We first show that firms may reach different steady states after being hit by AI. The precise nature of these steady states depends on what AI mainly does. If the main effect of AI is to increase the productivity of senior workers or junior workers, then, the pyramidal shape of the firm is maintained. If, however, AI makes it easier for juniors to learn what is needed to become senior workers, then, the span of control, i.e., the ratio between junior and senior workers, falls. This is tantamount to the firm assuming a rectangular or diamond shape.

The second main contribution is an analysis of the transition path from traditional ILMs to future ILMs in which AI is employed. In short run, whether AI increases productivity or accelerates learning, firms respond by hiring fewer juniors. This hiring freeze temporarily reshapes the hierarchy from a pyramid into a diamond, risking a costly break in the talent pipeline. Following this disruption, we identify different paths for the firm to reach a new steady state. In particular, transitions might be smooth, but they may also entail oscillations: the number of juniors might decrease and increase quite substantially, forcing the firm to alternate between pyramid and diamond structures before converging to the new steady state.

The ingredients of the model are inspired by a large literature on internal labor markets, a concept first introduced by Doeringer and Piore (1971). Their seminal work triggered a thriving theoretical and empirical literature, summarized by Gibbons (1997), Gibbons and Waldman (1999*a*), Lazear and Oyer (2013), and Waldman (2013).

A number of empirical papers, most notably, Baker, Gibbs and Holmstrom (1994*a*,*b*) provide a detailed case study of a firm indicating that ILM were conducive to human capital acquisition and providing promotion incentives (that are absent in our paper).[3]

[2]Indeed, we receive reports that around the world, consulting firms are reducing their junior hiring.

[3]Our work abstracts away from the incentive role of promotions, which are an important topic in ILM

Pastorino (2024) investigates the importance of human capital accumulation and learning about employees' abilities, especially for sorting employees to jobs, within one firm. Recent contributions underline the importance of internal labor markets and offer estimates that a large proportion of workers are covered by ILM practices despite increasing competitive pressure (Huitfeldt et al., 2023; Osterman, 2024).

Our work is related to the literature regarding vacancy chains and slot constraints in internal labor markets (Simon, 1951; White, 1970; Beckmann, 1978; Stewman and Konda, 1983; Rosenbaum, 1984; Ke, Li and Powell, 2018). These structural features impose constraints that bind the careers of workers across different levels. This interdependence creates career spillovers, where the progression of junior workers relies on the departure of seniors, a phenomenon documented empirically by Friebel and Panova (2008); Bianchi et al. (2023).

We are not aware of any paper that provides a framework to investigate the future shape of ILM. Our assumptions are simple, but grounded in the emerging empirical literature that shows that more tasks can be done by AI in a given time and leads to faster learning (Brynjolfsson, Li and Raymond, 2025; Chen et al., 2025).

2 Model Setup

Consider a firm consisting of three hierarchical layers: a CEO, senior workers, and junior workers. Time is discrete and indexed by $t = 0, 1, 2, \cdots, \infty$.

Workers In every period, the firm produces with n_t^h senior and n_t^l junior workers, where $n_t^h \geq 0$ and $n_t^l \geq 0$. Let the senior and junior workers' per-period wages be exogenously given as $w^h > 0$ and $w^l > 0$. Each senior worker can finish γ^h tasks, and each junior worker can finish γ^l tasks, where $0 < \gamma^l < \gamma^h$. We temporarily shut down the external labor market, and will include it and endogenize wages by allowing the firm to hire external seniors in Sections 5 and 6.

Firm The firm faces a production target of N tasks, each of which generates a revenue of y. We assume $y > w^l/\gamma^l$, implying that the marginal revenue of a task exceeds the marginal cost of junior labor. Hence, the firm always ensures all N tasks are completed; that is, we

(Lazear and Rosen, 1981; Rosen, 1986; Malcomson, 1984; MacLeod and Malcomson, 1988; Prendergast, 1993; Gibbons and Waldman, 1999*b*; Zabojnik and Bernhardt, 2001; Waldman, 2003; Kräkel and Schöttner, 2012; Auriol, Friebel and Von Bieberstein, 2016; Ke, Li and Powell, 2018; Bianchi et al., 2023).

have the following capacity constraint:

$$\text{(CC)} \qquad \gamma^h n_t^h + \gamma^l n_t^l \geq N.$$

The firm (or in fact the CEO, who acts as the residual claimant) gets the following per-period payoff

$$\pi_t = Ny - w^h n_t^h - w_l n_t^l.$$

Human Capital Accumulation At the end of each period, senior and junior workers leave the firm at an exogenous rate $d_h \in (0,1)$ and $d_l \in (0,1)$, respectively. After working in the firm for one period, juniors acquire the skills to become seniors with probability $q \in (0,1]$. Let h_t^l denote the net hiring to the junior rank in period t, which can be negative if the firm decides to fire some juniors. We hence have the following flow conditions that relate the numbers of workers on the two layers over time:

$$\text{(FC)} \qquad \begin{aligned} n_{t+1}^h &= (1-d_h)n_t^h + q(1-d_l)n_t^l, \\ n_{t+1}^l &= (1-d_l)(1-q)n_t^l + h_{t+1}^l. \end{aligned}$$

Timing Figure 1 shows the timeline of the stage game. At the beginning of period t, the firm decides whether they should hire or fire junior workers. Then the firm allocates the tasks between senior and junior workers. After that, the production happens and the output is realized. During the production process, junior workers acquire skills with certain probability. Finally, some workers leave the firm, and those remaining juniors who have acquired skills turn into seniors and get promoted.

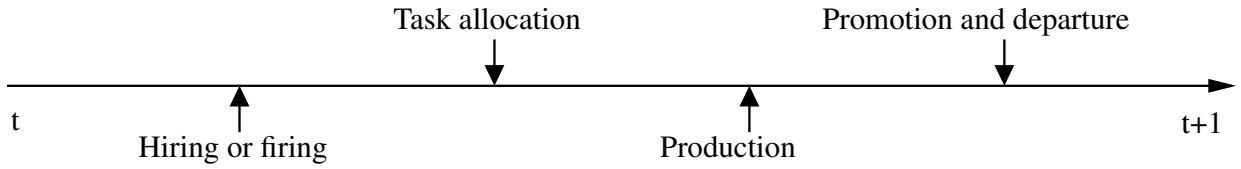


Figure 1: Stage Game Timeline

Remark 1. It is without loss of generality to assume that all workers are allocated with tasks at their full capacity. We could allow the firm to choose how many tasks are allocated to the workers. Let $D_t^l \in [0, \gamma^h n_t^h]$ denote the tasks that the CEO allocates to junior workers and $D_t^h \in [0, \gamma^l n_t^l]$ the tasks that are allocated to senior workers, where $D_t^h + D_t^l \leq N$. Notice that the firm pays the wages to workers, and juniors can only acquire skills from working.

Hence, it is inefficient to leave any worker idle. This means we must have $D_t^h = \gamma^h n_t^h$ and $D_t^l = \gamma^l n_t^l$.

3 Steady State Analysis

To study the steady state, we first investigate the firm's optimal hiring policy. In each period t, the firm has a stock of n_t^h seniors following from the flow condition (FC). Given these senior workers, the firm, facing the capacity constraint (CC), decides on how many juniors to hire to maximize her per-period payoff:

$$\begin{aligned} \max_{n_t^l} \quad & \pi_t(n_t^h) = Ny - w^h n_t^h - w_l n_t^l \\ \text{s.t.} \quad & \gamma^h n_t^h + \gamma^l n_t^l \geq N \end{aligned}$$

After solving the optimization problem, we have the following lemma describing the optimal hiring policy.

Remark 2. In principle, a firm should maximize the discounted value of the firm, and the optimization problem is

$$\max_{\{n_t^l\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \delta^t (Ny - w_t^h n_t^h - w_t^l n_t^l),$$

subject to the capacity constraint (CC) and flow conditions (FC). We show in Sections 5 and 6 that this dynamic problem simplifies to, or yields the same solution to a sequence of static problems. Thus, assuming period-by-period maximization here entails no loss of generality.

Lemma 1. *Given the stock senior workers (n_t^h), the firm hires $n_t^l = \max\{0, (N - \gamma^h n_t^h)/\gamma^l\}$ junior workers.*

Lemma 1 shows that the firm uses its seniors first, hiring just enough juniors to finish the remaining tasks. The firm does so because using the seniors it already has is cheaper than hiring new juniors.

Now we can investigate the steady state. In particular, all stocks and flows are constant over time in a steady state. Let n^{h*} and n^{l*} denote the number of high-ability and low-ability workers in a steady state. We have the following steady state constraints:

$$\text{(SS)} \qquad \begin{aligned} n_{t+1}^h &= n_t^h = n^{h*}; \\ n_{t+1}^l &= n_t^l = n^{l*}. \end{aligned}$$

Let $s_t := n_t^l/n_t^h$ denote the span of control in period t, and s^* represents that in a steady state. Together with Lemma 1 and the flow conditions (FC), we have the following lemma:

Lemma 2. *In the steady state, the span of control* (s^*)*, the number of juniors* (n^{l*})*, and the number of seniors* (n^{h*}) *are characterized as follows:*

$$s^* = \frac{d_h}{q(1-d_l)}, \quad n^{l*} = \frac{s^*}{\gamma^l s^* + \gamma^h} N \quad \textit{and} \quad n^{h*} = \frac{1}{\gamma^l s^* + \gamma^h} N.$$

Lemma 2 characterizes the organizational structure and the number of workers in each layer in the steady state. It shows the flow rates determine the span of control (s^*), or the ratio of juniors to seniors. This ratio, together with productivity (γ^l and γ^h) and the total number of tasks (N), determines the size of each layer (n^{l*} and n^{h*}).

First, the expression of s^* shows that the productivity does not affect the span of control. Only the flow rates (i.e., learning rate q and departure rates d_h and d_l) matter. The firm looks like a pyramid $(s^* > 1)$ when seniors leave often (high d_h), juniors leave often (high d_l), or juniors learn slowly (low q). In this case, the firm must hire many juniors just to keep the senior layer staffed. Conversely, if seniors stay longer (low d_h), juniors stay longer (low d_l), or juniors learn faster (high q), then firm needs fewer juniors to feed the senior layer. That is, the structure becomes more like a diamond $(s^* < 1)$.

Second, once the flow rates fix the span of control, the number of tasks (N) and workers' productivity (γ^h and γ^l) determine the size of each layer. The firm adjusts this fixed structure until the combined output of the two layers meet N.

How does AI change the firm structure? We distinguish two effects AI may have. First, AI may be a *productivity shock* that increases workers' capacity to take on tasks, γ^h or γ^l. Second, AI may be a *learning shock* that increases the probability q that a junior worker acquires high ability. Now we summarize how AI shocks affect the organizational structure in the long run.

Proposition 1. *The steady state reacts as follows to the AI shock:*

(*i*) *A positive productivity shock reduces the number of workers in both layers* $(n^{h*}$ *and* $n^{l*})$ *but leaves the span of control* (s^*) *constant.*

(*ii*) *A positive learning shock increases the number of high-ability workers* (n^{h*})*, reduces the number of low-ability workers* (n^{l*})*, and reduces the span of control* (s^*)*.*

Proposition 1 distinguishes how different AI shocks shape the internal labor market. First, consider AI as a positive productivity shock, i.e., an increase in γ^h or γ^l. This shock reduces the total workforce required to meet the production target N. While the total

workforce shrinks, the reduction occurs uniformly throughout the hierarchy. This is because the span of control (s^*) depends only on the flow rates (d_h, d_l, q), not on output per worker. Consequently, the organization shrinks in workforce but retain its structure in the long run; a pyramidal firm will shrink in size but still maintain its pyramidal structure.

Second, consider AI as a positive learning shock, i.e., an increase in q. A higher q raises the probability that a junior worker qualifies for promotion. This higher probability increases the flow of workers into the senior layer, increasing the number of seniors. As the flow upward accelerates, the firm needs fewer juniors to ensure the steady flow of talents. This reduces the span of control s^*. Consequently, a pyramidal firm will be more likely to adopt a diamond structure in the long run.

4 AI Shocks and Transition Dynamics

In this section, we investigate the transition of the organizational structure in details. We will first focus on the productivity shock brought by AI, and then discuss the learning shock.

4.1 Productivity Shock

Consider a firm initially in a steady state with productivity of senior workers γ_0^h and junior workers γ_0^l. At $t = 1$, an AI shock permanently increases their productivity to γ_{AI}^h and γ_{AI}^l, where $\gamma_0^l < \gamma_{AI}^l$ and $\gamma_0^h < \gamma_{AI}^h < \gamma_0^h + \frac{d_h}{q(1-d_l)}\gamma_0^l$[4]

Let $n^{h*}(\gamma^h, \gamma^l)$ and $n^{l*}(\gamma^h, \gamma^l)$ denote the steady-state number of seniors and juniors given the pair (γ^h, γ^l). Let $r^\gamma := \gamma^h/\gamma^l$ denote the ratio between the productivity of senior and junior workers. To ensure stability, we assume $r_{AI}^\gamma < \frac{2-d_h}{q(1-d_l)}$. The transition dynamics are characterized in the following proposition.

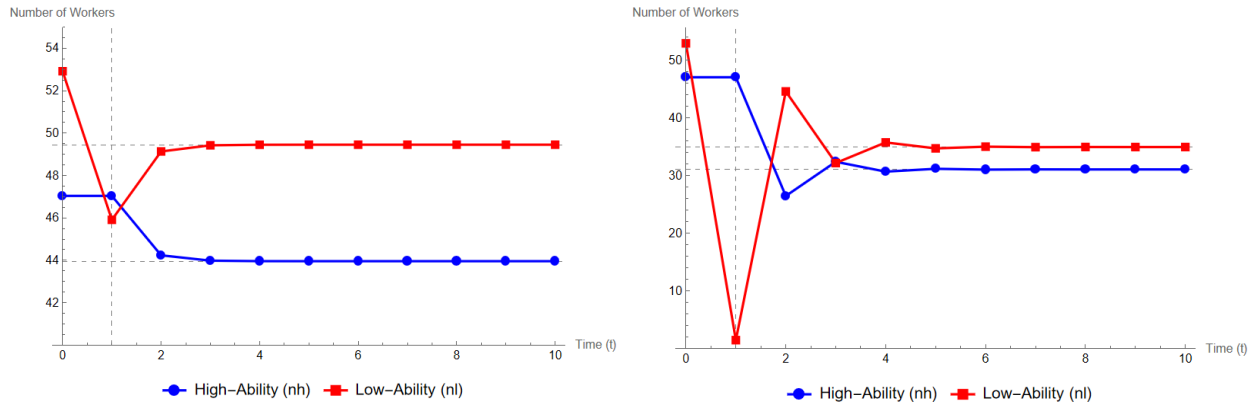
Proposition 2. *When $t = 1$, the number of seniors remains the same $(n^{h*}(\gamma_0^h, \gamma_0^l))$ while the number of juniors drops below the new steady-state level $(n^{l*}(\gamma_{AI}^h, \gamma_{AI}^l))$. When $t > 1$, the pattern of the transition is summarized as follows:*

(i) If $r_{AI}^\gamma < \frac{1-d_h}{q(1-d_l)}$, the number of seniors falls monotonically to the new steady state $n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l)$, while the number of juniors rises monotonically to $n^{l*}(\gamma_{AI}^h, \gamma_{AI}^l)$.

[4]If $\gamma_{AI}^h > \gamma_0^h + \frac{d_h}{q(1-d_l)}\gamma_0^l$, $(N - \gamma_{AI}^h n_1^h)/\gamma_{AI}^l < 0$. From Lemma 1 we have $n_1^l = 0$. Together with (FC) we know that $n_{t+1}^h < n_t^h$ if $n_t^l = 0$. Then, at certain period $T > 1$, we must have $n_T^l = (N - \gamma_{AI}^h n_T^h)/\gamma_{AI}^l > 0$. Therefore, in some periods, the firm will do without any juniors, but in the long run the structure converges to the new steady state. For simplicity, we add this restriction to maintain a three-layer structure over the transition.

(*ii*) *If* $r^{\gamma}_{AI} > \frac{1-d_h}{q(1-d_l)}$, *the numbers of workers in both layers converge to their new steady states by alternating above and below them in opposite directions* (*i.e.*, $[n^h_t - n^{h*}(\gamma^h_{AI}, \gamma^l_{AI})][n^l_t - n^{l*}(\gamma^h_{AI}, \gamma^l_{AI})] < 0$).

Proposition 2 shows that the number of juniors drops immediately after the shock while the number of seniors remains the same temporarily. Later on, the transition dynamics depends on how much the shock favors the seniors. In particular, if the productivity favors the juniors more ($r^{\gamma}_{AI} < \frac{1-d_h}{q(1-d_l)}$), then the transition will be smooth. However, if the productivity favors the seniors more ($r^{\gamma}_{AI} > \frac{1-d_h}{q(1-d_l)}$), then we will observe the oscillation pattern.



(**a**) Smooth Adjustment: $\gamma^h_{AI} = 1.15$ and $\gamma^l_{AI} = 1$.

(**b**) Oscillations: $\gamma^h_{AI} = 2.095$ and $\gamma^l_{AI} = 1$.

Figure 2: Transition Dynamics with Positive Productivity Shock. Parameters: $d_l = 0.6$, $d_h = 0.45$, $q = 1$, $\gamma^h_0 = 1.001$, $\gamma^l_0 = 1$, and $N = 100$.

Figure 2 illustrates the transition patterns. After the positive productivity shock, there is a surge in the capacity of the existing workers. Lemma 1 shows that the firm will allocate as many tasks as possible to seniors. Therefore, the number of seniors remains at its original steady-state level, while the number of juniors suffers a huge drop. This drastic drop in the number of juniors is temporary, and it will bounce back in the second period. This is because, after the initial shock on the juniors, the firm faces a shortage of workers with the seniors leaving the firm. Therefore, it has to hire back a group of juniors to achieve the production target.

If the shock favors the juniors more, the ratio r^{γ} is more likely to be smaller. In this case, we will observe a smooth adjustment shown in Figure 2a. After the initial shock, the productivity of juniors increases by a lot. Therefore, when the firm hire back juniors in the second period, the number of juniors will not increase dramatically. With the juniors acquiring necessary skills, the firm will gradually have enough seniors to replace those departing workers. Consequently, the senior stock glides down gradually to its new steady

state. This smooth adjustment continues until the rising promotion inflow matches the falling outflow of high ability workers. At this point, the organization reaches its new steady state.

In contrast, if the shock favors the seniors more, the ratio r^γ is more likely to be larger. In this case, we will observe oscillations shown in Figure 2b. After the initial shock, the seniors are much more productive while the juniors are not. Therefore, the firm has to hire a large number of juniors facing the departure of seniors. This surge of juniors, in turn, increases the talent supply dramatically, pushing the number of seniors above the steady state. The organization thus oscillates around the new steady state until convergence. This implies that a firm, which had a pyramid structure before the shock, may oscillate between a diamond shape and a pyramid shape for some periods after the shock. This shows a risk of an AI productivity shock. A firm may aggressively downsize the junior work force to capture efficiency gains, but this may deplete the talent pipeline needed for the future.

4.2 Learning Shock

Consider a firm initially in a steady state with a learning probability q_0 for the junior workers. At $t = 1$, an AI shock (e.g., an AI copilot that helps junior workers learn faster) permanently increases to q_{AI}, where $q_0 < q_{AI} < q_0 + \frac{\gamma^l}{\gamma^h(1-d_l)}$.[5]

Let $n^{h*}(q)$ and $n^{l*}(q)$ denote the steady-state number of senior and junior workers given q. To ensure stability, we assume $q_{AI} < \frac{\gamma^l(2-d_h)}{\gamma^h(1-d_l)}$. The transition dynamics are characterized in the following proposition.

Proposition 3. *When $t = 1$, the numbers of workers in both layers remain the same ($n^{h*}(q_0)$ and $n^{l*}(q_0)$). When $t = 2$, the number of seniors increases while the number of juniors drops. When $t > 2$, the pattern of the transition is summarized as follows:*

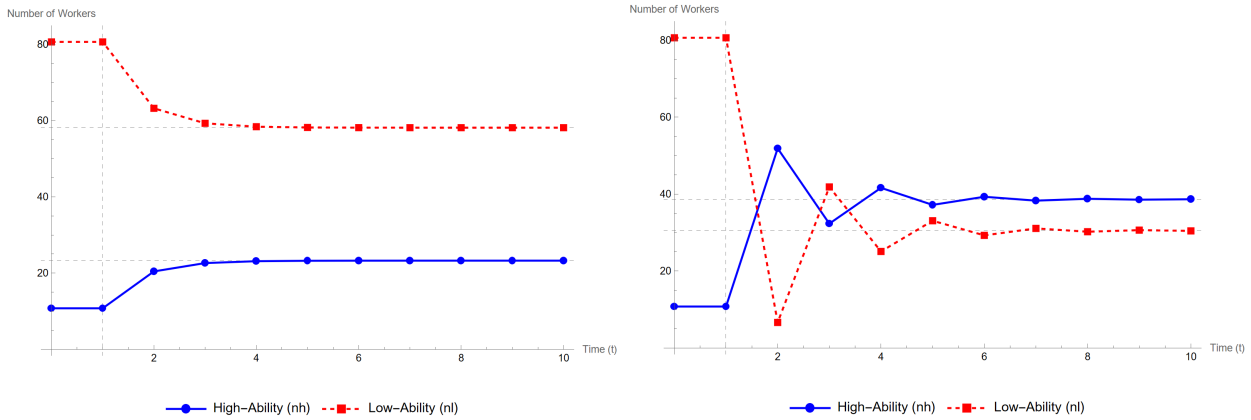
(*i*) *If $q_{AI} < \frac{\gamma^l(1-d_h)}{\gamma^h(1-d_l)}$, the number of seniors rises monotonically to the new steady state $n^{h*}(q_{AI})$, while the number of juniors falls monotonically to $n^{l*}(q_{AI})$.*

(*ii*) *If $q_{AI} > \frac{\gamma^l(1-d_h)}{\gamma^h(1-d_l)}$, the numbers of workers in both layers converge to their new steady states by alternating above and below them in opposite directions (i.e., $[n_t^h - n^{h*}(q_{AI})][n_t^l - n^{l*}(q_{AI})] < 0$).*

Proposition 3 shows that, unlike the productivity shock, the learning shock has a delayed effect. The numbers of workers in each layer remain the same in the first period and change

[5]With the same reason, we refrain the learning speed from growing too fast so that the firm can maintain the three-layer structure. Otherwise, the firm may not have the junior layer for several periods.

over time later. The transition pattern depends on the severity of the shock. In particular, if the shock is moderate ($q_{AI} < \frac{\gamma^l(1-d_h)}{\gamma^h(1-d_l)}$), then the transition will be smooth. However, if the shock is severe ($q_{AI} > \frac{\gamma^l(1-d_h)}{\gamma^h(1-d_l)}$), then we will observe the oscillation pattern.



(**a**) Smooth Adjustment: $q_{AI} = 0.3$.

(**b**) Oscillations: $q_{AI} = 0.95$.

Figure 3: Transition Dynamics with Positive Learning Shock. Parameters: $d_l = 0.4$, $d_h = 0.45$, $q_0 = 0.1$, $\gamma^h = 1.8$, $\gamma^l = 1$, and $N = 100$.

Figure 3 illustrates the transition patterns. After the initial shock, juniors learn faster from working, leading to a higher promotion rate at the end of the first period. This increases the stock of senior workers at $t = 2$ and leads to a reduced demand for junior workers. Unlike the productivity shock, after the initial learning shock, the number of juniors may not bounce back. This depends on the severity of the learning shock.

If the shock is moderate, the learning rate increases a bit, leading to a slight increase in the supply of seniors. In this case, we will observe a smooth adjustment shown in Figure 3a. Because of the moderate shock, the number of seniors will not jump dramatically to surpass the new steady state. Therefore, although the number of juniors goes down, this will not break down the talent pipeline. The juniors are sufficient to generate enough promotions to sustain the growth of the senior stock. Consequently, the senior stock glides up gradually to its new steady state.

In contrast, if the shock is severe, the learning rate increases a lot, leading to a significant increase in the supply of seniors. In this case, we will observe oscillations shown in Figure 3b. This severe learning shock brings a large wave of promotion, leading to a surge in the senior stock at $t = 2$. Facing a large number of productive seniors, the firm will reduce the number of juniors severely, breaking down the talent pipeline. When a large group of seniors depart in the next period, the firm is suddenly in a shortage of productivity and is forced to hire a surge of juniors to achieve the production target. This surge of juniors, in turn, pushes the number of seniors back above the steady state in next period. The

organization thus oscillates around the new steady state until convergence. This again implies that a firm, which had a pyramid structure before the shock, may oscillate between a diamond shape and a pyramid shape for some periods after the shock. This result also shows a risk of learning shock: if junior workers learn too fast and are promoted too quickly, it may also deplete the talent pipeline needed for the future.

5 General Equilibrium

So far, we have assumed that firms can neither hire nor fire senior workers, relying instead on internal promotion and attrition. We now relax these assumptions to show that our results on organizational structure and transitions hold in general equilibrium.

5.1 Market Setup

Consider an economy with a continuum of firms with mass M.

Workers There is an infinite supply of junior workers with outside option $\underline{v}$. Both juniors and seniors exit the market exogenously at rates d_l and d_h, returning to the same outside option $\underline{v}$. Let v_t^l and v_t^h denote the value functions for employed juniors and seniors:

$$(BE^{GE}) \qquad \begin{aligned} v_t^l &= w_t^l + \delta(1-d_l)[qv_{t+1}^h + (1-q)v_{t+1}^l] + \delta d_l \underline{v}, \\ v_t^h &= w_t^h + \delta(1-d_h)v_{t+1}^h + \delta d_h \underline{v}. \end{aligned}$$

Because of the infinite supply of juniors, free entry drives the value of being an employed junior down to the outside option: $v_t^l = \underline{v}$.

Firms Firms operate with the same production technology as defined in Section 2. Each firm i hires $n_{i,t}^h$ seniors and $n_{i,t}^l$ juniors to complete a fixed target of N tasks. Because the revenue per task greatly exceeds the cost of labor, firms always hire enough capacity to complete the target N. Consequently, each firm faces the capacity constraint:

$$(CC^{GE}) \qquad \gamma^h n_{i,t}^h + \gamma^l n_{i,t}^l \geq N.$$

Firms are forward-looking and maximize the present discounted value of profits. The firm's problem is:

$$\max_{\{n_{i,t}^l, n_{i,t}^h\}_{t=0}^{\infty}} \quad \sum_{t=0}^{\infty} \delta^t \left(Ny - w_t^h n_{i,t}^h - w_t^l n_{i,t}^l \right), \tag{1}$$

subject to the capacity constraint (CC^{GE}) for all t. Unlike the baseline model, firms here can hire senior workers from a competitive market. This market access eliminates the internal talent constraint.

Market Clearing Let N_t^h and N_t^l denote the aggregate supply of senior and junior workers. The aggregate supply of senior workers is affected by the human capital accumulation in the last period:

$$(FC^{GE}) \qquad N_{t+1}^h = (1 - d_h) N_t^h + q(1 - d_l) N_t^l.$$

Market clearing requires that aggregate demand from firms equals the aggregate supply of workers:

$$\int_0^M n_{i,t}^h di = N_t^h, \quad \int_0^M n_{i,t}^l di = N_t^l.$$

Equilibrium A competitive equilibrium consists of wage sequences $\{w_t^l, w_t^h\}_{t=0}^{\infty}$, firm hiring plans $\{n_{i,t}^l, n_{i,t}^h\}_{t=0}^{\infty}$, and aggregate labor stocks $\{N_t^l, N_t^h\}_{t=0}^{\infty}$ such that: Given wages, firm hiring plans maximize the present discounted value of profits (maximization problem (1)); Value functions satisfy the Bellman equations (BE^{GE}), and the free-entry condition holds; Aggregate demand equals aggregate supply; Aggregate supply evolves according to the flow equation (FC^{GE}).

5.2 Firm Optimization

The access to a competitive labor market allows firms to adjust their workforce size and composition immediately. This flexibility decouples the firm's decisions across periods: a hiring choice in period t does not constrain the firm in period $t+1$. This reduces the dynamic maximization problem (1) to a sequence of static problems:

$$\max_{n_{i,t}^l, n_{i,t}^h} \quad Ny - w_t^h n_{i,t}^h - w_t^l n_{i,t}^l \quad \text{subject to} \quad \gamma^h n_{i,t}^h + \gamma^l n_{i,t}^l \geq N.$$

The firm's production technology is linear, making junior and senior workers perfect substitutes. This linearity imposes an equilibrium condition on wages for firms to hire both types of workers.

Lemma 3. *In an equilibrium where firms employ strictly positive quantities of both senior and junior workers ($n^h_{i,t} > 0$ and $n^l_{i,t} > 0$), the ratio of wages must equal the ratio of productivities:*

$$\frac{w^h_t}{w^l_t} = \frac{\gamma^h}{\gamma^l}.$$

Lemma 3 shows a no-arbitrage condition $\frac{w^h_t}{w^l_t} = \frac{\gamma^h}{\gamma^l}$. Because the firm can instantly substitute γ^h/γ^l junior workers for one senior worker, the market wage premium for seniors must exactly match their productivity premium.

5.3 Stationary Equilibrium

Without loss of generality, we focus on a symmetric equilibrium in which aggregate supplies and wages remain constant. In this equilibrium, all firms make identical hiring decisions, and the aggregate demand for labor equals the aggregate supply determined by the economy's human capital accumulation technology.

Lemma 4. *The stationary symmetric equilibrium is characterized as follows:*

(i) *The equilibrium span of control and employment levels are:*

$$s^* = \frac{d_h}{q(1-d_l)}, \quad n^{l*} = \frac{s^*}{\gamma^l s^* + \gamma^h} N \quad \textit{and} \quad n^{h*} = \frac{1}{\gamma^l s^* + \gamma^h} N.$$

(ii) *The equilibrium wages are:*

$$w^{l*} = \underline{v}(1-\delta)\left[\frac{1-\delta(1-d_h)+\delta q(1-d_l)}{1-\delta(1-d_h)+\frac{\gamma^h}{\gamma^l}\delta q(1-d_l)}\right] \quad \textit{and} \quad w^{h*} = \frac{\gamma^h}{\gamma^l} w^{l*}.$$

The structural results in parts (i) mirror those in the single-firm analysis (Lemma 2). Even though firms in the general equilibrium framework have access to an external labor market, the aggregate flow constraint on human capital accumulation (FC^{GE}) pins down the same span of control, the ratio of juniors to seniors s^*. With the workforce composition fixed by this aggregate supply, the firm determines the absolute hiring levels (n^{l*}, n^{h*}) by scaling this specific mix of workers until their combined productivity meets the production target N.

While the organizational structure remains consistent with the partial equilibrium, the general equilibrium framework endogenizes wages (part ii). The junior wage w^l is anchored by the free-entry condition, which equates the value of a junior job to the outside option $\underline{v}$. However, w^l is not simply $\underline{v}(1-\delta)$ (the flow equivalent of the outside option). Instead, w^l is discounted relative to the outside option. This discount arises because junior workers accept lower current pay in exchange for the "option value" of acquiring skills and earning the senior wage w^h in the future.

The senior wage w^h is linked strictly to productivity. Because firms can instantly substitute between junior and senior labor, the no-arbitrage condition requires the senior wage premium to exactly match the productivity premium ($w^h/w^l = \gamma^h/\gamma^l$). If this condition failed, demand for one type of labor would collapse, breaking the equilibrium.

5.4 Equilibrium Analysis of AI

We now analyze how AI shocks affect the economy. We again consider three effects: a senior productivity shock (γ^h increases), a junior productivity shock (γ^l increases), and a learning shock (q increases). As before, we first analyze the steady state and then the transition path.

AI Effects on the Steady State In the long run, the impact of AI on employment levels (n^{l*}, n^{h*}) mirrors the single-firm analysis (Proposition 1). The general equilibrium framework, however, shows how AI changes wages.

Proposition 4. *In the steady state:*

(i) *If senior productivity γ^h increases, the junior wage w^l decreases, while the senior wage w^h increases.*

(ii) *If junior productivity γ^l increases, the junior wage w^l increases, while the senior wage w^h decreases.*

(iii) *If the learning probability q increases, both the junior wage w^l and the senior wage w^h decrease, but the ratio $\frac{w^h}{w^l}$ stays constant.*

Parts (i) and (ii) follow the no-arbitrage condition and the junior's career prospects. From Lemma 4, wages are driven by the productivity ratio γ^h/γ^l. When seniors become more productive (γ^h increases), this gap widens and the senior wage (w^h) rises. A higher senior wage, in turn, raises the continuation value of the junior position. Because of the higher continuation value and the free entry of juniors, firms offer lower junior wages (w^l).

If, instead, juniors become more productive (γ^l increases), the productivity gap narrows. This decreases the senior wage (w^h). A lower senior wage, in turn, decreases the continuation value of the junior position. Because of the lower continuation value and the free entry of juniors, firms must compensate juniors with higher current wages (w^l).

Part (iii) shows how faster learning reduces the scarcity of seniors. If AI accelerates learning (q increases), it increases the continuation value of a junior seat. Because of free entry, juniors accept lower wages (w^l) today. As juniors become cheaper, the senior wage must also decrease to satisfy the no-arbitrage condition.

AI Effects on the Transition Path We now turn to the transition path following an AI shock. This path depends on how the aggregate supply of senior workers evolves over time. To determine this evolution, we substitute the aggregate capacity constraint ($N_t^l = [MN - \gamma^h N_t^h]/\gamma^l$) into the aggregate flow equation (FC^{GE}). This substitution yields the flow equation for the symmetric firm's senior workforce:

$$n_{t+1}^h = \left[1 - d_h - \frac{\gamma^h}{\gamma^l} q(1 - d_l)\right] n_t^h + q(1 - d_l)\frac{N}{\gamma^l}.$$

This difference equation is identical to the one governing the single firm in Section 4. Because the mathematical structure corresponds exactly to the partial equilibrium case, the transition path dynamics, including the potential for oscillations, mirror those established in Propositions 2 and 3.[6]

6 General Equilibrium with Firm-Specific Human Capital

The previous section assumed that senior workers are perfect substitutes across firms. However, internal experience often creates firm-specific human capital (such as knowledge of workflows, internal protocols, or informal relationships) that external hires lack. We now extend the framework to include firm-specific human capital.

[6] If we consider a "pioneer" firm that adopts AI while others do not, then this firm faces fixed market wages determined by the non-AI firms. In this case, if AI increases senior productivity γ^h, the pioneer firm minimizes costs by hiring only seniors ($n^h = N/\gamma^h, n^l = 0$). If AI increases junior productivity γ^l, it hires only juniors ($n^l = N/\gamma^l, n^h = 0$). In these corner solutions, the firm specializes entirely in the labor type that offers the highest return per wage unit.

6.1 Market Setup

We distinguish between two types of senior workers. Internal seniors (promoted from within) possess specific capital and have productivity γ^h. External seniors (hired from the outside market) lack this capital and produce only $\gamma^h - \sigma$, where $\sigma > 0$. We assume $\gamma^h - \sigma > \gamma^l$, so that external seniors are still more productive than juniors.

Although productivity differs, the labor market remains competitive. Because seniors lose their specific capital upon switching firms, their outside option is the wage for generic external seniors. Firms therefore pay the market wage w_t^h to all seniors, extracting the rent σ generated by the internal ones.

Let $n_{i,t}^{h,in}$ denote the number of internal seniors, $n_{i,t}^{h,ex}$ the external seniors, and $n_{i,t}^{l}$ the juniors employed by firm i in period t. The firm maximizes the present discounted value of profits:

$$\max_{\{n_{i,t}^{l}, n_{i,t}^{h,in}, n_{i,t}^{h,ex}\}_{t=0}^{\infty}} \quad \sum_{t=0}^{\infty} \delta^t \left[Ny - w_t^h (n_{i,t}^{h,in} + n_{i,t}^{h,ex}) - w_t^l n_{i,t}^{l} \right],$$

subject to the capacity constraint

$$\gamma^h n_{i,t}^{h,in} + (\gamma^h - \sigma) n_{i,t}^{h,ex} + \gamma^l n_{i,t}^{l} \geq N,$$

and the talent constraint[7]

$$n_{i,t+1}^{h,in} \leq (1 - d_h)(n_{i,t}^{h,in} + n_{i,t}^{h,ex}) + q(1 - d_l) n_{i,t}^{l}.$$

The talent constraint distinguishes this problem from the benchmark general equilibrium. In the benchmark, periods were independent. Here, hiring a junior in period t generates the future value of a high-productivity senior in period $t + 1$.

6.2 Stationary Equilibrium

Without loss of generality, again, we focus on the symmetric equilibrium. In this setting, firms fill senior positions entirely from within.

Lemma 5. *In a stationary symmetric equilibrium, firms hire seniors exclusively through internal promotions* ($n^{h,ex} = 0$)*, and capacity and talent constraints bind.*

Lemma 5 shows that in a stationary symmetric equilibrium, the firm only hires from its internal labor market. Internal seniors possess specific capital, making them more

[7] Note that, we make an assumption here that the external seniors will become internal seniors in the next period.

productive than external hires. In an equilibrium, if firms employ external seniors, other firms must forgo some internal seniors, losing the rent from firm-specific human capital. Instead, this firm can retain all its internal seniors and then reduce its total workforce to match the target output N, strictly increasing profits. Thus, no external hiring occurs and the capacity constraint binds. Moreover, the talent constraint binds because otherwise there would be waste of talent.

In this equilibrium, with firm-specific human capital, the wage ratio between senior and junior includes not just current productivity, but the discounted value of the specific capital σ generated in the future.

Lemma 6. *In a stationary symmetric equilibrium, the wage ratio is given by:*

$$\frac{w^h}{w^l} = \frac{(\gamma^h - \sigma) + \delta(1-d_h)\sigma}{\gamma^l + \delta q(1-d_l)\sigma} =: \Gamma.$$

Lemma 6 shows that Γ generalizes the productivity ratio established in the previous section by incorporating two features. First, the senior wage is tied to the external market. While our earlier general equilibrium benchmark implied this relationship, the inclusion of specific human capital makes it clear that the senior wage must track the productivity of external hires to satisfy the no-arbitrage condition. Second, the ratio incorporates future value into current compensation. By adding the discounted term $\delta(1-d_h)\sigma$ to the numerator, we account for the future benefit of retaining an external senior (who becomes internal in the future). Correspondingly, the denominator also adds the future value of promotion $\delta q(1-d_l)\sigma$ for juniors. By including these terms, the ratio captures how today's wages reflect the long-term career path within the firm.

With the wage ratio of labor Γ, we characterize the equilibrium hiring and wage levels which mirrors those in the last section (Lemma 4).

Lemma 7. *The stationary symmetric equilibrium is characterized as follows:*

(i) *The equilibrium span of control and the equilibrium employment levels are:*

$$s^* = \frac{d_h}{q(1-d_l)}, \quad n^{l*} = \frac{s^*}{\gamma^l s^* + \gamma^h} N, \quad n^{h,in*} = \frac{1}{\gamma^l s^* + \gamma^h} N, \quad \textit{and} \quad n^{h,ex*} = 0.$$

(ii) *The equilibrium wages are:*

$$w^{l*} = \underline{v}(1-\delta)\left[\frac{1-\delta(1-d_h)+\delta q(1-d_l)}{1-\delta(1-d_h)+\Gamma\delta q(1-d_l)}\right] \quad \textit{and} \quad w^{h*} = \Gamma w^{l*}. \tag{2}$$

6.3 Equilibrium Analysis of AI

We now analyze how AI shocks affect the economy with firm-specific capital. We consider the same three effects: senior productivity (γ^h), junior productivity (γ^l), and learning (q).

Because firms in a stationary symmetric equilibrium hire exclusively from within, the AI shock affects the long-run organizational structure ($n^{h,in*}, n^{l*}, s^*$) and wages levels (w^{l*} and w^{h*}) as it does in the baseline model (Proposition 1 and Proposition 4). The impact on wages ratio, however, is different. In the general equilibrium without specific human capital, a learning shock left the wage ratio w^h/w^l constant. Here, faster learning compresses the wage gap.

Corollary 1. *In the steady state, if the learning probability q increases, the ratio between senior wage and junior wage $\frac{w^{h*}}{w^{l*}}$ decreases.*

This compression occurs because the juniors provide not just current production but also future firm-specific capital. A higher learning rate accelerates the creation of this capital, increasing the value of the junior layer relative to the seniors. Consequently, as the juniors become more valuable, the relative wage premium for senior workers declines.

The steady state identifies the destination, but the transition path depends on the firm's hiring strategy over time. The following lemma characterizes the optimal hiring for the firms in a symmetric equilibrium.

Lemma 8. *In a symmetric equilibrium, firms rely exclusively on internal promotion ($n_t^{h,ex} = 0$). Both the capacity and talent constraints bind throughout the transition.*

Lemma 8 shows that, along the transition path after the shock, firms prefer internal hiring. This is because promoted juniors possess specific capital σ, making them more productive than external seniors. Moreover, firms also refrain from hiring more workers than necessary to meet the current production target just for the future value. Because the revenue from any task completed beyond N is zero, the current marginal revenue of an extra worker is zero. Although a junior worker provides future value by filling the talent pipeline, this future value is offset by the worker's current wage. As a result, the firm minimizes costs by keeping the capacity constraint and talent constraint binding at every period.

With the firm constrained by its internal pipeline and a fixed output target, the evolution of the senior workforce follows a deterministic path. By substituting the aggregate capacity

constraint into the flow equation, we obtain the flow equation for senior workers:

$$n_{t+1}^{h,in} = \left[1 - d_h - \frac{\gamma^h}{\gamma^l} q(1 - d_l)\right] n_t^{h,in} + q(1 - d_l)\frac{N}{\gamma^l}.$$

This difference equation is identical to the one derived in Section 4. So the dynamics established in Propositions 2 and 3 apply here without modification.

The transition analysis shows that AI shocks can cause the organizational structure to oscillate, potentially creating periods where the internal talent pipeline is thin. One might be interested in how these structural disruptions affect the firm's value during the transition. To address this, we derive the firm's value function in the symmetric equilibrium.

Lemma 9. *In the stationary symmetric equilibrium, the value of a firm with $n_t^{h,in}$ internal senior workers is given by:*

$$V(n_t^{h,in}) = \lambda\sigma n_t^{h,in} + \frac{N(y - \lambda)}{1 - \delta},$$

where $\lambda = \frac{w^{l*}}{\gamma^l + \delta\sigma q(1-d_l)}$, with w^{l*} is the equilibrium junior wage in (2).

Lemma 9 decomposes the firm's value into two components, the internal talent value $\lambda\sigma n_t^{h,in}$ and the project value $\frac{N(y-\lambda)}{1-\delta}$. The project value represents the discounted stream of rents from the production target N, given the market cost of capacity λ. The talent value represents the premium the firm extracts from its current internal workforce $n_t^{h,in}$, who possess specific capital σ that external hires lack.

Next, we analyze the dynamic impact of AI on firm value throughout the transition.

Proposition 5. *Suppose that productivity (γ^h or γ^l) or the learning probability (q) permanently increases at $t = 1$.*

(i) At $t = 1$ the firm's value strictly increases ($V(n_1^{h,in}) > V(n_0^{h,in})$) .

(ii) The marginal value of an internal senior, $\lambda\sigma$, falls.

(iii) During the transition ($t > 1$), the firm's value $V(n_t^{h,in})$ comoves with the senior stock $n_t^{h,in}$.

Proposition 5 shows that the AI shock strictly increases the firm's value at $t = 1$, despite the chaotic transition. Because the market cost of capacity falls forever, the project value jumps immediately. This gain outweighs the temporal pipeline breakdown, the future shortage of seniors caused by hiring freezes or layoffs of juniors today. Yet, the transition entails a temporary inefficiency: when the senior stock dips below the new steady state, the firm captures fewer specific capital rents (σ).

7 Concluding Remarks

Despite its simplicity, our model captures some typical features of ILM, in particular, the need of firms to employ juniors to fill the talent pipeline. Our model is not a typical knowledge hierarchy as in Garicano (2000) and Garicano and Rossi-Hansberg (2015). It is rather a top-down hierarchy in which a CEO implements strategy by allocating tasks to agents in the different hierarchical layers, or, equivalently, delegates work packages that seniors can delegate further to the junior workers. The model also allows for thinking about the different ways in which AI affects workplaces, individual productivity gains on different layers and the speed of learning (here, modeled through the short cut of how likely a junior person acquires the skills needed to become senior). We have abstracted from incentive issues that are certainly an important element to talent management [8] and may need complementary analysis. We hope though that this model is a good start to think about these important issues and would see it as an advantage that it can readily be embedded in a labor market.

References

Auriol, Emmanuelle, Guido Friebel, and Frauke Von Bieberstein. 2016. "The firm as the locus of social comparisons: Standard promotion practices versus up-or-out." *Journal of Economic Behavior & Organization*, 121: 41–59. 3

Baker, George, Michael Gibbs, and Bengt Holmstrom. 1994*a*. "The internal economics of the firm: Evidence from personnel data." *The Quarterly Journal of Economics*, 109(4): 881–919. 1, 2

Baker, George, Michael Gibbs, and Bengt Holmstrom. 1994*b*. "The wage policy of a firm." *The Quarterly Journal of Economics*, 109(4): 921–955. 2

Bar-Isaac, Heski, and Clare Leaver. 2021. "Investment in Human Capital Under Endogenous Asymmetric Information." 20

Beckmann, Martin. 1978. *Rank in organizations.* Berlin:Springer. 3

Bentley University, and Gallup. 2025. "Bentley-Gallup Business in Society Survey: 2025 Report." Gallup, Inc. Accessed: January 14, 2026. 1

[8] There is a large literature on this, for instance, Waldman (1984); Milgrom and Oster (1987); Carmichael (1988); Fairburn and Malcomson (1994); Prendergast and Topel (1996); Bar-Isaac and Leaver (2021). Recently, many researchers investigate more on the role of manager's incentives (Hoffman and Tadelis, 2021; Friebel, Heinz and Zubanov, 2022; Minni, 2023; Haegele, 2024; Raith and Friebel, 2026)

Bianchi, Nicola, Giulia Bovini, Jin Li, Matteo Paradisi, and Michael Powell. 2023. "Career spillovers in internal labour markets." *The Review of Economic Studies*, 90(4): 1800–1831. 3

Brynjolfsson, Erik, Danielle Li, and Lindsey Raymond. 2025. "Generative AI at work." *The Quarterly Journal of Economics*, 140(2): 889–942. 3

Carmichael, H Lorne. 1988. "Incentives in academics: why is there tenure?" *Journal of political Economy*, 96(3): 453–472. 20

Chen, Yvonne Jie, Jie Gong, Jin Li, and Zibo Zhao. 2025. "Better Technology, Worse Motivation: GenAI's Mediocrity Trap." *Available at SSRN 5208163*. 3

Doeringer, Peter B., and Michael J. Piore. 1971. *Internal Labor Markets and Manpower Analysis.* Lexington, Mass:Heath. 2

Duncan, David S., Tyler Anderson, and Jeffrey Saviano. 2025. "AI Is Changing the Structure of Consulting Firms." *Harvard Business Review*. 1

Fairburn, James A, and James M Malcomson. 1994. "Rewarding performance by promotion to a different job." *European Economic Review*, 38(3-4): 683–690. 20

Friebel, Guido, and Elena Panova. 2008. "Insider Privatization and Careers: A Clinical Study of a Russian Firm in Transition." In *The Analysis of Firms and Employees: Quantitative and Qualitative Approaches*. , ed. Stefan Bender, Julia Lane, Kathryn L. Shaw, Fredrik Andersson and Till von Wachter, 253–266. Chicago, IL:University of Chicago Press. 3

Friebel, Guido, Matthias Heinz, and Nikolay Zubanov. 2022. "Middle managers, personnel turnover, and performance: A long-term field experiment in a retail chain." *Management Science*, 68(1): 211–229. 20

Garicano, Luis. 2000. "Hierarchies and the Organization of Knowledge in Production." *Journal of political economy*, 108(5): 874–904. 20

Garicano, Luis, and Esteban Rossi-Hansberg. 2015. "Knowledge-based hierarchies: Using organizations to understand the economy." *Annual Review of Economics*, 7(1): 1–30. 20

Gibbons, Robert. 1997. "Incentives and Careers in Organizations." In *Advances in Econometrics: Theory and Applications*. Vol. 2, , ed. David M. Kreps and Kenneth J. Wallis, Chapter 1, 1–37. Cambridge:Cambridge University Press. 2

Gibbons, Robert, and Michael Waldman. 1999*a*. "Careers in organizations: Theory and evidence." In *Handbook of Labor Economics*. Vol. 3, 2373–2437. Elsevier. 2

Gibbons, Robert, and Michael Waldman. 1999*b*. "A theory of wage and promotion dynamics inside firms." *The Quarterly Journal of Economics*, 114(4): 1321–1358. 3

Haegele, Ingrid. 2024. "Talent hoarding in organizations." *arXiv preprint arXiv:2206.15098*. 20

Hoffman, Mitchell, and Steven Tadelis. 2021. "People management skills, employee attrition, and manager rewards: An empirical analysis." *Journal of political economy*, 129(1): 243–285. 20

Huitfeldt, Ingrid, Andreas R Kostøl, Jan Nimczik, and Andrea Weber. 2023. "Internal labor markets: A worker flow approach." *Journal of Econometrics*, 233(2): 661–688. 3

Ke, Rongzhu, Jin Li, and Michael Powell. 2018. "Managing careers in organizations." *Journal of Labor Economics*, 36(1): 197–252. 3

Kräkel, Matthias, and Anja Schöttner. 2012. "Internal labor markets and worker rents." *Journal of Economic Behavior & Organization*, 84(2): 491–509. 3

Lazear, Edward P., and Paul Oyer. 2013. "Personnel Economics." In *Handbook of Organizational Economics*. , ed. Robert Gibbons and John Roberts, Chapter 12, 479–519. Princeton, NJ:Princeton University Press. 2

Lazear, Edward P, and Sherwin Rosen. 1981. "Rank-order tournaments as optimum labor contracts." *Journal of Political Economy*, 89(5): 841–864. 3

MacLeod, W Bentley, and James M Malcomson. 1988. "Reputation and hierarchy in dynamic models of employment." *Journal of Political Economy*, 96(4): 832–854. 3

Malcomson, James M. 1984. "Work incentives, hierarchy, and internal labor markets." *Journal of political economy*, 92(3): 486–507. 3

Milgrom, Paul, and Sharon Oster. 1987. "Job discrimination, market forces, and the invisibility hypothesis." *The Quarterly Journal of Economics*, 102(3): 453–476. 20

Minni, Virginia. 2023. "Making the invisible hand visible: Managers and the allocation of workers to jobs." 20

Osterman, Paul. 2024. "Reconsidering occupational internal labor markets: Incidence and consequences." *ILR Review*, 77(3): 366–395. 3

Pastorino, Elena. 2024. "Careers in firms: The role of learning about ability and human capital acquisition." *Journal of Political Economy*, 132(6): 1994–2073. 3

Prendergast, Canice. 1993. "The role of promotion in inducing specific human capital acquisition." *The Quarterly Journal of Economics*, 108(2): 523–534. 3

Prendergast, Canice, and Robert H Topel. 1996. "Favoritism in organizations." *Journal of Political Economy*, 104(5): 958–978. 20

Priest, Dan. 2025. "In the age of AI: Speed matters more, scale matters less, innovation matters most." *PwC Tech Effect*, Accessed: January 14, 2026. 1

Raith, Michael, and Guido Friebel. 2026. "Talent Management: The Role of Bosses." *Journal of Labor Economics*. Forthcoming. 20

Rosenbaum, James. 1984. *Career mobility in a corporate hierarchy.* New York:Academic Press. 3

Rosen, Sherwin. 1986. "Prizes and incentives in elimination tournaments." *American Economic Review*, 76(4): 701–715. 3

Sharwood, Simon. 2025. "AWS CEO says using AI to replace junior staff is 'Dumbest thing I've ever heard'." 1

Simon, Herbert A. 1951. "The analysis of promotional opportunities." *Personnel*, 27(4): 282–285. 3

Singla, Alex, Alexander Sukharevsky, Lareina Yee, Michael Chui, and Bryce Hall. 2025. "The state of AI: How organizations are rewiring to capture value." McKinsey & Company Survey. Accessed: January 13, 2026. 1

Stewman, Shelby, and Suresh Konda. 1983. "Careers and organizational labor markets: Demographic models of organizational behavior." *American Journal of Sociology*, 88(4): 637–685. 3

Waldman, Michael. 1984. "Job assignments, signalling, and efficiency." *The Rand journal of economics*, 15(2): 255–267. 20

Waldman, Michael. 2003. "Ex ante versus ex post optimal promotion rules: The case of internal promotion." *Economic Inquiry*, 41(1): 27–41. 3

Waldman, Michael. 2013. "Theory and evidence in internal labor markets." In *The handbook of organizational economics*. , ed. Robert Gibbons and John Roberts, Chapter 13, 520–571. Princeton University Press. 2

White, Harrison C. 1970. *Chains of opportunity: System models of mobility in organizations.* Harvard University Press. 3

Zabojnik, Jan, and Dan Bernhardt. 2001. "Corporate tournaments, human capital acquisition, and the firm Size—Wage relation." *The Review of Economic Studies*, 68(3): 693–716. 3

A Proof

Proof of Lemma 1. Consider the CEO's optimization problem:

$$\max_{n_t^l} \pi_t(n_t^h) = Ny - w^h n_t^h - w_l n_t^l$$
$$\text{s.t.} \quad \gamma^h n_t^h + \gamma^l n_t^l \geq N$$

Notice that $n_t^l \geq 0$. If $n_t^l > 0$, then the capacity constraint (CC) must be binding. Otherwise, we can safely reduce n_t^l by adjusting h_t^l and improve the payoff. Therefore, given the senior stock n_t^h, we must have

$$n_t^l = \max\{0, (N - \gamma^h n_t^h)/\gamma^l\}.$$

■

Proof of Lemma 2. From Lemma 1, we know that $n_t^l = \max\{0, (N - \gamma^h n_t^h)/\gamma^l\}$. We first show that in the steady state, $n_t^l > 0$. Suppose not, then $n_t^l = 0$. From the first equation of flow condition (FC), we have

$$n_{t+1}^h = (1 - d_h) n_t^h + 0 < n_t^h,$$

which contradicts the steady state constraints (SS). Therefore, $n^{l*} = (N - \gamma^h n^{h*})/\gamma^l > 0$. Combing (SS) with the first equation of (FC), we have

$$n^{h*} = (1 - d_h) n^{h*} + q(1 - d_l) n^{l*}.$$

After rearranging it, we can get

$$s^* = \frac{n^{l*}}{n^{h*}} = \frac{d_h}{q(1 - d_l)}.$$

Substituting n^{l*}, we can get

$$n^{h*} = (1 - d_h) n^{h*} + q(1 - d_l)[(N - \gamma^h n^{h*})/\gamma^l],$$

or $n^{h*} = \frac{q(1-d_l)}{\gamma^l d_h + \gamma^h q(1-d_l)} N = \frac{1}{\gamma^l s^* + \gamma^h} N$. Hence, $n^{l*} = (N - \gamma^h n^{h*})/\gamma^l = \frac{d_h}{\gamma^l d_h + \gamma^h q(1-d_l)} N = \frac{s^*}{\gamma^l s^* + \gamma^h} N$. ■

Proof of Proposition 1. Part (i): From Lemma 2, we have

$$\frac{\partial n^{h*}}{\partial \gamma^h} < 0, \quad \frac{\partial n^{l*}}{\partial \gamma^h} < 0, \quad \text{and} \quad \frac{\partial s^*}{\partial \gamma^h} = 0;$$
$$\frac{\partial n^{h*}}{\partial \gamma^l} < 0, \quad \frac{\partial n^{l*}}{\partial \gamma^l} < 0, \quad \text{and} \quad \frac{\partial s^*}{\partial \gamma^l} = 0.$$

Part (ii): From Lemma 2, we have

$$\frac{\partial n^{h*}}{\partial q} > 0, \quad \frac{\partial n^{l*}}{\partial q} < 0, \quad \text{and} \quad \frac{\partial s^*}{\partial q} < 0.$$

■

Before proving Proposition 2, we first prove the following lemma that characterizes the evolution of the firm's structure for $t \geq 1$.

Lemma A.1. *Suppose that $r^{\gamma}_{AI} < \frac{2-d_h}{q(1-d_l)}$. The number of the senior workers at time $t \geq 1$ is given by*

$$n^h_t = n^{h*}(\gamma^h_{AI}, \gamma^l_{AI}) + \Phi^{t-1}\left[n^{h*}(\gamma^h_0, \gamma^l_0) - n^{h*}(\gamma^h_{AI}, \gamma^l_{AI})\right],$$

where $\Phi := 1 - d_h - \frac{\gamma^h_{AI}}{\gamma^l_{AI}} q(1-d_l)$. The corresponding number of junior workers employed is given by

$$n^l_t = (N - \gamma^h_{AI} n^h_t)/\gamma^l_{AI}.$$

Proof of Lemma A.1. We first assume that $n^l_t > 0$ for all $t \geq 0$. We will verify this assumption later.

Given that $n^l_t > 0$, we have $n^l_t = (N - \gamma^h n^h_t)/\gamma^l$ directly from Lemma 1. Together with the first equation in (FC), for $t \geq 1$, we have

$$n^h_{t+1} = (1-d_h)n^h_t + q(1-d_l)[(N - \gamma^h_{AI} n^h_t)/\gamma^l_{AI}],$$

or

$$n^h_{t+1} = \underbrace{\left[1 - d_h - \frac{\gamma^h_{AI}}{\gamma^l_{AI}} q(1-d_l)\right]}_{\Phi} n^h_t + \frac{1}{\gamma^l_{AI}} q(1-d_l)N.$$

After rearranging, we have

$$n^h_{t+1} - n^{h*}(\gamma^h_{AI}, \gamma^l_{AI}) = \Phi[n^h_t - n^{h*}(\gamma^h_{AI}, \gamma^l_{AI})].$$

Notice that $r^{\gamma}_{AI} = \gamma^h/\gamma^l < \frac{2-d_h}{q(1-d_l)}$, we have $|\Phi| < 1$. Hence, the system will eventually

converge to the new steady state. Since the shock happens at $t = 1$, we have

$$n_1^h = n_0^h = n^{h*}(\gamma_0^h, \gamma_0^l).$$

Therefore, for $t \geq 1$ the transition of the number of seniors can be described as

$$n_t^h = n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l) + \Phi^{t-1}\left[n^{h*}(\gamma_0^h, \gamma_0^l) - n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l)\right].$$

Now we verify that $n_t^l > 0$. From Proposition 1, we know that $n^{h*}(\gamma_0^h, \gamma_0^l) > n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l)$. Since $n_1^h = n^{h*}(\gamma_0^h, \gamma_0^l)$ and $|\Phi| < 1$, we must have $n_t^h < n_1^h$ for all $t > 1$. Since $\gamma_{AI}^h < \gamma_0^h + \frac{d_h}{q(1-d_l)}\gamma_0^l$, we have $\gamma_{AI}^h n_1^h < N$. Therefore, for all $t \geq 1$, we must have

$$(N - \gamma_{AI}^h n_t^h)/\gamma_{AI}^l > 0.$$

This means $n_t^l = (N - \gamma_{AI}^h n_t^h)/\gamma_{AI}^l$ for $t \geq 1$. Since $n_0^l = n^{l*}(\gamma_0^h, \gamma_0^l) > 0$, we have $n_t^l > 0$ for all $t \geq 0$. ■

Proof of Proposition 2. **Step 1.** When $t = 1$, we have $n_1^h = n^{h*}(\gamma_0^h, \gamma_0^l)$ since the shock does not affect the talent flow in the last period. Hence, from Lemma A.1, we can calculate

$$n_1^l = (N - \gamma_{AI}^h n_1^h)/\gamma_{AI}^l.$$

It is easy to verify that $n_1^l < n^{l*}(\gamma_{AI}^h, \gamma_{AI}^l)$.

Step 2. When $t > 1$ and $r_{AI}^\gamma < \frac{1-d_h}{q(1-d_l)}$, we have $\Phi \in (0, 1)$. From Lemma A.1, we have

$$n_t^h - n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l) = \Phi^{t-1}\left[n^{h*}(\gamma_0^h, \gamma_0^l) - n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l)\right].$$

Since $n^{h*}(\gamma_0^h, \gamma_0^l) > n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l)$, we must have $n_t^h - n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l)$ decreases in t and is positive for all $t > 1$. This means n_t^h is decreasing over time and is always larger than $n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l)$. As a result, $n_t^l = (N - \gamma_{AI}^h n_t^h)/\gamma_{AI}^l$ is increasing over time and is always smaller than $n^{l*}(\gamma_{AI}^h, \gamma_{AI}^l)$.

Step 3. When $t > 1$ and $r_{AI}^\gamma > \frac{1-d_h}{q(1-d_l)}$, we have $\Phi \in (-1, 0)$. Following the same logic in Step 2, we have $n_t^h - n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l) > 0$ if and only if t is an odd number. Since $n_t^l = (N - \gamma_{AI}^h n_t^h)/\gamma_{AI}^l$, we have

$$n_t^l - n^{l*}(\gamma_{AI}^h, \gamma_{AI}^l) = -\frac{\gamma_{AI}^h}{\gamma_{AI}^l}[n_t^h - n^{h*}(\gamma_{AI}^h, \gamma_{AI}^l)].$$

Therefore, $n_t^l - n^{l*}(\gamma_{AI}^h, \gamma_{AI}^l) > 0$ if and only if t is an even number. Moreover, $[n_t^h -$

$n^{h*}(\gamma^h_{AI}, \gamma^l_{AI})][n^l_t - n^{l*}(\gamma^h_{AI}, \gamma^l_{AI})] < 0.$ ■

Before proving Proposition 3, we first prove the following that characterizes the evolution of the firm's structure for $t \geq 1$.

Lemma A.2. *Suppose that $q_{AI} < \frac{\gamma^l(2-d_h)}{\gamma^h(1-d_l)}$. The number of the senior workers at time $t \geq 1$ is given by*

$$n^h_t = n^{h*}(q_{AI}) + \Psi^{t-1}\left[n^{h*}(q_0) - n^{h*}(q_{AI})\right],$$

where $\Psi := 1 - d_h - \frac{\gamma^h}{\gamma^l} q_{AI}(1-d_l)$. The corresponding number of junior workers employed is given by

$$n^l_t = (N - \gamma^h n^h_t)/\gamma^l.$$

Proof of Lemma A.2. We first assume that $n^l_t > 0$ for all $t \geq 0$. We will verify this assumption later.

Given that $n^l_t > 0$, we have $n^l_t = (N - \gamma^h n^h_t)/\gamma^l$ directly from Lemma 1. Together with the first equation in (FC), for $t \geq 1$, we have

$$n^h_{t+1} = (1 - d_h) n^h_t + q_{AI}(1 - d_l)[(N - \gamma^h n^h_t)/\gamma^l],$$

or

$$n^h_{t+1} = \underbrace{\left[1 - d_h - \frac{\gamma^h}{\gamma^l} q_{AI}(1-d_l)\right]}_{\Psi} n^h_t + \frac{1}{\gamma^l} q_{AI}(1-d_l)N.$$

After rearranging, we have

$$n^h_{t+1} - n^{h*}(q_{AI}) = \Phi[n^h_t - n^{h*}(q_{AI})].$$

Notice that $q_{AI} < \frac{\gamma^l(2-d_h)}{\gamma^h(1-d_l)}$, we have $|\Psi| < 1$. Hence, the system will eventually converge to the new steady state. Since the shock happens at $t = 1$, we have

$$n^h_1 = n^h_0 = n^{h*}(q_0).$$

Therefore, for $t \geq 1$ the transition of the number of seniors can be described as

$$n^h_t = n^{h*}(q_{AI}) + \Psi^{t-1}\left[n^{h*}(q_0) - n^{h*}(q_{AI})\right].$$

Now we verify that $n^l_t > 0$. From Proposition 1, we know that $n^{h*}(q_0) < n^{h*}(q_{AI})$. Suppose that $\Psi > 0$, then we must have $n^h_t < n^{h*}(q_{AI})$ for all $t \geq 1$. This means $n^l_t > n^{l*}(q_{AI}) > 0$ for all $t \geq 1$. Suppose that $\Psi < 0$, then we must have $n^h_t > n^{h*}(q_{AI})$ if and

only if t is an even number. Given that $|\Psi| < 1$, we must have $n_t^h \leq n_2^h$ for all $t \geq 1$. Since $q_{AI} < q_0 + \frac{\gamma^l}{\gamma^h(1-d_l)}$, we have $\gamma^h n_2^h < N$. This means $n_t^l > 0$ for all $t \geq 1$. Therefore, for all $t \geq 1$, we must have

$$(N - \gamma^h n_t^h)/\gamma^l > 0.$$

This means $n_t^l = (N - \gamma^h n_t^h)/\gamma^l$ for $t \geq 1$. Since $n_0^l = n^{l*}(q_0) > 0$, we have $n_t^l > 0$ for all $t \geq 0$. ■

Proof of Proposition 3. **Step 1.** When $t = 1$, we have $n_1^h = n^{h*}(q_0)$ and $n_1^l = n^{l*}(q_0)$ since the shock does not affect the productivity in the current period and the talent flow in the last period.

Step 2. When $t = 2$, the talent pipeline is affected. From the first part of the flow conditions (FC), we have

$$n_2^h = (1 - d_h)n_1^h + q_{AI}(1 - d_l)n_1^l > (1 - d_h)n_1^h + q_0(1 - d_l)n_1^l = n^{h*}(q_0).$$

This means the number of seniors increases. Meanwhile, we have

$$n_2^l = (N - \gamma^h n_2^h)/\gamma^l < (N - \gamma^h n^{h*}(q_0))/\gamma^l = n^{l*}(q_0).$$

Hence, the number of juniors decreases.

Step 3. When $t > 2$ and $q_{AI} < \frac{\gamma^l(1-d_h)}{\gamma^h(1-d_l)}$, we have $\Psi \in (0, 1)$. From Lemma A.2, we have

$$n_t^h - n^{h*}(q_{AI}) = \Phi^{t-1}\left[n^{h*}(q_0) - n^{h*}(q_{AI})\right].$$

Since $n^{h*}(q_0) < n^{h*}(q_{AI})$, we must have $n_t^h - n^{h*}(q_{AI})$ increases in t and is negative for all $t > 1$. This means n_t^h is increasing over time and is always smaller than $n^{h*}(q_{AI})$. As a result, $n_t^l = (N - \gamma^h n_t^h)/\gamma^l$ is decreasing over time and is always greater than $n^{l*}(q_{AI})$.

Step 4. When $t > 2$ and $q_{AI} > \frac{\gamma^l(1-d_h)}{\gamma^h(1-d_l)}$, we have $\Psi \in (-1, 0)$. Following the same logic in Step 3, we have $n_t^h - n^{h*}(q_{AI}) > 0$ if and only if t is an even number. Since $n_t^l = (N - \gamma^h n_t^h)/\gamma^l$, we have

$$n_t^l - n^{l*}(q_{AI}) = -\frac{\gamma^h}{\gamma^l}[n_t^h - n^{h*}(q_{AI})].$$

Therefore, $n_t^l - n^{l*}(q_{AI}) > 0$ if and only if t is an odd number. Moreover, $[n_t^h - n^{h*}(q_{AI})][n_t^l - n^{l*}(q_{AI})] < 0$. ■

Proof of Lemma 3. Minimizing cost subject to the linear capacity constraint implies that the firm compares the marginal cost per unit of output for each worker type. The cost per task using seniors is w_t^h/γ^h, and the cost using juniors is w_t^l/γ^l. If $w_t^h/\gamma^h < w_t^l/\gamma^l$, the firm

sets $n_{i,t}^l = 0$. If $w_t^h/\gamma^h > w_t^l/\gamma^l$, the firm sets $n_{i,t}^h = 0$. An interior solution where both $n_{i,t}^h$ and $n_{i,t}^l$ are positive requires equality. ■

Proof of Lemma 4. For parts (i) and (ii), the expressions for s^*, n^{h*}, and n^{l*} follow directly from the binding capacity constraint and the aggregate flow equation (FC^{GE}) in the steady state. For part (iii), the expressions for w^{l*} and w^{h*} follow from solving the system of equations formed by the no-arbitrage condition (Lemma 3), the workers' Bellman equations, and the free-entry condition ($v^l = \underline{v}$). ■

Proof of Lemma 5. Suppose, by contradiction, that firms hire external seniors ($n^{h,ex} > 0$). In a symmetric equilibrium, the supply of external seniors corresponds to the seniors that firms fail to retain. Thus, a firm hiring $n^{h,ex}$ external seniors could explicitly choose to retain these workers instead.

Consider this alternative strategy: the firm sets $\tilde{n}^{h,in} = n^{h,in} + n^{h,ex}$ and $\tilde{n}^{h,ex} = 0$. This reallocation maintains the total senior workforce and the wage bill $w^h(n^{h,in} + n^{h,ex})$. However, because internal seniors possess firm-specific capital while external seniors do not, total capacity increases:

$$\gamma^h \tilde{n}^{h,in} + \gamma^l n^l > \gamma^h n^{h,in} + (\gamma^h - \sigma) n^{h,ex} + \gamma^l n^l \geq N.$$

The capacity constraint becomes slack. Because wages are positive, the firm can reduce its workforce (reduce $\tilde{n}^{h,in}$ by $\varepsilon > 0$ and n^l by $\frac{d_h}{q(1-d_l)}\varepsilon$) to restore the binding capacity constraint and keep the talent constraint satisfied. This reduction lowers costs and increases profits, contradicting the optimality of hiring external seniors. Thus, $n^{h,ex}$ must be zero.

We prove that the talent constraint binds when proving Lemma 7. ■

Proof of Lemma 6. Consider a firm maximizes the present discounted value of profits subject to the capacity constraint and the talent flow constraint. Let λ_t and μ_t denote the Lagrange multipliers for the capacity and talent constraints, respectively. The Lagrangian is:

$$\mathcal{L} = \sum_{t=0}^{\infty} \delta^t \left\{ \begin{array}{l} Ny - w^h(n_t^{h,in} + n_t^{h,ex}) - w^l n_t^l \\ + \lambda_t[\gamma^h n_t^{h,in} + (\gamma^h - \sigma) n_t^{h,ex} + \gamma^l n_t^l - N] \\ + \mu_t[(1-d_h)(n_t^{h,in} + n_t^{h,ex}) + q(1-d_l) n_t^l - n_{t+1}^{h,in}] \end{array} \right\}.$$

The first-order condition for junior workers n_t^l is:

$$\frac{\partial \mathcal{L}}{\partial n_t^l} = -w^l + \lambda_t \gamma^l + \mu_t q(1-d_l) = 0.$$

The first-order condition for external seniors $n_t^{h,ex}$ is:

$$\frac{\partial \mathcal{L}}{\partial n_t^{h,ex}} = -w^h + \lambda_t(\gamma^h - \sigma) + \mu_t(1 - d_h) \leq 0.$$

In an equilibrium where firms can freely access the external market, the condition holds with equality:

$$w^h = \lambda_t(\gamma^h - \sigma) + \mu_t(1 - d_h). \tag{3}$$

Otherwise, lowering the firm's own senior wage can strictly raise the profit of the firm.

The first-order condition for internal seniors $n_{t+1}^{h,in}$ is:

$$\frac{\partial \mathcal{L}}{\partial n_{t+1}^{h,in}} = -\mu_t + \delta\left[-w^h + \lambda_{t+1}\gamma^h + \mu_{t+1}(1 - d_h)\right] = 0.$$

In the stationary symmetric equilibrium, multipliers are constant ($\lambda_t = \lambda, \mu_t = \mu$). Rearranging yields:

$$\mu = \delta\left[\lambda\gamma^h + \mu(1 - d_h) - w^h\right].$$

Substitute the expression for w^h from (3) into the bracketed term:

$$\mu = \delta\left[\lambda\gamma^h + \mu(1 - d_h) - (\lambda(\gamma^h - \sigma) + \mu(1 - d_h))\right] = \delta\lambda\sigma.$$

This result shows that the shadow value of the talent constraint equals the discounted marginal value of the specific capital σ.

We now express wages in terms of λ. Substituting $\mu = \delta\lambda\sigma$ back into the wage equations:

$$w^h = \lambda(\gamma^h - \sigma) + (\delta\lambda\sigma)(1 - d_h) = \lambda[\gamma^h - \sigma + \delta\sigma(1 - d_h)],$$
$$w^l = \lambda\gamma^l + (\delta\lambda\sigma)q(1 - d_l) = \lambda[\gamma^l + \delta\sigma q(1 - d_l)].$$

The multiplier λ is positive because the capacity constraint binds. Dividing w^h by w^l eliminates λ and yields the result Γ. ■

Proof of Lemma 7. We first show that the talent constraint binds. Suppose, by contradiction, that the constraint is slack ($d_h n^{h,in} < q(1 - d_l)n^l$). The firm could substitute juniors for internal seniors. The firm can increase seniors by $\varepsilon > 0$ and decrease junior $\frac{\gamma^h}{\gamma^l}\varepsilon$ such

that it maintains the capacity. This alternative strategy changes the wage bill by

$$w^h\varepsilon - w^l\varepsilon\frac{\gamma^h}{\gamma^l} = w^l\varepsilon\left(\frac{\gamma^h - \sigma + \delta\sigma(1-d_h)}{\gamma^l + \delta\sigma q(1-d_l)} - \frac{\gamma^h}{\gamma^l}\right) < 0.$$

Thus, the firm profits by substituting juniors for seniors until the talent constraint binds.

With the binding talent constraint ($d_h n^{h,in} = q(1-d_l)n^l$) and the binding capacity constraint, the structural results (i) and (ii) follow algebraically. Part (iii) follows from substituting the ratio Γ into the value functions and applying the free-entry condition ($v^l = \underline{v}$). ■

Proof of Lemma 8. We first prove that, if there is a symmetric equilibrium involving external hiring, we can always construct a symmetric equilibrium that achieves the same profit for the firm without external hiring but strictly increase firm's capacity.

Suppose that firms hire external seniors during the transition ($n_t^{h,ex} > 0$). In a symmetric equilibrium, the supply of external seniors corresponds to the seniors that firms fail to retain. Thus, a firm hiring $n_t^{h,ex}$ external seniors could explicitly choose to retain these workers instead.

Consider this alternative strategy: the firm sets $\tilde{n}_t^{h,in} = n_t^{h,in} + n_t^{h,ex}$ and $\tilde{n}_t^{h,ex} = 0$. This reallocation maintains the total senior workforce and the wage bill $w_t^h(n_t^{h,in} + n_t^{h,ex})$. Furthermore, because the total number of seniors remains unchanged, the firm carries the exact same talent pool into the next period, satisfying all dynamic constraints. However, because internal seniors possess firm-specific capital while external seniors do not, total capacity increases. Therefore, it is without loss of generality to focus on the symmetric equilibrium with internal hiring only. We next prove that the slack capacity constraint implies that the original plan is suboptimal if the firm's optimal strategy requires the capacity constraint to bind.

To show that the capacity and talent constraints bind, we analyze the firm's dynamic optimization problem. The firm maximizes discounted profits subject to both constraints. Although firms hire only internal seniors in equilibrium, the availability of external hires determines the market wage w_t^h. Let λ_t be the Lagrange multiplier for the capacity constraint

and μ_t be the multiplier for the talent constraint. The Lagrangian is:

$$\begin{aligned}\mathcal{L} = \sum_{t=0}^{\infty} \delta^t \Bigg\{ & Ny - w_t^h n_t^{h,ex} - w_t^h n_t^{h,in} - w_t^l n_t^l \\ & + \lambda_t \left[(\gamma^h - \sigma) n_t^{h,ex} + \gamma^h n_t^{h,in} + \gamma^l n_t^l - N \right] \\ & + \mu_t \left[(1-d_h)(n_t^{h,in} + n_t^{h,ex}) + q(1-d_l) n_t^l - n_{t+1}^{h,in} \right] \Bigg\}.\end{aligned}$$

The first-order conditions for $n_t^{h,ex}$, n_t^l, and $n_{t+1}^{h,in}$ characterize the firm's optimal hiring:

$$\begin{aligned} -w_t^h + \lambda_t(\gamma^h - \sigma) + \mu_t(1-d_h) &= 0, \\ -w_t^l + \lambda_t \gamma^l + \mu_t q(1-d_l) &= 0, \\ \delta \left[-w_{t+1}^h + \lambda_{t+1}\gamma^h + \mu_{t+1}(1-d_h) \right] - \mu_t &= 0. \end{aligned}$$

These three conditions imply $\mu_t = \delta\lambda\sigma$ and

$$\begin{aligned} w_t^h &= \lambda_t(\gamma^h - \sigma) + \delta\sigma(1-d_h)\lambda_{t+1}, \\ w_t^l &= \lambda_t \gamma^l + \delta\sigma q(1-d_l)\lambda_{t+1}. \end{aligned}$$

That is, in a dynamic symmetric equilibrium, the no-arbitrage condition becomes $\frac{w_t^h}{w_t^l} = \frac{\lambda_t(\gamma^h-\sigma)+\delta\sigma(1-d_h)\lambda_{t+1}}{\lambda_t\gamma^l+\delta\sigma q(1-d_l)\lambda_{t+1}}$. In the linear economy with competitive labor markets, shadow prices adjust instantly to the new steady-state levels following a permanent shock: $\lambda_t = \lambda$. As a result, $\frac{w^h}{w^l} = \frac{(\gamma^h-\sigma)+\delta(1-d_h)\sigma}{\gamma^l+\delta q(1-d_l)\sigma}$.

Finally, we need to prove λ (and hence μ) is positive. This is done by solving the wages out from the the no-arbitrage condition, workers' value equations, and the free-entry condition as before. The expressions of the wages are the same with those in Lemma 7 except the γ^h, γ^l, and q changes to the new values after the shock. Because market wages are strictly positive, the shadow price λ (and consequently μ) must be strictly positive. By the Kuhn-Tucker conditions, $\lambda > 0$ and $\mu > 0$ imply that both constraints bind at the optimum. This concludes the proof. ■

Proof of Lemma 9. We conjecture a linear value function $V(n^{h,in}) = A + Bn^{h,in}$. Substituting this into the Bellman equation $V(n_t^{h,in}) = Ny - w^h n_t^{h,in} - w^l n_t^l + \delta V(n_{t+1}^{h,in})$ and using the wage expressions from Lemma 6, we match coefficients.

First, the binding capacity constraint implies $n_t^l = (N - \gamma^h n_t^{h,in})/\gamma^l$. The transition law is $n_{t+1}^{h,in} = \Phi n_t^{h,in} + C$, where Φ and C are constants defined by the binding flow constraint. Matching the slope coefficients yields $B = \lambda\sigma$. Matching the constant terms

yields $A = \frac{N(y-\lambda)}{1-\delta}$.

The expression λ follows from the junior wage equation derived in the proof of Lemma 6, $w^l = \lambda[\gamma^l + \delta\sigma q(1 - d_l)]$. ■

Proof of Proposition 5. For parts (i) and (ii), differentiating λ with respect to the parameters shows that $\frac{\partial\lambda}{\partial\gamma^h} < 0$, $\frac{\partial\lambda}{\partial\gamma^l} < 0$, and $\frac{\partial\lambda}{\partial q} < 0$. This implies the shadow cost of capacity falls with AI improvements.

Using the value function $V(n_t^{h,in}) = \lambda\sigma n_t^{h,in} + \frac{N(y-\lambda)}{1-\delta}$, we differentiate with respect to λ:

$$\frac{\partial V}{\partial \lambda} = \sigma n_t^{h,in} - \frac{N}{1-\delta}.$$

Since the firm's total capacity is N and internal seniors produce $\gamma^h > \sigma$, the total specific capital rent cannot exceed the total capacity value, so $\sigma n_t^{h,in} < N < \frac{N}{1-\delta}$. Thus, $\frac{\partial V}{\partial \lambda} < 0$. Because AI lowers λ, and lower λ increases V, the firm value increases instantly. Simultaneously, the marginal value of an incumbent is $\lambda\sigma$, which strictly falls as λ falls.

For part (iii), from Lemma 9, $V(n_t^{h,in})$ is linear and increasing in $n_t^{h,in}$ (since $\lambda\sigma > 0$). From the transition analysis (Propositions 2 and 3), we know $n_t^{h,in}$ can oscillate during the transition. Therefore, $V(n_t)$ can also oscillate. Specifically, if $n_t^{h,in} < n^{h,in*}$, then $V(n_t^{h,in}) < V(n^{h,in*})$. ■