Artificial Intelligence Investments and Expertise Erosion

Pierre Chaigneau
Queen's University

Evan Jo
Queen's University

September 4, 2025

Abstract

We study the feedback effects between automation and expertise. Building on Autor and Thompson's (2025) "expertise framework", we develop an overlapping generations model with endogenous on-the-job learning. Automation reduces expertise acquisition in entry-level jobs, which worsens the expertise distribution over time. A concern is that this might increase the value of subsequent improvements in technologies that substitute for human expertise, potentially leading to self-reinforcing advances in artificial intelligence that further erode expertise. We show, however, that this conjecture is incorrect. Intuitively, further improvements in automation raise the productivity of workers with high expertise, whose share declines with prior automation.

"AI tools are doing the types of simple coding and debugging tasks that junior software developers did to gain experience. AI is also doing work that young employees in the legal and retail sectors once did. And Wall Street firms are reportedly considering steep cuts to entry-level hiring." *Fortune, May 2025*

Artificial Intelligence (AI) progressively automates tasks that used to require expertise, starting with entry-level tasks. Recent empirical evidence suggests that AI is starting to replace some entry-level professional jobs (Berger et al. (2024), Brynjolfsson, Chandar, Chen (2025), Lichtinger and Hosseini (2025), Hampole et al. (2025)).[1] These jobs typically involve basic tasks, but also opportunities for learning that prepare employees for higher-level jobs. In this paper, we study how the deployment of powerful AI changes worker expertise, which in turn affects the value of future improvements in AI.[2]

To understand the effect of the automation of some tasks on human capital investments, it is important to understand how jobs are structured around tasks. In his famous description of the pin factory, Adam Smith explained how each worker has a specific task to accomplish. A worker specializes in a clearly circumscribed task, and only needs to be capable of accomplishing this task. Models of atomistic task assignment capture this division of labor (e.g. Autor, Levy, and Murnane (2003), Acemoglu and Autor (2011), Acemoglu and Restrepo, (2018), Acemoglu et al., (2024)). By contrast, professional services workers – such as lawyers, consultants, or financial advisors – are in charge of a multitude of tasks, so that each job can be viewed as a bundle of tasks (Autor and Handel (2013), Brynjolfsson, Mitchell, and Rock (2018)). Each worker needs broad expertise that allows her to accomplish all these tasks. This expertise is acquired during the educational process, but also on-the-job, via formal training and informal learning (Loewenstein and Spletzer (1994)).

Autor and Thompson (2025) develop an "expertise framework" based on a model of occupational task bundling which captures these features. In their model, workers are

[1]Several recent headlines also draw attention to this trend. They include: "For Some Recent Graduates, the A.I. Job Apocalypse May Already Be There" *New York Times* May 30 2025; "AI May Be Erasing Entry-Level Jobs" *Forbes*, June 1 2025; "AI is 'breaking' entry-level jobs that Gen Z workers need to launch careers, LinkedIn exec warns." *Fortune*, May 2025; "Anthropic CEO warns AI could eliminate half of all entry-level white-collar jobs" *Fortune*, May 28 2025. There are many examples of specific tasks that get automated. For instance, Kate Smaje, McKinsey's global leader of technology and AI pointed out: "Do we need armies of business analysts creating PowerPoints? No, the technology could do that" Source: "McKinsey leans on AI to make PowerPoints faster, draft proposals" *Fortune*, June 2 2025.

[2]Recent trends in Venture Capital (VC) investments are suggestive of large-scale AI development and deployment: "Early stage investments by the top VC firms in Q1 [2025] consolidated around AI agents and the underlying infrastructure to deploy them. With these high-profile backers, emerging AI startups are set to usher in a new era of enterprise operations." PitchBook newsletter, June 26 2025.

vertically differentiated in their expertise ("hierarchical expertise") and each occupation involves multiple tasks ("indivisibility"). They use this conceptual framework to study the effect of automation on the demand for labor. To do so, they postulate that technological progress allows to automate additional tasks, but to a different degree in various occupations. This model allows them to develop new predictions and explain puzzling empirical facts. In particular, they show that occupations similarly exposed to automation will nevertheless experience different wage and employment outcomes depending on whether automation increases or decreases the average expertise level in the remaining non-automated tasks.[3]

In Autor and Thompson (2025), workers' expertise is exogenous. To better understand the effect of automation on expertise generation, we extend this model to allow for endogenous expertise acquisition with on-the-job learning in an overlapping generations (OLG) model with young and old workers.

To do so, we make two assumptions on the expertise distribution. First, in the benchmark case without automation, worker expertise is uniformly distributed on $[0, 1]$, as in Autor and Thompson (2025). Second, we let a young worker whose occupation involves some non-automated expert tasks learn with probability p, in which case her expertise when old is uniformly distributed between her expertise when young and the highest level (1), and not learn with the complementary probability, in which case she has the same expertise when young and when old. Thus, the on-the-job learning process is simply described by one parameter – the probability of learning.

These assumptions allow us to derive the unique expertise distribution for young workers, which is a function of p. They imply that, without automation, the density of the expertise distribution of young (old) workers is a strictly decreasing (increasing) function of expertise.

We rely on the same learning process to incorporate the effects of automation on skills accumulation via on-the-job learning. When all expert tasks in an occupation are automated, workers do not learn as much in this occupation. Accordingly, we assume that the learning process is as described above, except that there is learning with probability $q < p$.

Even though AI increases the value of high-level expertise,[4] we show that it adversely

[3]This contrasts with the previous literature. For example, Acemoglu and Restrepo (2018) analyze how technology results in a substitution of capital for workers in specific tasks: the direct effect is to displace workers and reduce employment. In the model of Autor and Thompson (2025), occupations that involve more expert tasks benefit more from automation, consistently with the empirical evidence on AI investments across firms (Babina et al. (2024)) and the impact of AI on firm valuation (Eisfeldt, Schubert and Zhang (2023)).

[4]In the literature, this is referred to as AI "augmenting" rather than "displacing" skilled labor

impacts the accumulation of expertise among workers. Specifically, we show that the direct effect of an increase in automation is to worsen the distribution of expertise in a first-order stochastic dominance (FOSD) sense.

This can be decomposed into two effects. First, automation worsens the expertise distribution below the automation threshold. This is inconsequential from a social welfare perspective because these workers are assigned to tasks that do not require expertise. Second, automation worsens the expertise distribution above the threshold. This second effect leads to a shortage of workers with high levels of expertise, which reduces production. Intuitively, young workers in occupations where tasks are fully automated only perform generic tasks, so that they do not improve their expertise, and they cannot be assigned to tasks that require high-level expertise when they are old. By contrast, the few workers whose expertise when young is above the automation threshold have learning opportunities. Automation increases their pay because it makes them more productive, even when they are young, but also because it increases the scarcity of their high-level expertise when they are old.

One might presume that the endogenous change in the expertise distribution that stems from widespread automation would make it all the more valuable to further advance automation. Indeed, a worsening distribution of expertise across workers presumably increases the value of a technology that substitutes for human expertise. Surprisingly, our main result establishes that this conjecture is wrong.

This can be explained as follows. First of all, the distribution of worker expertise is determined by the level of automation last period, while the market wages for workers of different expertise are determined in a market equilibrium and are taken as given. In addition, the change in a firm's workforce composition following a small change in the level of automation has a second-order effect on the objective function since it was chosen optimally. As a result, the value of a marginal improvement in automation simply depends on the effect of this improvement on the productivity of workers across all occupations.

How does this depend on the level of automation last period? We show that this can be decomposed into two effects, both of which go in the same direction. First, a higher level of automation last period worsens the expertise distribution in a FOSD sense. In particular, it uniformly decreases the density of the expertise distribution for workers with expertise above the automation threshold. But these are precisely the workers whose productivity would be improved by further advances in automation. By contrast, further advances in automation do not improve the productivity of workers below the

(Agrawal, Gans, Goldfarb (2019), Tschang and Almirall (2021), Chen and Wang (2024)).

automation threshold, since all their expert tasks have already been automated. Second, since advances in automation result in more discrepancies across occupations, their value is even lower in the aggregate than when considering occupations in isolation because occupations are imperfect substitutes in the production function.

This result has implications for the legitimate concern that the widespread deployment of powerful AI might erode the expertise distribution to such an extent that it would require additional advances in AI. In other words, advances in AI would be self-reinforcing, which might be problematic in light of the potential disruptive effects associated with a transformative technology that is still largely misunderstood (Jones (2024), Acemoglu and Lensman (2024)). On the contrary, we show that while reaching a certain level of automation makes this level of automation all the more valuable in the future, it does not increase the value of *additional* advances in automation. However, this suggests that advances in automation must be sustainable.

In standard task models, workers specialize in tasks according to their comparative advantage (Autor, Levy, and Murnane (2003), Acemoglu and Autor (2011), Acemoglu and Restrepo (2018), Acemoglu et al. (2024)). These models involve atomistic task assignment: a given task (e.g. giving a presentation) is only performed by one skill group. By contrast, in models of occupational task bundling, a worker has multiple skills (Heckman and Scheinkman (1987)), and a given task will be performed in multiple occupations by different types of workers (Autor and Thompson (2025)).

Empirical studies show that almost all occupations have automatable tasks, but with large differences in the fraction of automatable tasks (Frey and Osborne (2017), Brynjolfsson, Mitchell, and Rock (2018)). For example, Xu et al. (2024) simulate a small software company, and find that 30% of tasks can be automated to AI agents.[5] The empirical study of Berger et al. (2024) finds that generative AI complements "high-level white collar jobs" but substitutes for low-level ones. This heterogeneous effect is consistent with our framework, in which low-level expert tasks get automated, whereas workers with expertise above the automation threshold get much more productive thanks to automation.

In a related paper, Ide (2025) focuses on the macroeconomic consequences of inadequate knowledge transmission due to automation by estimating the resulting decline in long-term economic growth. In his model, novices with identical skill levels learn by performing routine tasks, as opposed to expert tasks, and routine tasks get automated.

[5]There is also anecdotal evidence of practical adoption. For example, the Bank of New York Mellon reportedly "employs dozens of artificial intelligence-powered 'digital employees' that have company logins and work alongside its human staff." Source: "Digital Workers Have Arrived in Banking", *Wall Street Journal*, June 30 2025.

In our model, young workers with different skill levels learn by performing a range of expert tasks that get progressively automated. Ide and Talamàs (2025) incorporate AI agents in a model with workers who perform tasks and solvers who assist them. AI agents can potentially be co-workers or solvers/co-pilots, with different consequences for workers-solvers interactions.

Mo and Ouyang (2025) review the financial economics literature on AI, including the labor market implications of AI in section 6. Maug (2025) and Nishesh, Ouimet, and Simintzi (2024) review the corporate finance literature on human capital. We contribute to this literature by studying the feedback effect between automation and expertise.

1 The model

We build on the occupational task bundling model of Autor and Thompson (2025). We start by presenting a simplified version of their model, which we then extend by endogeneizing skill acquisition over a worker's life cycle. Since our focus is on professional occupations which do not require much capital, including coding, legal advice, financial advice, etc., we abstract from capital allocation and the combination of occupational outputs in the production of the "final consumption good". Instead, we are interested in the technology level which determines which occupations are automated. This allows to substantially simplify the model.

We rely on this model rather than on the canonical model of atomistic task assignment in which different workers specialize according to their comparative advantage. A model of task bundling captures important features of professional firms. For example, an expert lawyer will not just work on a narrow subset of tasks that are all at the frontier of her expertise level. Instead, she will typically be responsible for drafting a contract, or for a client relationship. Thus, even the most expert workers must also perform some tasks that only require a basic level of expertise, as well as more generic tasks such as meeting clients or presenting their findings.

Each worker must specialize in some "occupation". An occupation involves a bundle of tasks that each require a different level of expertise. Occupation $\phi \in [0, 1]$ involves a bundle of tasks t in the continuum $[-\theta(1-\phi), \phi]$, where $\theta > 0$.[6] A worker must perform all non-automatable tasks in her occupation. In occupation ϕ, tasks $t \in [-\theta(1-\phi), 0]$ are "generic tasks" that require labor input but no expertise, whereas tasks in $t \in (0, \phi]$

[6]Occupations are identified by the minimum expertise level required: an occupation where the most expert task requires an expertise level of ϕ is referred to as occupation ϕ.

are expert tasks. An expert task only requires labor input if it is not automated. Non-automated expert task t can only be completed by a worker with expertise $e_i \geq t$. This implies that a worker must have a level of expertise that allows her to perform the most expert task in her occupation.

Given a state of automation $I \in [0, 1]$, all expert tasks $t \in (0, I]$ can be automated. This means that more basic expert tasks are automated first, and those that require the highest level of expertise are automated last.

Several price-taking firms are competing for workers, whose measure is for now normalized to 1. In the baseline model, worker expertise e_i is uniformly distributed over $[0, 1]$ – we endogenize the expertise distribution in section 3.

As in Autor and Thompson (2025), a task can be completed using either labor or automation. If it is completed using labor, the production function is:

$$x_L(t, \ell_i(t)) = \begin{cases} \ell_i(t) & \text{if } t \in (-\theta, e_i] \\ 0 & \text{if } t \notin (-\theta, e_i] \end{cases} \tag{1}$$

If it is automated, the production function is:

$$x_K(t, \kappa_i(t)) = \begin{cases} \kappa_i(t) & \text{if } t \in (0, I] \\ 0 & \text{if } t \notin (0, I] \end{cases} \tag{2}$$

In our model, AI is used for the automation of some expert tasks. For simplicity, we assume a zero cost of AI.[7] In equilibrium, automated tasks are completed by AI. Production in task t by worker i and the AI allocated to this worker is:

$$x(t, \kappa_i(t), \ell_i(t)) = \begin{cases} x_L(t, \ell_i(t)) & \text{if } t \in (-\theta, 0] \cup (I, 1) \\ x_K(t, \kappa_i(t)) & \text{if } t \in (0, I] \end{cases} \tag{3}$$

In sum, there are three types of tasks:

1. Generic tasks $t \in (-\theta, 0]$, which require labor and can be complemented by any worker;

2. Automated tasks $t \in (0, I]$, which do not require labor;

3. Expert tasks $t \in (I, 1)$, which require labor and minimum expertise $e_i \geq t$.

The tasks involved in an occupation are perfect complements: all tasks need to be

[7] The crucial assumption is that tasks that can be automated will be automated.

completed for output to be produced.[8] This can be modelled with a Leontief production function, as in equation (5). Production $y_i(\phi)$ by worker i in occupation ϕ is:

$$y_i(\phi) = \min\{\hat{y}_i(\phi), \overline{y}\} \tag{4}$$

$$\text{where } \hat{y}_i(\phi) = \int_{\theta(1-\phi)}^{\phi} \min_t x(t, \kappa_i(t), \ell_i(t))dt \tag{5}$$

The parameter $\overline{y}$ is the maximum output possible in an occupation. This can be justified based on limited demand for this occupation. For automation to be beneficial, $\overline{y}$ must be sufficiently large. As a result of complementary tasks within an occupation, a worker will allocate her unit of labor uniformly across non-automated tasks:

$$\ell_i(t) = \begin{cases} \frac{1}{\phi+\theta(1-\phi)-\min\{I,\phi\}} & \text{if } t \in (-\theta(1-\phi), 0] \cup (I, 1) \\ 0 & \text{if } t \notin (-\theta(1-\phi), 0] \cup (I, 1) \end{cases} \tag{6}$$

In any occupation, the AI used will match the quantity of labor used for each task:

$$\kappa_i(t) = \begin{cases} \frac{1}{\phi+\theta(1-\phi)-\min\{I,\phi\}} & \text{if } t \in (0, I] \\ 0 & \text{if } t \notin (0, I] \end{cases} \tag{7}$$

We assume that AI is available to all firms and can be scaled as needed. This yields the same implication as in Ide and Talamàs (2025) that compute is abundant relative to human time, so that human time is the binding constraint.

There is a continuum $\phi \in (0, 1)$ of occupations in a "firm".[9] This means that each firm faces a spectrum of work to be accomplished, including challenging work that requires the highest level of expertise, and less challenging work that requires only limited expertise. For example, in a law firm, some contracts are very standard and involve boilerplate language (very low ϕ), whereas others are highly complex and involve rare legal issues (very high ϕ). Likewise, a financial intermediary can offer different services and serve different clienteles that require different levels of expertise. As in Autor and Thompson (2025), worker-level outputs $y_i(\phi)$ in each occupation aggregate linearly to total occupational output $y(\phi)$, and the final output produced by firm j is

[8] This is for simplicity. Autor and Thompson (2025), who also study the allocation of scarce capital, postulate a Cobb-Douglas structure instead. This yields the same result that a worker allocates her labor uniformly across non-automated tasks (i.e. generic tasks and expert tasks) in her occupation.

[9] This is different from the model of Autor and Thompson (2025) in which a firm is only involved in one occupation.

described by the following CES production function:

$$Y_j = \left(\int_0^1 y(\phi)^{\frac{\lambda-1}{\lambda}} d\phi \right)^{\frac{\lambda}{\lambda-1}} \tag{8}$$

where λ is the elasticity of substitution between occupation-level outputs.

Examples of occupations include product manager, CFO, store manager, engineer, financial advisor, and lawyer. All these occupations involve a worker having to perform many different tasks. Even though some tasks can be delegated, and of course automated, a single worker still needs to perform many different tasks. For example, a financial advisor needs to understand the circumstances of a client, to either generate expert advice or make sense of automated advice, and be available to meet with the client. Likewise, a product manager is responsible for coordinating and supervising all aspects of product development and marketing.

As in Autor and Thompson (2025), we assume $\lambda > 1$, so that various occupations are imperfect substitutes. Specifically, each occupation faces decreasing returns to scale. For example, a law firm which only offered very standardized legal services would not become a "one-stop-shop" for corporate and individual clients, and would therefore not sell as much standardized legal work as it would if it offered a more comprehensive range of services. This makes it more profitable to have one firm covering all occupations rather than two separate firms which together cover all occupations but which specialize in different subsets of occupations.

As in Autor and Thompson (2025) we assume $\theta > 1$. This implies that the productivity of workers whose expert tasks are partly automated is higher than the productivity of workers who only perform generic tasks. Generic tasks do not require field-specific expertise. Examples include interacting with customers and with computer programs, travelling, giving a presentation, and managing a team. Thus, generic tasks are not necessarily menial tasks. Many are arguably related to general managerial skills.

The cost of labor c_i has two components: the wage w_i, and other costs of employment h_i such as providing an office and support services (e.g. benefits, human resources, supervision) to the employee: $c_i = w_i + h_i$. The wage w_i is endogenous and determined by market conditions, whereas h_i is exogenously given.

Figures 1 and 2 depict the space of occupations and tasks. Occupations $\phi \in (0, 1)$ are on the x-axis, and tasks $t \in (-\theta, 1)$ are on the y-axis. The state of automation is the dashed line I. Each figure depicts two occupations, ϕ' and ϕ''. The former occupation involves tasks $t \in (t'_l, t'_h)$, and the latter occupation involves tasks $t \in (t''_l, t''_h)$. When automation is low, as in Figure 1, both occupations involve non-automated expert

tasks. When automation is high, as in Figure 2, only the latter occupation involves non-automated expert tasks.

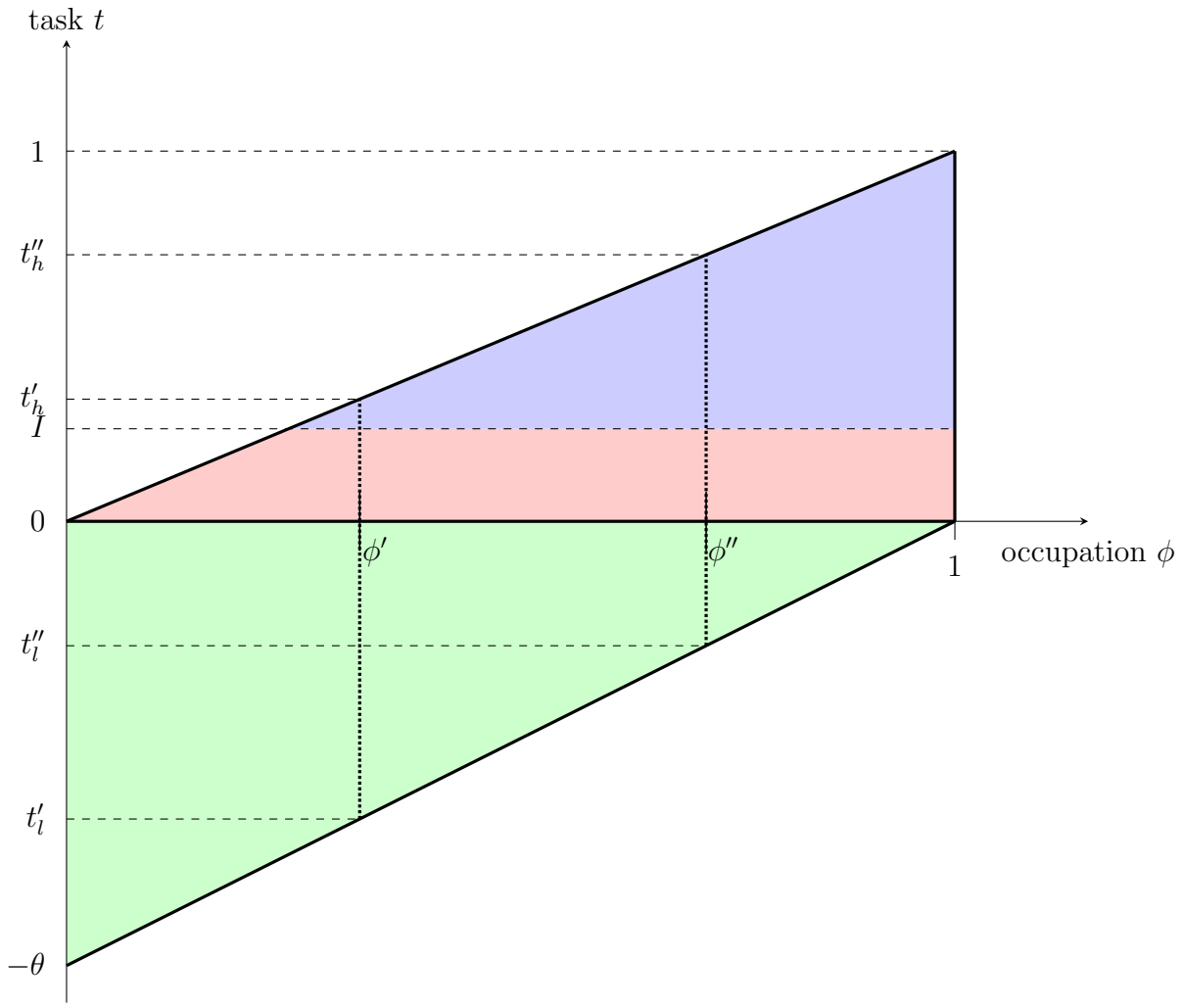


Figure 1: Expertise level required in each task t (y-axis), for all occupations ϕ (x-axis). In occupation ϕ', tasks range from t'_l to t'_h; in occupation ϕ'', tasks range from t''_l to t''_h. Green area: generic tasks. Red area: automated expert tasks. Blue area: expert tasks requiring human expertise.

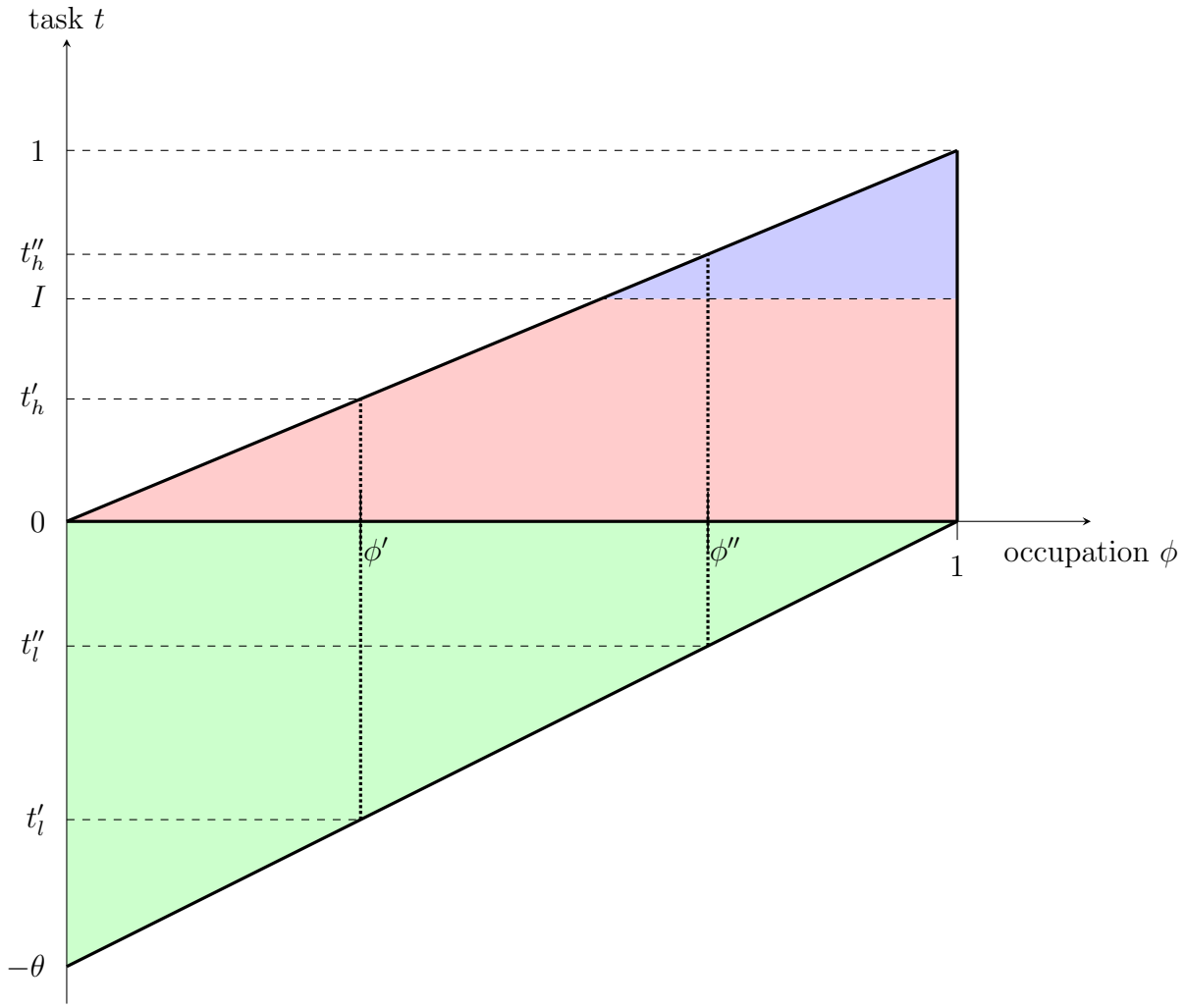


Figure 2: Expertise level required in each task t (y-axis), for all occupations ϕ (x-axis). In occupation ϕ', tasks range from t_l' to t_h'; in occupation ϕ'', tasks range from t_l'' to t_h''. Green area: generic tasks. Red area: automated expert tasks. Blue area: expert tasks requiring human expertise.

2 Analysis

Atomistic firms take the distribution of worker expertise as given in any period. Moreover, since they compete for workers with transferable expertise, they face a simple optimization problem in which they optimize by only taking into account the current expertise distribution.

Start with the case $I = 0$ in every period. In this case, in occupation ϕ with a unit

measure of workers per occupation:

$$y(\phi) = \int_{-\theta(1-\phi)}^{\phi} \frac{1}{\phi + \theta(1-\phi)} dt = 1 \qquad \forall \phi \in (0,1). \tag{9}$$

The equilibrium allocation of workers in the model of Autor and Thompson (2025), and in our model with $I = 0$,[10] involves a worker with expertise e_i working in occupation ϕ such that $e_i = \phi$. Workers are then uniformly distributed across occupations. In addition, equation (9) shows that each worker has the same productivity. Then:

$$Y = \left(\int_0^1 y(\phi)^{\frac{\lambda-1}{\lambda}} d\phi \right)^{\frac{\lambda}{\lambda-1}} = 1. \tag{10}$$

Now let $I > 0$. In this case, in occupation ϕ with a unit measure of workers per occupation:

$$\begin{aligned} y(\phi) &= \int_{-\theta(1-\phi)}^{\phi} \frac{1}{\phi + \theta(1-\phi) - \min\{\phi, I\}} dt \\ &= \begin{cases} \int_{-\theta(1-\phi)}^{\phi} \frac{1}{\theta(1-\phi)} dt & \text{if } \phi < I \\ \int_{-\theta(1-\phi)}^{\phi} \frac{1}{\phi+\theta(1-\phi)-I} dt & \text{if } \phi \geq I \end{cases} \\ &= \begin{cases} \frac{\phi+\theta(1-\phi)}{\theta(1-\phi)} & \text{if } \phi < I \\ \frac{\phi+\theta(1-\phi)}{\phi+\theta(1-\phi)-I} & \text{if } \phi \geq I \end{cases} \end{aligned} \tag{11}$$

Equation (11) implies that, with $\theta > 1$, the productivity $y(\phi)$ of a worker with $\phi > I$ is higher than the productivity of a worker with $\phi \leq I$.

An increase in I from $I = 0$ increases worker productivity in all occupations, because it allows each worker to increase her labor supply to all non-automated tasks in her occupation. With $\theta > 1$, workers with high expertise e_i benefit the most, because the production technology becomes even more AI-intensive in their occupations. As a result, a firm would ideally like to reassign some workers from low-ϕ occupations to high-ϕ occupations. However, each firm is constrained by the inability of workers with low expertise to work in occupations that require high expertise.

For occupations $\phi > I$, the equilibrium allocation of workers still involves a worker with expertise e_i working in occupation ϕ such that $e_i = \phi$. Intuitively, workers with $e_i < I$ cannot be reassigned to occupations that involve non-automated expert tasks (such that $\phi > I$) since they do not have the required expertise. Among workers with

[10] Autor and Thompson (2025) also allow for $I > 0$, but in their model it does not have implications for expertise acquisition, whereas it does in our model.

$e_i > I$, those with relatively low e_i are more productive than those with relatively high e_i if θ is sufficiently low, but the former cannot be reassigned to occupations with higher ϕ since they do not possess the required expertise.

For occupations $\phi < I$, workers such that $e_i < I$ are perfect substitutes, since occupations $\phi < I$ only require generic tasks (expert tasks in these occupations are fully automated). The firm will allocate these workers to equate their marginal products.

Lemma 1 *A worker with expertise $e_i > I$ is assigned to occupation $\phi = e_i$. Workers with expertise $e_i \leq I$ are assigned to occupations $\phi \leq I$, with a measure of workers $L(\phi)$ in occupation ϕ:*

$$L(\phi) = \frac{I\left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-1}}{\int_0^I \left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-1} d\phi} \tag{12}$$

In particular, the relative measures of workers in occupations ϕ' and ϕ'' both below the automation threshold I are:

$$\frac{L(\phi')}{L(\phi'')} = \left(\frac{\frac{\phi'+\theta(1-\phi')}{\theta(1-\phi')}}{\frac{\phi''+\theta(1-\phi'')}{\theta(1-\phi'')}}\right)^{\lambda-1} \tag{13}$$

Lemma 2 *When $\overline{y}$ is sufficiently large, for occupations $\phi < I$ whose expert tasks have been automated: $\frac{dL(\phi)}{d\phi} > 0$ for $I > 0$, and $\frac{dL(0)}{dI} < 0$.*

Lemma 2 shows how changes in automation affect the allocation of workers to occupations. First, among occupations whose expert tasks are automated, the measure of workers allocated to each occupation is an increasing function of ϕ. Intuitively, in this subset of occupations, occupations with a higher ϕ are the most productive ones. When occupations are (imperfect) substitutes in overall production, firms optimally reallocate workers to these occupations. Since in addition the measure of workers allocated to occupations above the automation threshold ($\phi > I$) does not change, this implies that occupations whose expert tasks have just been automated are the ones that employ the highest measure of workers. Workers may be even more productive in occupations whose expert tasks have not been automated (such that $\phi > I$); however, workers with expertise e_i less than I cannot perform some tasks in these occupations, and therefore they cannot be allocated to these occupations. The second result of Lemma 2 is a straightforward implication of the first: increases in automation reduce the measure of workers allocated to the most basic occupation (such that $\phi = 0$).

In sum, there are still as many workers in occupations with relatively low ϕ, which can be viewed as "entry-level jobs", but these workers with expertise below the automation threshold (i.e. $e_i \leq I$) are reassigned toward occupations that involve fewer generic tasks (i.e. with a higher ϕ), where they are more productive. However, as the level of automation I increases, jobs below the automation threshold are not just "entry-level jobs". For $I = 0.8$, Figure 3 shows that, compared to the case without automation ($I = 0$), the measure of workers allocated to occupations $\phi < 0.5029$ decreases, it increases for workers allocated to occupations $\phi \in (0.5029, 0.8)$, and it remains unchanged for workers allocated to occupations $\phi > 0.8$ above the automation threshold.

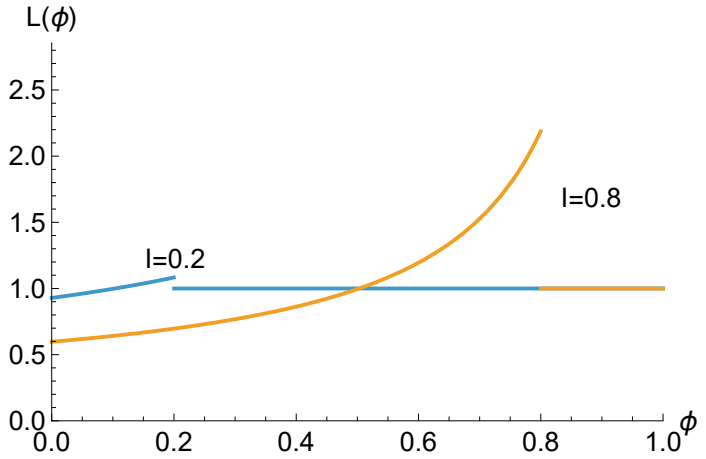


Figure 3: Measure $L(\phi)$ of workers assigned to each occupation ϕ as a function of ϕ, for two levels of automation I, when the capacity constraint is not binding (high $\overline{y}$), with $\lambda = 2$, $\theta = 1.5$.

Figure 3 depicts the function $L(\phi)$ for two different levels of automation I. An incremental rise in automation will fully automate all expert tasks in some occupation. When occupations are substitutes in the firm's production function, this initially *increases* employment in this occupation. Autor and Thompson (2025) get the same result. Intuitively, this occupation enjoys a productivity boost relative to other occupations where expert tasks have already been automated. Since occupations are substitutes, this productivity boost results in a greater resource allocation to this occupation. For the same reason, an increase in automation from a positive level reduces employment in occupations with low ϕ ("entry-level jobs"). This is consistent with recent empirical evidence (Brynjolfsson, Chandar, Chen (2025)). Automation also involves changes in workforce composition: workers with moderately low expertise (i.e. young educated workers) cannot perform any non-automated expert task. They can

then be displaced by workers with even lower expertise, since occupations with low ϕ only involve generic tasks.

In the baseline model presented in this section, the measure of workers $L(\phi)$ in occupations that still have non-automated expert tasks (i.e., for $\phi > I$) is a constant function of ϕ, as in Autor and Thompson (2025). This result will be overturned in section 3 below, when expertise is endogenized.

In sum, when $I > 0$:

$$\begin{aligned} Y &= \left(\int_0^I L(\phi) \left(\frac{\phi + \theta(1-\phi)}{\theta(1-\phi)} \right)^{\frac{\lambda-1}{\lambda}} d\phi + \int_I^1 \left(\frac{\phi + \theta(1-\phi)}{\phi + \theta(1-\phi) - I} \right)^{\frac{\lambda-1}{\lambda}} d\phi \right)^{\frac{\lambda}{\lambda-1}} \\ &= \left(\int_0^I \frac{I \left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)} \right)^{\lambda-\frac{1}{\lambda}}}{\int_0^I \left(\frac{z+\theta(1-z)}{\theta(1-z)} \right)^{\lambda-1} dz} d\phi + \int_I^1 \left(\frac{\phi + \theta(1-\phi)}{\phi + \theta(1-\phi) - I} \right)^{\frac{\lambda-1}{\lambda}} d\phi \right)^{\frac{\lambda}{\lambda-1}} \end{aligned} \tag{14}$$

When maximum production in an occupation, $\overline{y}$, is sufficiently small, even small productivity gains in an occupation due to automation will result in a reduction in this occupation's workforce. This effect is all the more pronounced that there are few generic tasks, i.e. θ is small. In this case, automation will largely shrink the workforce. For example, coding arguably involves few generic tasks and it is characterized by a small $\overline{y}$: once the required program is written, there is no longer a need for extensive coding expertise.

The salary of a worker corresponds to her marginal value in equilibrium, depending on the equilibrium allocation of workers to occupations.

Lemma 3 *The salary of a worker in occupation $\phi \leq I$ is:*

$$w(\phi) = \left(\int_0^I \left(\frac{\phi + \theta(1-\phi)}{\theta(1-\phi)} \right)^{\lambda-1} d\phi \right)^{\frac{1}{\lambda}} \left(\frac{Y}{I} \right)^{\frac{1}{\lambda}} \tag{15}$$

The salary of a worker in occupation $\phi > I$ such that $e_i = \phi$ is:

$$w(\phi) = \left(\frac{\phi + \theta(1-\phi)}{\phi + \theta(1-\phi) - I} \right)^{\frac{\lambda-1}{\lambda}} Y^{\frac{1}{\lambda}} \tag{16}$$

All workers in occupations with fully-automated expert tasks are interchangeable and therefore receive the same salary, i.e. $w(\phi)$ is constant for $\phi \leq I$. By contrast, workers in occupations where expert tasks are not fully automated need a minimum level of expertise. They are assigned to occupations according to their expertise level

and paid correspondingly. Figure 4 depicts market salaries for two different levels of automation.

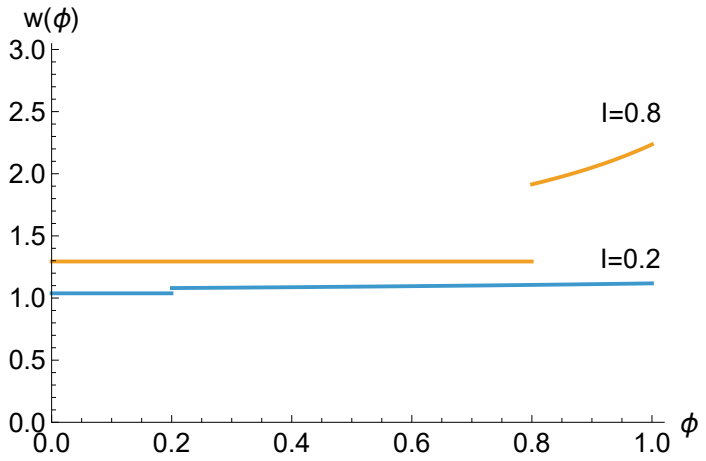


Figure 4: Salary $w(\phi)$ of a worker assigned to occupation ϕ as a function of ϕ, for two levels of automation I, when the capacity constraint is not binding (high $\overline{y}$), with $\lambda = 2$, $\theta = 1.5$.

3 Endogenous expertise acquisition

3.1 Baseline model

We now extend the model by replacing the unit measure of workers by overlapping generations (OLG) in which each cohort lives for two periods. An individual is young in the first period, and old in the second period. One unit of workers is born every period. Thus, there is a unit measure of young workers and a unit measure of old workers.

Assumption 1 describes the expertise distribution across generations and the learning process in occupations with non-automated expert tasks:

Assumption 1

(i) In the benchmark case $I = 0$, the expertise of all workers is uniformly distributed on $[0, 1]$, as in Autor and Thompson (2025);

(ii) The expertise e_i^o of old worker i with expertise e_i^y when young who worked in an occupation ϕ such that $e_i^y \geq \phi > I$:

(a) With probability $1 - p$, it is $e_i^o = e_i^y$;

(b) With probability p, it is uniformly distributed on $[e_i^y, 1]$.

Assumption 1 implies that the density of the expertise distribution of young workers is decreasing in e_i (see Appendix A). It also implies that, in the benchmark case $I = 0$, the density of the expertise distribution of old workers is increasing in e_i. The parameter p captures the intensity of on-the-job learning, which leads to differences between the expertise distribution of young workers and the expertise distribution of old workers. We also allow for partial learning in occupations in which expert tasks are fully automated (such that $\phi \leq I$), as described in Assumption 2 below.

Assumption 1 also pins down the density of young workers' expertise. As shown in Appendix A, the expertise e_i of young workers is continuously distributed on $[0, 1]$ with the following density:

$$f_y(e_i) \quad = \quad \alpha_y \left((-1)^{\frac{p}{p-2}} (x-1)^{-\frac{p}{p-2}} \right) \tag{17}$$

where the scaling constant α_y is set so that the measure of young workers is equal to 1. With $I = 0$, Assumption 1 implies that the expertise of old workers is continuously distributed on $[0, 1]$ with density:

$$f_o(e_i) \quad = \quad p \int_0^{e_i} f_y(z) \frac{1}{1-z} dz + (1-p) f_y(e_i) \tag{18}$$

As opposed to Assumption 1, which takes the level of expertise when young as a given, equation (18) takes the level of expertise when old as a given, and integrates over levels of expertise when young. The first term on the RHS captures the measure of workers with expertise less than e_i when young who learned and ended up with expertise e_i when old; the second term on the RHS captures the measure of workers with expertise equal to e_i when young who did not learn and therefore kept the same level of expertise when old. This captures a learning process in the form of a compound lottery, as described in Assumption 1. With probability $1 - p$, a young worker does not gain expertise on the job, so that her expertise when old is the same as her expertise when young. With probability $p \in (0, 1)$, she does. In this case, her expertise when old is uniformly distributed between her expertise when young and 1: $e_i^o \in [e_i^y, 1]$ with a uniform density equal to $\frac{1}{1-e_i}$. This captures learning by doing by young workers who use their expertise in their occupation when expert tasks have not been automated.

By construction, the measure of old workers is equal to 1:

$$\int_0^1 f_o(e_i) de_i \quad = \quad 1 \tag{19}$$

Moreover, the density functions used are such that we have: $f_y(e_i) + f_o(e_i) = 2$ for

every $e_i \in [0,1]$ when $I = 0$. See Appendix A for more details.

Assumption 2 describes the learning process in occupations where all expert tasks are automated.

Assumption 2

A worker i in occupation ϕ such that $\phi \leq I$ does not perform any expert tasks. This worker learns as in Assumption 1 part (ii) but with probability $q \in [0, p)$.

Assumption 2 captures "learning by proximity" to automated expert tasks and to expert workers. First, generic tasks that are related to automated expert tasks may provide some learning opportunities. Second, proximity with other workers who perform expert tasks may generate *knowledge spillovers* within the organization (horizontal learning). In the former case, there is more learning if the occupation involves higher-level expert tasks. In the latter case, there is more learning if the occupation is closer to occupations where expert tasks are not fully automated. In any case, this can be formalized with the same learning process already used, except that the intensity of learning is now parameterized by q instead of p, where $q < p$ to account for reduced learning opportunities in occupations where expert tasks are fully automated.

Parameter q will be higher when knowledge spillovers are facilitated, as with internships and flat hierarchies, but lower when they are impeded, for example because of remote work and other working arrangements that prevent frequent interactions between junior and senior employees.

In sum, with $I > 0$, the expertise of old workers is continuously distributed on $[0,1]$ with density:

$$f_o(e_i) = \begin{cases} q\int_0^{e_i} f_y(z)\frac{1}{1-z}dz + (1-q)f_y(e_i) & \text{if } \phi_i \leq I \\ q\int_0^{I} f_y(z)\frac{1}{1-z}dz + p\int_I^{e_i} f_y(z)\frac{1}{1-z}dz + (1-p)f_y(e_i) & \text{if } \phi_i > I \end{cases} \quad (20)$$

On the second line, the first integral is the measure of workers with expertise when young less than I whose expertise when old is equal to e_i; the second integral is the measure of workers with expertise when young higher than I whose expertise when old is equal to e_i.

In sum, with endogenous expertise acquisition, the allocation of workers to occupations affects expertise production and the availability of expertise next period.

Figure 5 depicts the expertise density of young workers, old workers, and across cohorts. The first row of Figure 5 depicts the case without automated tasks. In this case, worker expertise is uniformly distributed on $[0,1]$, as in Autor and Thompson (2025). The second row depicts the case when half of occupations have fully automated

expert tasks ($I = 0.5$), and when workers who only complete generic tasks still improve their expertise to some extent. In this case, young workers in occupations $\phi \in [0, \frac{1}{2}]$ only partially learn a fraction $\frac{q}{p} = \frac{2}{3}$ of what they would have learned without automation. As a result, the expertise density of old workers is flatter on $e_i \in [0, \frac{1}{2}]$ than in the case without any automation. The third row depicts the case when half of occupations have fully automated expert tasks ($I = 0.5$), as in the previous row, but also when workers who only complete generic tasks do not improve their expertise. In this case, young workers with expertise below the automation threshold (i.e. $e_i < I = 0.5$) do not learn and therefore have the same expertise when old, whereas young workers with expertise above the automation threshold still learn. The fourth row depicts the case when all occupations have fully automated expert tasks ($I = 1$), and when workers who only complete generic tasks still improve their expertise to some extent. With $I = 1$, there is no longer the threshold effect that characterized the second and third rows where $I = 0.5$. The fifth row of Figure 5 depicts the case without on-the-job learning. The expertise distribution of old workers is then identical to the expertise distribution of young workers.

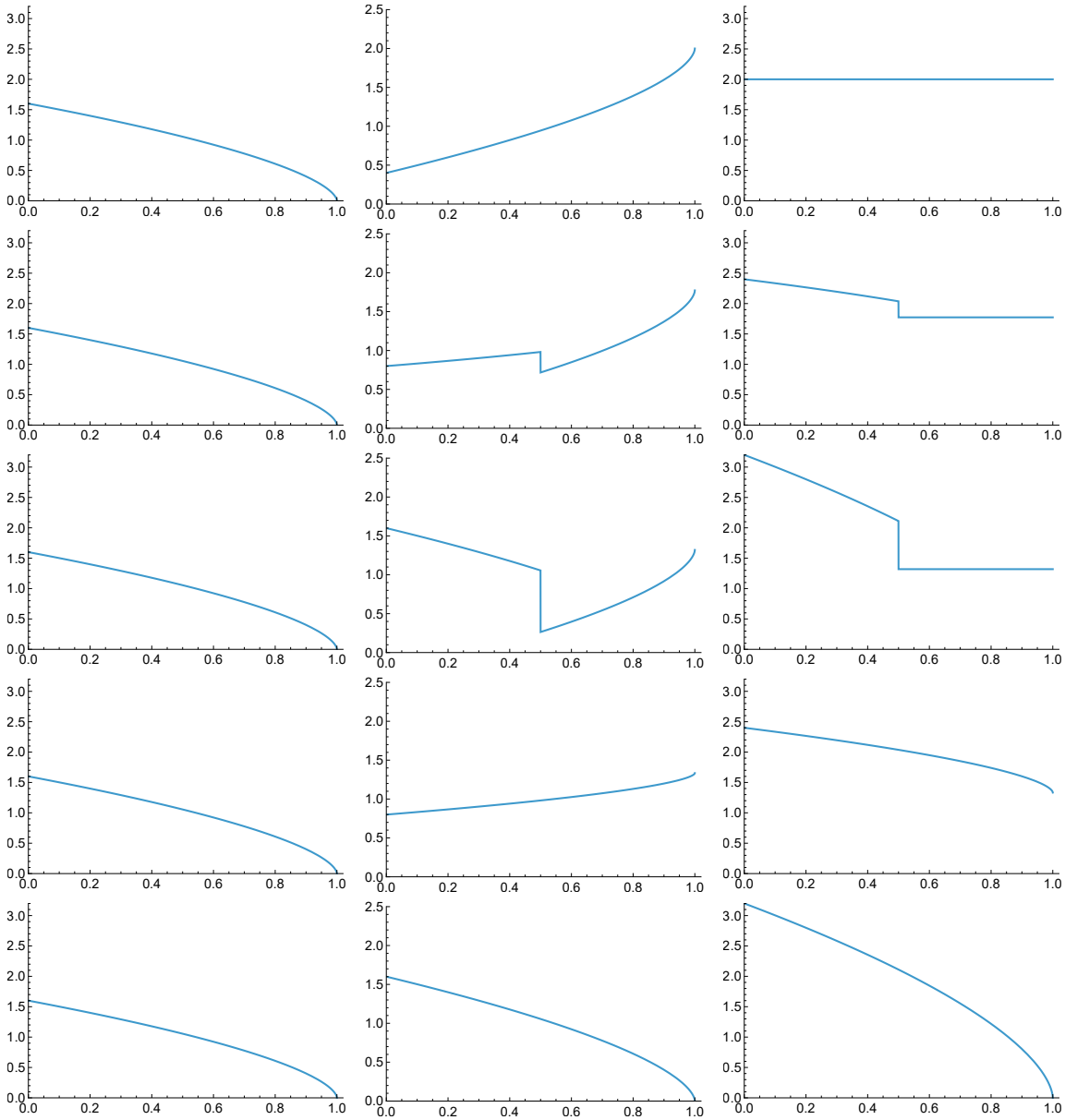


Figure 5: Left panel: density $f_y(e_i)$ of young workers with expertise e_i as a function of e_i. Middle panel: density $f_o(e_i)$ of old workers with expertise e_i as a function of e_i. Right panel: $f_y(e_i) + f_o(e_i)$ as a function of e_i. First row: $I_0 = 0$. Second row: $I_0 = 0.5$, $q = 0.5$. Third row: $I_0 = 0.5$, $q = 0$. Fourth row: $I_0 = 1$, $q = 0.5$. Fifth row: $I_0 = 1$, $q = 0$. In all rows and all panels, $p = \frac{3}{4}$.

The expertise distribution of young workers is fixed. Proposition 1 show that increases in automation worsen the overall expertise distribution in a FOSD sense.

Proposition 1 *A higher I worsens the distribution of old workers' expertise in a FOSD sense.*

For workers with expertise such that they do not perform any expert task, a marginal increase in I does not change anything: their learning process is still described by Assumption 2. However, a marginal increase in I increases the measure of workers who do not perform any expert task, and who consequently learn less than they otherwise would. This worsens the expertise distribution in a FOSD sense for levels of expertise higher than the initial level I below which expert tasks are automated.

We now study how the firm's willingness to pay for a marginal increase in automation I in the current period depends on the state of automation I_0 in the preceding period. The latter matters because it affects the expertise distribution, which firms take as given in the current period.

The relation is not obvious for several reasons. First, firms compete for workers, so that workers are paid their marginal product, and firm profits only account for a fraction of the total surplus. Second, an increase in automation in a period worsens the expertise distribution next period regardless of automation choices in this period. Third, automation increases the productivity of all workers, not just those above the automation threshold: for workers below the automation threshold, productivity increases because they can devote more time to generic tasks while the expert tasks involved in their occupation are automated; for workers above the automation threshold, productivity increases because they can devote more time to high-level expert tasks as well as generic tasks. The latter effect is often referred to as the augmentation effect of AI.

Proposition 2 shows that the previously reached level of automation has an unambiguous effect on further advances in automation.

Proposition 2 *The value of marginal advances in automation this period is decreasing in the level of automation reached last period. The magnitude of the effect is decreasing in* λ.

The higher the pre-existing level of automation, the lower the value of improving automation further. Intuitively, a marginal increase in automation does not affect the productivity of workers whose expertise is below the automation threshold, since all their expert tasks have already been automated. On the contrary, it increases the productivity of all workers whose expertise is above the automation threshold. The measure of these workers is decreasing in the level of automation in the preceding period (see Figures 5 for examples). This is the direct effect of the previous level of automation on the value of further advances in automation.

The indirect effect, which goes in the same direction, is related to the imperfect substitutability of occupations in the production process, and can be explained as follows.

Production involves imperfect substitutability across occupations ($\lambda > 1$). Occupations are almost perfect substitutes when λ is very high, whereas they are very imperfect substitutes when λ is low. In this latter case, it is especially valuable to reach similar levels of production across all occupations. But advances in automation mostly increase productivity in occupations above the automation threshold, where it is already higher to start. Thus, further advancing automation from a high level is less valuable, especially when λ is low.

4 Discussion

Autor and Thompson (2025) interpret Y as the output of an economy. Accordingly, different occupations are assumed to be imperfect substitutes in their model. In the model presented in this paper, where Y is interpreted as the output of a firm, it might be worth also considering the case when different occupations are complements, i.e. allow for $\lambda < 1$. This is especially relevant in firms that produce goods rather than professional services.

Indeed, the output of some firms relies on each occupation being performed sufficiently well. For example, a rocket ship will not work if some component is poorly designed or manufactured. Likewise, a pharmaceutical company will not successfully develop a new drug if it does not successfully complete all steps of drug development (discovery and development, preclinical research, clinical research, FDA review), where each of these steps involves several phases.

When occupations are complements in the production function, the direct effect of automation on the expertise distribution described in Proposition 1 can be amplified by a worker reallocation effect. Indeed, when workers in occupations that require high expertise become much more productive due to automation, there is a need for additional workers to perform generic tasks in occupations where expert tasks have been automated. Firms may then reallocate workers to these occupations that offer fewer opportunities for learning.[11]

[11]Of course, firms could in principle reallocate old expert workers to these tasks to still ensure that young workers learn when performing expert tasks. However, they will only be incentivized to do so in the presence of labor market imperfections that leads them to capture some of the gains from enhanced worker learning. In addition, firm-level institutions, including implicit and explicit contracts as well as intermediaries such as unions, may prevent worker reassignments that would be viewed as demotions.

5 Conclusion

In this paper, we study the potentially self-reinforcing nature of advances in automation because of their effect on expertise accumulation among workers. We extend the occupational task bundling model of Autor and Thompson (2025) to an OLG framework by incorporating a simple learning process. This allows us to derive the distribution of worker expertise as a function of the level of automation. We show that automation leads to a redistribution of worker expertise which is a combination of two effects. The first is a redistribution of expertise below the automation threshold. It is inconsequential because no expertise is needed in these occupations thanks to automation. The second effect, which is consequential, is a decrease of expertise above the automation threshold. Indeed, young workers with a low level of expertise do not complete any expert task and therefore cannot improve their expertise. This highlights the special nature of investments in AI, which do not just change the production process but also have longer-term effects on expertise generation.

We use this framework to determine whether the value of further advances in automation is increasing in the level of automation. We show that it is not. On the one hand, reaching a certain level of automation increases the value of *sustaining* this level of automation in the future. On the other hand, the value of further *advancing* automation is negatively related to the level of automation. Intuitively, further advances in automation improve the productivity of workers with a high level of expertise. But the fraction of these workers in the population is decreasing in the level of automation already reached.

References

Acemoglu, D., Autor, D., 2011. Skills, Tasks and Technologies: Implications for Employment and Earnings. In: *Handbook of Labor Economics*, Elsevier, 4, 1043-1171.

Acemoglu, D., Kong, F. and Restrepo, P., 2024. Tasks at work: comparative advantage, technology and labor demand. NBER working paper 32872.

Acemoglu, D., Kong, F. and Restrepo, P., 2024. Tasks at work: comparative advantage, technology and labor demand. NBER Working paper 32872.

Acemoglu, D. and Lensman, T., 2024. Regulating transformative technologies. *American Economic Review: Insights*, 6(3), 359-376.

Acemoglu, D., Restrepo, P. 2018. Artificial intelligence, automation, and work. In: *The economics of artificial intelligence: An agenda*. University of Chicago Press, 197-236.

Agrawal, A., Gans, J.S. and Goldfarb, A., 2019. Artificial intelligence: the ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives*, 33(2), 31-50.

Autor, D.H. and Handel, M.J., 2013. Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labor Economics*, 31, S59-S96.

Autor, D. and Thompson, N., 2025. Expertise. *Journal of the European Economic Association*, forthcoming.

Autor, D.H., Levy, F. and Murnane, R.J., 2003. The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118, 1279-1333.

Babina, T., Fedyk, A., He, A. and Hodson, J., 2024. Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151, 103745.

Baccara, M., Lee, S. and Yariv, L., 2023. Task allocation and on-the-job training. *Journal of Economic Theory*, 207, 105587.

Berger, P.G., Cai, W., Qiu, L. and Shen, C.X., 2024. Employer and employee responses to generative AI: Early evidence. Working paper, University of Chicago.

Brynjolfsson, E., Mitchell, T., Rock, D., 2018. What can machines learn and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, 108, 43-47.

Brynjolfsson, E., Chandar, B., Chen, B., 2025. Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence. Working paper, Stanford University.

Chen, M.A. and Wang, J., 2024. Displacement or augmentation? The effects of AI innovation on workforce dynamics and firm value. Working paper, Georgia State

University.

Eisfeldt, A.L., Schubert, G. and Zhang, M.B., 2023. Generative AI and firm values. NBER Working paper 31222.

Frey, C.B. and Osborne, M.A., 2017. The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254-280.

Hampole, M., Papanikolaou, D., Schmidt, L.D. and Seegmiller, B., 2025. Artificial intelligence and the labor market. NBER Working paper 33509.

Heckman, J. and Scheinkman, J., 1987. The importance of bundling in a Gorman-Lancaster model of earnings. *Review of Economic Studies*, 54, 243-255.

Ide, E. and Talamàs, E., 2025. Artificial intelligence in the knowledge economy. *Journal of Political Economy*, forthcoming.

Ide, E., 2025. Automation, AI, and the Intergenerational Transmission of Knowledge. Working paper, IESE.

Jones, C.I., 2024. The AI dilemma: Growth versus existential risk. *American Economic Review: Insights*, 6(4), 575-590.

Lowenstein, M. and Spletzer, J., 1994. Informal training: A review of existing data and some new evidence. Working paper, Bureau of Labor Statistics.

Lichtinger, G. and Hosseini, S.M., 2025. Generative AI as Seniority-Biased Technological Change: Evidence from U.S. Résumé and Job Posting Data. Working paper, Harvard University.

Maug, E., 2025. Financing Human Capital: A Survey and Synthesis. Working paper, University of Mannheim.

McAdams, J.M., 2019. Non-compete agreements: a review of the literature. Available at SSRN 3513639.

Mo, H., Ouyang, S., 2025. (Generative) AI in Financial Economics. Working paper, Oxford University.

Nishesh, N., Ouimet, P. and Simintzi, E., 2024. Labor and corporate finance. *Handbook of Corporate Finance*, 647-673.

Tschang, F.T. and Almirall, E., 2021. Artificial intelligence as augmenting automation: Implications for employment. *Academy of Management Perspectives*, 35(4), 642-659.

Xu, F.F., Song, Y., Li, B., Tang, Y., Jain, K., Bao, M., Wang, Z.Z., Zhou, X., Guo, Z., Cao, M. and Yang, M., 2024. TheAgentCompany: benchmarking LLM agents on consequential real world tasks. arXiv preprint arXiv:2412.14161.

Appendix

A Expertise distribution of young and old workers

If $I = 0$, the density of the expertise distribution of all workers at $e_i \in [0, 1]$ is such that, up to a scaling constant:

$$f_y(e_i) + f_o(e_i) \quad = \quad f_y(e_i) + p \int_0^{e_i} f_y(z) \frac{1}{1-z} dz + (1-p) f_y(e_i) \tag{21}$$

Having $f_y(e_i) + f_o(e_i)$ be constant with respect to e_i requires:

$$f_y'(e_i) + p f_y(e_i) \frac{1}{1-e_i} + (1-p) f_y'(e_i) \quad = \quad 0 \tag{22}$$

Solving this differential equation with the arbitrary boundary condition $f(0) = 1$ yields:

$$f_y(e_i) \quad = \quad (-1)^{\frac{p}{p-2}} (e_i - 1)^{-\frac{p}{p-2}} \tag{23}$$

so that:

$$f_y'(e_i) \quad = \quad \frac{p}{2-p} (-1)^{\frac{p}{p-2}} (e_i - 1)^{\frac{2-2p}{p-2}} \tag{24}$$

which is strictly negative for any $p \in (0, 1)$ and $e_i \in (0, 1)$.

The scaling constant α_y solves:

$$\alpha_y \int_0^1 \left((-1)^{\frac{p}{p-2}} (x-1)^{-\frac{p}{p-2}} \right) de_i \quad = \quad 1.$$

It can be verified that, for any parameter values:

$$\int_0^1 f_o(e_i) de_i \quad = \quad 1 \tag{25}$$

where $f_o(e_i)$ is defined in equation (18).

B Proofs

Proof of Lemma 1:

Consider occupations $\phi < I$, with workers such that $e_i < I$ who are assigned to occupations to equate their marginal products. When marginal products across occupations $\phi < I$ are the same, there exists a constant k such that:

$$\begin{aligned} \frac{\lambda-1}{\lambda}\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\left(L(\phi)\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{-\frac{1}{\lambda}} &= k \\ \Leftrightarrow \quad (L(\phi))^{\frac{1}{\lambda}} &= \left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\frac{\lambda-1}{\lambda}} \Big/ k \\ \Leftrightarrow \quad L(\phi) &= \left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-1} \Big/ k^{\lambda} \end{aligned} \tag{26}$$

When there is a measure I of workers with expertise e_i less than I, we must have:

$$\int_0^I L(\phi)d\phi = I \tag{27}$$

Plugging $L(\phi)$ from equation (26) and rearranging yields:

$$k = \left(\int_0^I \left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-1} d\phi \Big/ I\right)^{\frac{1}{\lambda}} \tag{28}$$

Plugging this value of k in equation (26):

$$L(\phi) = \frac{I\left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-1}}{\int_0^I \left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-1} d\phi} \tag{29}$$

■

Proof of Lemma 3:

For $\phi \leq I$, a marginal increase in $L(\phi)$ increases output by:

$$\begin{aligned}
\frac{dY}{dL(\phi)} &= \frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\left(L(\phi)\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{-\frac{1}{\lambda}} \\
&\quad\times\left(\left(\int_0^I \frac{I\left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-\frac{1}{\lambda}}}{\int_0^I\left(\frac{z+\theta(1-z)}{\theta(1-z)}\right)^{\lambda-1}dz}d\phi+\int_I^1\left(\frac{\phi+\theta(1-\phi)}{\phi+\theta(1-\phi)-I}\right)^{\frac{\lambda-1}{\lambda}}d\phi\right)\right)^{\frac{\lambda}{\lambda-1}-1} \\
&= \frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\left(\frac{I\left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-1}}{\int_0^I\left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-1}d\phi}\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{-\frac{1}{\lambda}}Y^{\frac{1}{\lambda}} \\
&= I^{-\frac{1}{\lambda}}\left(\int_0^I\left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-1}d\phi\right)^{\frac{1}{\lambda}}Y^{\frac{1}{\lambda}} \qquad (30)
\end{aligned}$$

Since firms compete for workers, this expression gives the equilibrium salary of a worker with expertise $e_i=\phi$ such that $\phi\leq I$.

For $\phi>I$, a marginal increase in $L(\phi)$ increases output by:

$$\begin{aligned}
\frac{dY}{dL(\phi)} &= \frac{\phi+\theta(1-\phi)}{\phi+\theta(1-\phi)-I}\left(L(\phi)\frac{\phi+\theta(1-\phi)}{\phi+\theta(1-\phi)-I}\right)^{-\frac{1}{\lambda}} \\
&\quad\times\left(\left(\int_0^I \frac{I\left(\frac{\phi+\theta(1-\phi)}{\theta(1-\phi)}\right)^{\lambda-\frac{1}{\lambda}}}{\int_0^I\left(\frac{z+\theta(1-z)}{\theta(1-z)}\right)^{\lambda-1}dz}d\phi+\int_I^1\left(\frac{\phi+\theta(1-\phi)}{\phi+\theta(1-\phi)-I}\right)^{\frac{\lambda-1}{\lambda}}d\phi\right)\right)^{\frac{\lambda}{\lambda-1}-1} \\
&= L(\phi)^{-\frac{1}{\lambda}}\left(\frac{\phi+\theta(1-\phi)}{\phi+\theta(1-\phi)-I}\right)^{\frac{\lambda-1}{\lambda}}Y^{\frac{1}{\lambda}} \\
&= \left(\frac{\phi+\theta(1-\phi)}{\phi+\theta(1-\phi)-I}\right)^{\frac{\lambda-1}{\lambda}}Y^{\frac{1}{\lambda}} \qquad (31)
\end{aligned}$$

Indeed, for $\phi>I$, $L(\phi)=1$ in the equilibrium of the baseline model. ■

Proof of Proposition 1:

We must show that, for $I'>I$, we have:

$$F_o^I(e_i)\leq F_o^{I'}(e_i) \qquad (32)$$

where F_o^I is the CDF of old worker's expertise with automation level I. It is defined as:

$$F_o^I(e_i) \equiv \int_0^{e_i} f_o^I(z)dz \qquad (33)$$

where:

$$f_o(e_i) = \begin{cases} q\int_0^{e_i} f_y(z)\frac{1}{1-z}dz + (1-q)f_y(e_i) & \text{if } e_i \leq I \\ q\int_0^{I} f_y(z)\frac{1}{1-z}dz + p\int_I^{e_i} f_y(z)\frac{1}{1-z}dz + (1-p)f_y(e_i) & \text{if } e_i > I \end{cases} \quad (34)$$

Thus, we have:

$$\begin{aligned} F_o^I(e_i) &= \int_0^{\min\{e_i,I\}} \left(q\int_0^x f_y(z)\frac{1}{1-z}dz + (1-q)f_y(x) \right) dx \\ &\quad + \int_{\min\{e_i,I\}}^{e_i} \left(q\int_0^I f_y(z)\frac{1}{1-z}dz + p\int_I^x f_y(z)\frac{1}{1-z}dz + (1-p)f_y(x) \right) dx \end{aligned} \quad (35)$$

First, for $e_i \leq I$:

$$\frac{dF_o^I(e_i)}{dI} = 0 \quad (36)$$

Second, for $e_i > I$:

$$\begin{aligned} \frac{dF_o^I(e_i)}{dI} &= ((1-q)-(1-p))\, f_y(I) + \int_I^{e_i} (q-p) f_y(I)\frac{1}{1-I}dx \\ &= (p-q)\, f_y(I)\left(1 - \frac{e_i - I}{1-I}\right) \geq 0 \end{aligned} \quad (37)$$

with a strict inequality for $e_i < 1$, which completes the proof. ■

Proof of Proposition 2:

For a given level of automation I and equilibrium measure of workers $L(\phi)$, use Lemma 3 to rewrite firm profits as:

$$\begin{aligned} \Pi &= \left(\int_0^1 L(\phi)\,(y(\phi))^{\frac{\lambda-1}{\lambda}}\, d\phi \right)^{\frac{\lambda}{\lambda-1}} - \int_0^1 L(\phi) w(\phi) d\phi \\ &= \left(\int_0^I L(\phi) \left(\frac{\phi + \theta(1-\phi)}{\theta(1-\phi)} \right)^{\frac{\lambda-1}{\lambda}} d\phi + \int_I^1 L(\phi) \left(\frac{\phi + \theta(1-\phi)}{\phi + \theta(1-\phi) - I} \right)^{\frac{\lambda-1}{\lambda}} d\phi \right)^{\frac{\lambda}{\lambda-1}} \\ &\quad - \int_0^I L(\phi) \left(\int_0^I \left(\frac{z + \theta(1-z)}{\theta(1-z)} \right)^{\lambda-1} dz \right)^{\frac{1}{\lambda}} \left(\frac{Y(I)}{I} \right)^{\frac{1}{\lambda}} d\phi \\ &\quad - \int_I^1 L(\phi) \left(\frac{\phi + \theta(1-\phi)}{\phi + \theta(1-\phi) - I} \right)^{\frac{\lambda-1}{\lambda}} \left(\frac{Y(I)}{L(\phi)} \right)^{\frac{1}{\lambda}} d\phi \end{aligned} \quad (38)$$

For $I > 0$ at time $t-1$, we have:

$$L(\phi) \begin{cases} \geq 2 & \text{for } \phi \leq I \\ < 2 & \text{for } \phi > I \end{cases} \tag{39}$$

At $\phi = I$, we have:

$$\frac{\phi + \theta(1-\phi)}{\phi + \theta(1-\phi) - I} = \frac{\phi + \theta(1-\phi)}{\theta(1-\phi)} \tag{40}$$

Moreover, with $\theta > 1$, we have:

$$\begin{aligned} \frac{d}{d\phi}\frac{\phi + \theta(1-\phi)}{\theta(1-\phi)} &= \frac{(1-\theta)\theta(1-\phi) + \theta(\phi + \theta(1-\phi))}{(\theta(1-\phi))^2} = \frac{1}{(\theta(1-\phi))^2} > 0 \\ \frac{d}{d\phi}\frac{\phi + \theta(1-\phi)}{\phi + \theta(1-\phi) - I} &= \frac{(1-\theta)(\phi + \theta(1-\phi) - I) - (1-\theta)(\phi + \theta(1-\phi))}{(\phi + \theta(1-\phi) - I)^2} \\ &= -(1-\theta)\frac{I}{(\phi + \theta(1-\phi) - I)^2} > 0 \end{aligned}$$

since $\theta > 1$. As a result, for any $\phi' \leq I < \phi''$, we have:

$$\frac{\phi' + \theta(1-\phi')}{\theta(1-\phi')} < \frac{\phi'' + \theta(1-\phi'')}{\phi'' + \theta(1-\phi'') - I} \tag{41}$$

We consider a marginal change in I from its current level. A firm takes the workers' wages as given and the current level of I as an exogenous parameter, and it maximizes its profits by choosing the allocation of workers to occupations. In equilibrium, this allocation is $L(\phi)$ for all firms. By the envelope theorem, the effect of a marginal change in I on the firm's objective function is given by the partial effect of a marginal change in I holding $L(\phi)$ constant. From equation (38), the last two lines describe the wage costs which remain unchanged.

We now simplify notations. Define:

$$\begin{aligned} P(\phi) &\equiv \begin{cases} \left(\frac{\phi + \theta(1-\phi)}{\theta(1-\phi)}\right)^{\frac{\lambda-1}{\lambda}} & \text{for } \phi \leq I \\ \left(\frac{\phi + \theta(1-\phi)}{\phi + \theta(1-\phi) - I}\right)^{\frac{\lambda-1}{\lambda}} & \text{for } \phi > I \end{cases} \\ f(x) &\equiv x^{\frac{\lambda}{\lambda-1}} \end{aligned} \tag{42}$$

so that $P(\phi) > 0$, $P'(\phi) > 0$, $f'(x) > 0$, $f''(x) > 0$, and:

$$\frac{\partial P(\phi)}{\partial I} \equiv \begin{cases} 0 & \text{for } \phi \leq I \\ \frac{1}{\phi+\theta(1-\phi)-I}\left(\frac{\phi+\theta(1-\phi)}{\phi+\theta(1-\phi)-I}\right)^{\frac{\lambda-1}{\lambda}} & \text{for } \phi > I \end{cases}$$

Thus, taking the expectation over a uniform distribution on $[0,1]$, we can rewrite profits in equation (38) as:

$$\Pi = f\left(\mathbb{E}[L(\phi)P(\phi)]\right) - \mathbb{E}[L(\phi)w(\phi)] \tag{43}$$

so that, bearing in mind that the function $L(\phi)$ depends on I_0 (not I), we have:

$$\frac{\partial \Pi}{\partial I} = \mathbb{E}\left[L(\phi)\frac{\partial P(\phi)}{\partial I}\right] f'\left(\mathbb{E}[L(\phi)P(\phi)]\right) \tag{44}$$

The function $L(\phi)$ is discontinuous at $\phi = I_0$. Denote $\lim_{\phi \nearrow I_0} L(\phi) = L(\phi^-)$ and $\lim_{\phi \searrow I_0} L(\phi) = L(\phi^+)$. The RHS of equation (44) includes two terms. The first term captures improvements in productivity from automation. The second term captures decreasing returns to scale from automation. Both terms depend on $L(\phi)$ which is determined by the level of automation last period, I_0. By contrast, the level of automation this period, I, affects $P(\phi)$. Marginal improvements in automation I are less valuable when the past level of automation I_0 is higher if:

$$\begin{aligned} \frac{\partial^2 \Pi}{\partial I \partial I_0} &= \mathbb{E}\left[\frac{\partial L(\phi)}{\partial I_0}\frac{\partial P(\phi)}{\partial I}\right] f'\left(\mathbb{E}[L(\phi)P(\phi)]\right) \\ &\quad + \mathbb{E}\left[L(\phi)\frac{\partial P(\phi)}{\partial I}\right]\mathbb{E}\left[\frac{\partial L(\phi)}{\partial I_0}P(\phi)\right] f''\left(\mathbb{E}[L(\phi)P(\phi)]\right) < 0 \\ \Leftrightarrow \quad & \underbrace{\mathbb{E}\left[L(\phi)\frac{\partial P(\phi)}{\partial I}\right]}_{>0}\underbrace{\mathbb{E}\left[\frac{\partial L(\phi)}{\partial I_0}P(\phi)\right]}_{<0}\underbrace{\frac{f''\left(\mathbb{E}[L(\phi)P(\phi)]\right)}{f'\left(\mathbb{E}[L(\phi)P(\phi)]\right)}}_{>0} + \underbrace{\mathbb{E}\left[\frac{\partial L(\phi)}{\partial I_0}\frac{\partial P(\phi)}{\partial I}\right]}_{<0} < 0 \end{aligned} \tag{45}$$

We get the inequalities below the various terms in equation (45) because:

$$\mathbb{E}\left[\frac{\partial L(\phi)}{\partial I_0}\frac{\partial P(\phi)}{\partial I}\right] = \underbrace{cov\left(\frac{\partial L(\phi)}{\partial I_0}, \frac{\partial P(\phi)}{\partial I}\right)}_{<0} + \underbrace{\mathbb{E}\left[\frac{\partial L(\phi)}{\partial I_0}\right]}_{=0}\mathbb{E}\left[\frac{\partial P(\phi)}{\partial I}\right] \tag{46}$$

$$\mathbb{E}\left[\frac{\partial L(\phi)}{\partial I_0}P(\phi)\right] = \underbrace{cov\left(\frac{\partial L(\phi)}{\partial I_0}, P(\phi)\right)}_{<0} + \underbrace{\mathbb{E}\left[\frac{\partial L(\phi)}{\partial I_0}\right]}_{=0}\mathbb{E}\left[P(\phi)\right] \tag{47}$$

In equation (46), we used $\frac{\partial L(\phi)}{\partial I_0}$ decreasing in ϕ and $\frac{\partial P(\phi)}{\partial I}$ increasing in ϕ for the first term, and $\mathbb{E}\left[\frac{\partial L(\phi)}{\partial I_0}\right] = \frac{\partial}{\partial I_0}\mathbb{E}\left[L(\phi)\right] = 0$ since $\mathbb{E}\left[L(\phi)\right] = 2$. In equation (47), we used $\frac{\partial L(\phi)}{\partial I_0}$ decreasing in ϕ and $P'(\phi) \geq 0$ for the first term, and the same equality used in equation (46) for the second term.

Using the definition of f from equation (42), we have:

$$\frac{f''\left(\mathbb{E}[L(\phi)P(\phi)]\right)}{f'\left(\mathbb{E}[L(\phi)P(\phi)]\right)} = \frac{\frac{\lambda}{\lambda-1} - 1}{\mathbb{E}[L(\phi)P(\phi)]} \tag{48}$$

which is strictly decreasing in λ for $\lambda \in (1, \infty)$. Also:

$$\lim_{\lambda \to 1} \frac{f''\left(\mathbb{E}[L(\phi)P(\phi)]\right)}{f'\left(\mathbb{E}[L(\phi)P(\phi)]\right)} = \infty \tag{49}$$

$$\lim_{\lambda \to \infty} \frac{f''\left(\mathbb{E}[L(\phi)P(\phi)]\right)}{f'\left(\mathbb{E}[L(\phi)P(\phi)]\right)} = 0 \tag{50}$$

Applying these results to equation (45) completes the proof. ■

Acknowledgements

Conflicts of interests: none.

Sources of research support: none.