Automation, AI, and the Intergenerational Transmission of Knowledge

Enrique Ide*

March 27, 2026

Abstract

Motivated by concerns that Artificial Intelligence (AI) may deprive new generations of work experience, this paper analyzes how technological change affects the intergenerational transmission of tacit knowledge—practical, hard-to-codify skills acquired through workplace interaction. I develop a task-based overlapping-generations model in which novices acquire tacit knowledge by working alongside experts. Knowledge-transfer contracts are incomplete because the knowledge being transmitted is tacit and embodied. In equilibrium, only the most knowledgeable experts manage production, and endogenous growth arises because they transmit their expertise to multiple novices. I show that improvements in entry-level automation increase short-run output but can reduce long-run growth and welfare, even without reducing entry-level employment. This occurs when such improvements reallocate novices away from the most productive experts, weakening the diffusion of best practices and lowering the quality of future experts. By contrast, technological improvements that expand the span of control of the most productive experts strengthen knowledge transmission and raise growth. The analysis suggests that the span of control of leading experts is a useful summary measure of how technological change affects the transmission of tacit knowledge.

*Department of Economics, IESE Business School, Carrer d'Arnús i de Garí 3-7, 08034 Barcelona, Spain (eide@iese.edu).
I am grateful for the valuable comments and suggestions of David Autor, Heski Bar-Isaac, Matthew Beane, David De la Croix, Andrea Galeotti, Luis Garicano, Jie Gong, Dan Gross, Chad Jones, Anton Korinek, Jin Li, Suraj Malladi, Ben Moll, Cristóbal Otero, Sebastián Otero, Esteban Rossi-Hansberg, Raffaella Sadun, Eduard Talamàs, Chris Tonetti, and seminar participants at Chicago FED, IESE, IMF, Kellogg Strategy, Luohan Academy, and NBER Org. Econ. I also acknowledge the financial support of IESE through the High Impact Initiative-course 2024/2025. I declare I have no relevant or material financial interests that relate to the research described in this paper.

1 Introduction

Recent advances in Artificial Intelligence (AI) have fueled optimistic forecasts of dramatic increases in productivity and economic growth. Unlike earlier technological innovations—such as robots or conventional computers—AI systems can learn and adapt from examples, enabling them to perform work once considered inherently human (Autor, 2024; Brynjolfsson et al., 2025b; Ide and Talamàs, 2025). As a result, many technologists foresee an economic transformation that could rival or even surpass that of the Industrial Revolution (Amodei, 2024; Levy, 2025; Altman, 2025).

However, the widespread adoption of AI may carry unintended long-term costs. By allowing senior workers to accomplish more tasks independently, these technologies may reduce entry-level opportunities and threaten the implicit contract in which younger workers exchange their labor for training and the prospect of future promotion (Roose, *The New York Times*, 2025; The Journal Podcast, *The Wall Street Journal*, 2025; BloombergTV, *Bloomberg.com*, 2025). Such arrangements have historically been crucial for transmitting tacit knowledge—practical insights that cannot be fully articulated yet remain indispensable for carrying out complex work—raising concerns about how future generations will acquire essential expertise (Beane, 2024a,b; Garicano, 2025).

Motivated by these concerns, this paper develops a framework for analyzing how different forms of technological change—automation, task creation, and labor- and capital-augmenting technologies—reshape the intergenerational transmission of tacit knowledge and, through that channel, long-run growth. The central premise is that technological change affects long-run productivity not only by changing how tasks are performed, but also by changing who works with whom and, therefore, who learns from whom.

Existing research on automation and AI has mainly focused on how new technologies directly affect productivity, labor-market outcomes, and the distribution of income and wealth (e.g. Autor et al., 2003; Acemoglu and Autor, 2011; Moll et al., 2022; Acemoglu et al., 2024; Autor and Thompson, 2025; Ide and Talamàs, 2025). However, much less attention has been paid to how these innovations reshape the interpersonal interactions through which valuable skills are transmitted. This paper addresses that gap by embedding different technological advances within a growth model of knowledge diffusion and learning (e.g., Lucas, 2009; Lucas and Moll, 2014; Perla and Tonetti, 2014; De la Croix et al., 2018). In doing so, it provides a framework for analyzing how new technologies jointly affect current output, knowledge transmission, and long-run growth.

The baseline model features overlapping generations, with individuals living for two periods: first as novices, then as experts. Novices are identical at birth, have no wealth, and supply one unit of labor inelastically. Experts, by contrast, differ in skill. This skill level is observable, but the knowledge underlying it is tacit. Outsiders can therefore assess an expert's productivity, but the insights and processes through which it is achieved are accessible only through close interaction.

Production combines expert skill with routine tasks that can be performed either by novices or by machines. Machines are more productive than novices, but can perform only a subset of tasks; the remaining tasks must be carried out by novices. Because each task requires a novice's full unit

of labor, production gives rise to many-to-one matching: each novice works for exactly one expert, while each expert hires multiple novices. By working alongside an expert, a novice gains access to that expert's tacit knowledge and enters the next period with the same skill level as her mentor.

Since future earnings rise with mentor skill, novices are willing to accept lower wages to work alongside more skilled experts. However, experts' skills are tacit: they cannot be fully articulated and are therefore non-verifiable. Moreover, because these skills are tied to personal experience and thus embodied in the individual, they cannot be seized or repossessed. Consequently, novices cannot borrow against future human capital to pay for access to better mentors. Lacking initial wealth, they face a zero lower bound on wages. Thus, concurrent labor becomes the only means of compensating experts for the opportunity to learn, giving rise to an apprenticeship-like arrangement.

In equilibrium, all automatable tasks are assigned to machines, and each expert cohort is partitioned endogenously by skill. The lowest-skilled experts are **inactive** because hiring novices at equilibrium wages is unprofitable.[1] Above them lies an **active region**, which itself splits into two: a *price-cleared* tier in which experts hire novices at strictly positive wages, and an *oversubscribed* tier at the top, where the zero lower bound binds and wages can no longer clear the market for those positions.[2] This endogenous partition then governs the evolution of expertise over time: the next generation's skill distribution is obtained by truncating the current distribution at the marginal active expert and reweighting the surviving upper tail by experts' span of control—the number of novices each active expert employs.

When the initial distribution of expert skill is fat-tailed,[3] this recursive evolution sustains strictly positive asymptotic growth. Active experts replicate their skills by passing them on to multiple novices, while the fat upper tail prevents the pool of mentors from becoming too sparse as the economy moves to higher skill levels. The asymptotic growth rate is therefore pinned down by experts' span of control and the thickness of the initial skill distribution. Furthermore, once skills are normalized relative to the least-skilled expert in each cohort, the asymptotic cross-section converges to a Pareto distribution.

I then use the framework to study several technological shocks, beginning with automation. I first show that expanding the set of automatable tasks can raise output on impact, but can also—through the displacement of novices from high-skill mentors—lower long-run growth. The immediate output gain reflects the reallocation of routine tasks from novices to more productive machines. At the same time, as active experts hire fewer novices, displaced novices are reallocated toward lower-skill experts, allowing a new tier of inactive experts to become active. Some novices are therefore matched

[1] The inactivity of the lowest-skilled experts stems from the assumption that experts cannot supply routine labor. If experts can do so, the lowest-skilled experts need not be inactive, but the main results are unchanged; see Section 6.3.

[2] The oversubscribed region is consistent with the severe rationing of entry-level positions—and the corresponding willingness of junior associates to accept unusually long hours relative to pay—often observed at top professional services firms (see Sections 2 and 4.3 for details).

[3] Although assuming an unbounded, fat-tailed initial skill distribution may appear strong at first sight, it is better interpreted as a reduced-form representation of an economy in which agents with bounded skills draw ideas from an unbounded latent distribution with fat tails ("Plato's realm of Forms" in the language of Buera and Lucas, 2018).

with worse mentors, acquire lower-quality tacit knowledge, and later become worse mentors themselves. The best practices of top experts thus diffuse more slowly, imposing a persistent drag on growth. Importantly, this slowdown does not require the automation improvement to reduce entry-level employment or deprive novices of hands-on practice; a compositional deterioration in the pool of mentors is sufficient.

Through this displacement channel, the automation shock can reduce aggregate welfare, defined as the discounted sum of the income of current and future generations. This inefficiency stems from the zero lower bound on wages. In a frictionless environment, novices would borrow against future earnings to preserve access to socially efficient training—especially since the same automation shock makes future expertise more valuable. Instead, novices cannot directly pay oversubscribed experts for access, so those experts fail to internalize the full social value of mentorship, and automation becomes socially excessive. The zero lower bound thus forces novices to finance elite mentorship solely through their labor, the very currency that automation devalues.

Displacement, however, is only one side of the story. Automation can also affect experts' scale of operation: by reducing the supervision time required per project, it may allow them to oversee more concurrent projects. The resulting increase in experts' scale can more than offset the decline in novices per project, accelerating the diffusion of best practices and raising long-run growth. Even so, at least in these early stages of adoption, the preliminary evidence on generative AI discussed below points more clearly to displacement than to offsetting scale effects. This may be because AI still operate as "so-so" automation (Acemoglu and Restrepo, 2019), generating limited time savings, or because expert supervision is constrained by cognitive or attention limits.

The forces identified above can also be used to understand other forms of technological change. Technologies that create new labor-intensive tasks—for example, data annotation or model evaluation in the case of AI—raise output on impact and increase long-run growth, even without scale effects. They do so by inducing the most skilled experts to hire more novices, thereby accelerating the diffusion of best practices. By contrast, technologies that raise the productivity of novices or machines in the tasks they currently perform increase long-run growth only if they expand experts' scale of operation.

Taken together, these results show that new technologies—especially general-purpose ones—can set multiple, and often competing, forces in motion. Their net effect on knowledge transmission is therefore theoretically ambiguous. The model nevertheless yields a simple organizing principle: what matters for knowledge transmission is whether a technology expands or contracts the span of control of the most skilled experts. Admittedly, this result rests on stark simplifying assumptions, most notably perfect skill inheritance. Even so, the core logic remains: the span of control of top experts is an informative and practical guide to a technology's effect on the intergenerational transmission of tacit knowledge.

While applicable to many technologies, these results are especially relevant to the recent advent of generative AI. Early evidence points to a weakening of entry-level hiring pipelines in AI-exposed, knowledge-intensive activities, precisely where the transmission of tacit skill is most im-

portant (Berger et al., 2024; Hosseini and Lichtinger, 2025; Brynjolfsson et al., 2025a; Klein Teeselink, 2025; Atkinson and Yamco, 2026).[4] The source of this shift remains contested, but the observable consequence is a smaller inflow of novices into exposed sectors. Anecdotal reports are likewise consistent with the mentor-composition mechanism emphasized above, suggesting that graduates are increasingly accepting positions at smaller, mid-tier firms that previously struggled to attract such talent (Ellis, 2026; Lahart, 2026).

In this sense, the early evidence suggests that current iterations of generative AI may constitute a form of "seniority-biased technological change" (Hosseini and Lichtinger, 2025): a technology that substitutes for entry-level execution while complementing expert judgment. Holding expertise fixed, such a change is unambiguously beneficial. But expertise is endogenous, and its intergenerational transmission is governed by incomplete contracts. As a result, seniority-biased technological change can ultimately erode the future supply of the very expertise it complements. Later in the paper, I offer tentative back-of-the-envelope calculations to assess whether this erosion may be quantitatively relevant, and in the Final Remarks, I discuss several policy implications that follow from these results.

1.1 Related Literature

This paper bridges three strands of literature. It connects the task-based literature on technology and labor demand with the literature on idea flows and economic growth. It does so by examining how new technologies reshape the workplace interactions through which tacit skills are transmitted across generations. It also contributes to the literature on the economics of apprenticeships. In particular, it shows how technological change can reshape early-career learning and, through that channel, affect aggregate productivity and long-run growth.

The task-based approach to technology and labor demand was developed by Zeira (1998), Autor et al. (2003), and Acemoglu and Autor (2011). In this framework, production requires the completion of a range of tasks allocated across factors of production according to comparative advantage and available technologies. Technological change therefore affects labor demand, wages, inequality, and productivity by altering who performs which tasks and at what cost (Acemoglu and Restrepo, 2018, 2019, 2022; Acemoglu et al., 2024). Autor and Thompson (2025) extends this framework by modeling occupations as bundles of tasks and showing that automation can be expertise-leveling, lowering an occupation's expertise requirements in a way standard task-based models do not capture.[5]

More recently, Ide and Talamàs (2024, 2025, 2026) build on the literature on knowledge hierarchies developed by Garicano (2000) and Garicano and Rossi-Hansberg (2004, 2006) to study AI's distinc-

[4]Some studies report more optimistic findings. Johnston and Makridis (2025) document employment and wage gains for younger workers following the widespread adoption of AI, though Brynjolfsson et al. (2025a) argue these results may reflect insufficiently granular data. de Souza (2025) finds mixed effects in Brazil: AI appears to automate low-wage administrative jobs while augmenting younger and less-experienced workers in production roles.

[5]Other important contributions that use the task-based approach to study automation, AI, and technological change include Aghion et al. (2017), Moll et al. (2022), Korinek and Suh (2024), Acemoglu and Loebbing (2024), Freund and Mann (2025), Jones and Tonetti (2026), and Demirer et al. (2026). For analyses of AI's effects on output and growth outside the task-based framework, see Gans (2025) and Beraja and Talamàs (2026).

tive implications. Knowledge hierarchies are a specialized version of the task-based framework in which the links between tasks emerge endogenously from organizational choices aimed at deploying tacit knowledge efficiently (Garicano and Rossi-Hansberg, 2015; Ide and Talamàs, 2025). Using this framework, Ide and Talamàs (2024, 2025, 2026) examine AI's implications for occupational choices, organizational structures, labor income, and international trade patterns. Their findings reveal mechanisms absent from conventional task-based analyses, especially the importance of AI autonomy: autonomous AI primarily benefits the most knowledgeable individuals, whereas nonautonomous AI disproportionately benefits the least knowledgeable.

I contribute to the broader task-based literature on technological change by studying how different forms of technological change reshape the intergenerational transmission of tacit knowledge and, thus, future productivity and growth. My model builds more directly on the canonical task-based treatment of technological change. However, it also draws inspiration from the knowledge-hierarchies literature in emphasizing the economic importance of tacit knowledge and the role of organizations in determining who works with whom.

In related work, Kosmyna et al. (2025) provide neuroscientific evidence that reliance on AI for essay writing reduces cognitive engagement and attenuates neural connectivity. More closely related, Acemoglu et al. (2026) develop a dynamic model in which agentic AI substitutes for human learning effort and can erode the stock of public knowledge, potentially generating a knowledge-collapse steady state. I study a different channel. New technologies reshape who learns from whom and thereby alter the diffusion of tacit practices across workers and over time. The focus is therefore on mentor composition, incomplete contracts, and long-run growth, rather than on the cognitive effects of AI use or the erosion of public knowledge through reduced learning effort.

The literature on idea flows and economic growth emphasizes how knowledge diffuses through interactions among individuals and firms, shaping productivity and growth (Kortum, 1997; Eaton and Kortum, 1999, 2002; Lucas, 2009; Lucas and Moll, 2014).[6] My paper contributes to this literature by studying how technological change reshapes who learns from whom in production, and thereby weakens or accelerates the diffusion of tacit practices, a margin absent from existing models.

In the context of idea flows and economic growth, the papers closest to mine are Perla and Tonetti (2014), De la Croix et al. (2018), and Caicedo et al. (2019). In both my paper and Perla and Tonetti (2014), long-run growth arises from the endogenous evolution of a fat-tailed productivity distribution as knowledge diffuses from high- to low-productivity types. In Perla and Tonetti (2014), however, diffusion is anonymous rather than workplace-based: low-productivity firms imitate higher-productivity firms by drawing from the productivity distribution of producing firms. Hence, they have no workplace transmission of knowledge, no scarce-mentor bottleneck, and no role for technological change in reshaping who learns from whom.

Like Caicedo et al. (2019), my paper embeds organizations in a model of idea flows and knowledge diffusion. In Caicedo et al. (2019), however, organizations affect output and wages, not learning. In

[6]Other important contributions include Buera and Oberfield (2020), Jarosch et al. (2021), Perla et al. (2021), and Benhabib et al. (2021). For a survey on the literature, see Buera and Lucas (2018).

my paper, workplace assignment itself shapes learning. De la Croix et al. (2018) is closer on intergenerational transmission: it also studies an overlapping-generations economy in which tacit knowledge passes from experts to novices. But De la Croix et al. (2018) assumes random matching and focuses on preindustrial apprenticeship institutions that mitigate masters' incentives to shirk on training effort. By contrast, my paper assumes that novices compete to learn from the best experts and asks how technological change reshapes who learns from whom.

Finally, my paper contributes to the literature on the economics of apprenticeships. As Mokyr (2019) emphasizes, apprenticeship contracts are inherently incomplete because they govern the transfer of tacit knowledge. That incompleteness can distort training and lead to inefficient knowledge transmission (De la Croix et al., 2018; Fudenberg and Rayo, 2019; Fudenberg et al., 2021).

I contribute to this literature by studying how technological change reshapes apprenticeship-like arrangements and, through them, the intergenerational transmission of tacit knowledge. Closest is the complementary work of Garicano and Rayo (2025), which extends Garicano and Rayo (2017) to analyze how AI affects apprenticeship viability. Their analysis, however, focuses on the dynamics of a single expert–novice relationship. By contrast, my paper studies how technological change reshapes apprenticeship-like arrangements in general equilibrium, altering the allocation of novices across experts, the intergenerational transmission of tacit knowledge, and ultimately long-run growth.

1.2 Roadmap

The remainder of the paper is organized as follows. Section 2 lays out the conceptual foundations of the analysis. Section 3 presents the baseline model and defines competitive equilibrium. Section 4 characterizes the long-run equilibrium and the conditions under which the economy sustains growth. Section 5 studies technological shocks in the baseline model. Section 6 develops the main extensions and offers tentative quantitative illustrations. Section 7 concludes.

2 Conceptual Foundations

This section outlines the conceptual foundations of the analysis: the tacit dimension of expertise, the inherent constraints on its transfer, and the mechanics of its transmission across generations. Section 3 formalizes these concepts to analyze how technological shocks impact long-term productivity and economic growth.

The Tacit Dimension of Expertise.— Codifiable knowledge refers to explicit rules and procedures that can be easily transmitted via manuals, books, and databases. Tacit knowledge, in contrast, encompasses intuitive skills and insights that are difficult to articulate precisely (Polanyi, 1966; Foray, 2004). These two forms of knowledge are interdependent: codified instructions usually require tacit knowledge to be properly interpreted and adapted to real-world complexities. As Mokyr (2002, pp. 14-15) explains:

> Tacit knowledge is needed to obtain inexpensive, reliable access to codified instructions [...] no set of instructions [...] can ever be complete. It would be too expensive to write a complete set of instructions for every technique. Judgement, dexterity, experience, and other forms of tacit knowledge inevitably come into play when technique is executed.

While professional capability relies on both forms of knowledge, codifiable elements can be standardized and scaled. This paper therefore focuses on the tacit dimension. Hereafter, I use "expertise" and "tacit knowledge" interchangeably to refer to the unwritten insights required to execute and orchestrate well-defined tasks. In investment banking, for instance, populating financial models and verifying due diligence documents are largely standardized procedures. However, transforming their completion into financial advice requires the judgment to select appropriate models, identify anomalies that standard procedures might overlook, and adjust methods when unforeseen contingencies arise.

Because tacit knowledge is tied to personal experience, it is embodied in the individual (Polanyi, 1966; Garicano, 2000). Thus, while tacit insights can be transferred, their transmission requires the expert's direct involvement. This makes the expert's time a key bottleneck to the use and diffusion of this knowledge (Foray, 2004; Garicano and Rossi-Hansberg, 2015; Ide and Talamàs, 2025).

Labor as a Mechanism and Currency for Acquiring Tacit Knowledge.— Lacking a codifiable structure, tacit knowledge must be acquired experientially through active execution and collaboration with more experienced individuals. As Polanyi (1962, p. 55) observes:

> An art which cannot be specified in detail cannot be transmitted by prescription [...] It can be passed on only by example from master to apprentice [...] By watching the master and emulating his efforts, the apprentice unconsciously picks up the rules of the art, including those which are not explicitly known to the master himself.

Thus, labor supplied alongside an expert serves a dual role: it is both a productive input and a mechanism for assimilating tacit knowledge (Lave and Wenger, 1991; Brown and Duguid, 1991; Beane, 2019). This apprenticeship model is pervasive in expertise-intensive professions (e.g., medicine, law, finance), where novices perform routine tasks under expert supervision, observing how the expert applies judgment to integrate these inputs into a cohesive whole (Garicano and Rayo, 2017; Beane, 2024b).

Independent trial-and-error can also help acquire these tacit skills, but remains an inefficient substitute for mentorship. This is evident in white-collar entry-level labor markets, where novices routinely accept demanding hours and low effective pay to work alongside top practitioners.[7] While this phenomenon could also reflect other mechanisms—such as novices trying to acquire a prestigious signal of high ability—Oyer (2006) and Arellano-Bover (2024) support knowledge acquisition

[7]For example, junior investment bankers routinely work 100- to 110-hour weeks, systematically underreporting to evade 80-hour institutional limits (Glickman and Saeedy, 2025; Saeedy, 2024b). In extreme cases, the physical toll of these conditions has proved fatal, as seen in the exhaustion-related death of an intern in 2013 after consecutive overnight shifts and the 2024 death of an associate following repeated 100-hour workweeks (Saeedy, 2024a).

as the primary driver. Both studies provide causal evidence that initial placement at elite institutions leads to higher long-term productivity and lifetime earnings driven by superior skill development.[8]

The contracts governing the transfer of tacit knowledge, however, are inherently incomplete (Mokyr, 2019). Because such expertise resists precise articulation, it is non-verifiable. Moreover, because it is embodied within the individual, it cannot be seized or repossessed. Consequently, novices cannot credibly pledge future human capital to finance this type of training (Becker, 1964; Garicano and Rayo, 2017). Concurrent labor thus remains the only viable means of payment.

Intergenerational Transmission of Tacit Knowledge.— Because tacit knowledge is embodied in individuals and shared primarily through joint production, its preservation requires the continual renewal of practitioners (Foray, 2004). The transmission of tacit knowledge thus operates as an intergenerational cycle: novices acquire this experiential knowledge by working alongside experts and subsequently train the next cohort. When novices learn from the most productive experts, this mechanism diffuses best practices over time, driving aggregate productivity growth.

As the historical embodiment of this dynamic, apprenticeship systems have functioned as central engines for economic growth. Humphries (2006), for example, identifies England's apprenticeship system as a "neglected factor" underpinning the British Industrial Revolution. Similarly, De la Croix et al. (2018) argue that European apprenticeship institutions significantly contributed to the continent's relative rise in technological creativity, population growth, and per-capita income in the centuries preceding the Industrial Revolution.

Together, the features outlined above characterize a distinctive environment that is central to macroeconomic outcomes. Moreover, by altering the conditions in which learning occurs—and the value of the labor used to pay for it—technological advances can affect skill accumulation, impacting both current and future productivity. Section 3 develops a model to study these issues.

3 The Baseline Model

Section 3.1 introduces the environment and maps the model's key assumptions to the conceptual foundations of Section 2. Section 3.2 then describes agents' problems, defines the competitive equilibrium, and establishes its existence and uniqueness. Doing so upfront allows the analysis to focus directly on long-run outcomes and the effects of technological shocks. Finally, Appendix A provides a glossary of mathematical symbols used throughout the paper.

3.1 The Environment and Its Interpretation

A Broad Overview of the Environment.— Time is discrete and infinite, indexed by $t = 0, 1, 2, \ldots$. The economy consists of overlapping generations of risk-neutral individuals who live for two periods:

[8] Furthermore, related benefits of elite placement—such as professional connections and access to exclusive networks—can be broadly interpreted as part of tacit knowledge transfer.

first as novices, then as experts. Each cohort has unit mass, so the total population remains constant at two. All agents discount future payoffs by the common discount factor $\beta \in (0, 1)$.

Novices are identical at birth, lack wealth, and inelastically supply one unit of labor. Experts, in contrast, differ in their skill q. This skill level is fully observable, but the knowledge behind it is tacit. Consequently, external observers can perfectly assess an expert's productivity but cannot discern the insights or processes that drive those results without closer interaction.

Let $F_t(q)$ denote the cumulative distribution function (CDF) of expert skills at time t, with $x_t \equiv \min \text{support}\{F_t\}$. This distribution is the aggregate state of the economy and evolves endogenously as novices work alongside experts and assimilate their expertise. The initial distribution F_0 is exogenously given, has a finite mean, and admits a continuous and strictly positive probability density function $f_0(q)$ on its support $[x_0, +\infty)$.

There is a single final good, which serves as the numeraire. An expert produces output by combining her skill with a continuum of routine tasks in the $[0, N]$ interval. Tasks are not scalable: each task must be completed exactly once per expert, either by a single unit of novice labor or, when feasible, by a single machine. For simplicity, experts do not supply routine labor themselves.

Novices can perform all tasks and contribute a factor h per task to output, while machines can perform only a subset $[0, I] \subseteq [0, N]$ of tasks but contribute a larger factor $m > h$ per task. Output is the product of the expert's skill and the contributions of all completed tasks. Accordingly, if an expert with skill q decides to produce (is "active") and assigns a feasible measure $k \in [0, I]$ of automatable tasks to machines—and the remaining tasks to novices—her output is given by:[9]

$$qm^k h^{N-k}$$

I restrict attention to the case in which experts require multiple novices to produce. As I formally show in Section 4.1, this assumption is necessary for the economy to exhibit long-run growth:

Assumption 1 (Experts Need Multiple Novices)**.** *The measure of non-automatable tasks exceeds one:*

$$N - I > 1$$

Following Acemoglu et al. (2024), each machine is produced instantaneously using $\rho > 0$ units of the final good and fully depreciates after one period, implying a rental rate of ρ. On the labor side, there is many-to-one matching: each novice works for exactly one expert, dedicating her single unit of labor to complete one of that expert's routine tasks, while each active expert hires multiple novices to execute all tasks not assigned to machines (a measure strictly greater than one, per Assumption 1).

By working alongside an expert, a novice observes and assimilates the tacit insights behind the expert's skill. Hence, a novice matched with an expert of skill q becomes an expert of identical skill in the subsequent period. Anticipating higher future income from superior training, novices are therefore willing to accept lower wages to work alongside more skilled experts, yielding a decreasing wage schedule, $w_t(q)$.

[9]See Appendix B.1 for the relationship between this specification and standard task-based models (e.g. Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2018, 2019). Unlike those settings, tasks in this framework are non-scalable.

However, the non-verifiability of tacit knowledge and the inalienability of the resulting human capital prevent novices from pledging future earnings as collateral. Lacking initial wealth, they face a liquidity constraint that forces wages to be non-negative ($w_t(q) \geq 0$). Thus, concurrent labor becomes the only means of compensating an expert for the opportunity to learn, giving rise to an apprenticeship-like arrangement.

The economic environment is thus parsimoniously described by a small set of primitives: the discount factor β, the initial distribution of skills F_0, the rental rate of machines ρ, the parameters of the production technology (N, I, h, m), and the zero-lower bound on wages.

Interpretation.— Having described the economic environment, I now relate the model's core assumptions to the conceptual framework of Section 2.

First, the production process features a strict division of labor between execution and judgment: novices and machines perform routine tasks, while experts apply tacit skills. This structure mirrors the reality of knowledge-intensive industries, where early-career professionals (and, more recently, AI) execute well-structured tasks, and senior practitioners apply judgment to integrate these inputs into a cohesive whole. While imposing this exogenous division of labor simplifies the baseline analysis, I explain in Section 6 that the main results hold when experts endogenously choose whether to orchestrate production or supply routine labor.

Second, the assumption that tasks are non-scalable reflects the limited time and attention of experts: each expert can only oversee the measure N of tasks required for a single production process.[10] This formalizes a core principle of Section 2: because expertise is embodied in the individual, its application and transmission cannot be easily scaled. In investment banking, for instance, a senior partner's capacity to review financial models and direct junior associates is strictly bounded; they can oversee only a finite number of concurrent transactions, regardless of skill. Admittedly, a strictly fixed capacity is a stark simplification. Section 6 therefore explores the implications of relaxing it.

Third, the model assumes that joint production is necessary for skill acquisition. By executing routine tasks, a novice gains intimate insight into how the expert transforms these inputs into output, unconsciously learning "the rules of the art." As a first approximation, I abstract from the finer details of this process, assuming that all routine tasks provide uniform insight into the expert's methods and that novices perfectly inherit their mentor's skill. To capture skill development beyond direct mentorship, Online Appendix V also extends the framework to incorporate independent innovation.

Fourth, because novices lack initial wealth and cannot borrow, labor is their only means to pay for training ($w_t(q) \geq 0$). Rather than explicitly microfounding the contractual frictions that preclude such borrowing, I impose this non-negativity constraint as a reduced-form proxy for the contract incompleteness inherent in the transfer of tacit knowledge, as discussed in Section 2.[11] A zero wage

[10] Given the assumption of non-scalability, the fundamental unit of production is best interpreted as an expert-led team, rather than a firm that expands by hiring labor. In this framework, firm boundaries are indeterminate—a "firm" is simply an aggregation of independent teams. The model therefore offers no rationale for multi-team firms, making the single-expert team the relevant unit of analysis.

[11] For instance, the unverifiable and embodied nature of expertise might prevent novices from credibly committing to

can be interpreted literally as unpaid labor (e.g., culinary *stages*), or more broadly as novices accepting demanding hours relative to nominal pay, resulting in low effective compensation (e.g., medical residencies, or junior associates in law and finance).[12]

Fifth, the model assumes that an expert's skill level is observable, but the insights behind it remain hidden from the outside. This assumption aligns with the observation that a novice in a white-collar profession often sacrifices current compensation to work alongside a top practitioner (see Section 2). To evaluate the role of this directed sorting, Online Appendix III examines the opposite scenario, where skill levels are unobservable and matching is completely random.

Finally, the initial skill distribution, F_0, is assumed to have unbounded support. Following Buera and Lucas (2018), this assumption serves as a reduced-form approximation for an economy where an initial population with bounded skills draws ideas from an unbounded latent distribution ("Plato's realm of Forms"). Online Appendix IV formalizes this extension.

3.2 Optimization, Matching, and Equilibrium

Experts' Problem.—Given the machine rental rate ρ and the wage schedule w_t, an expert with skill q at time t maximizes final-period income. The expert chooses whether to produce or remain inactive. If active, the expert assigns a measure $k \in [0, I]$ of automatable tasks to machines. Optimal income is therefore:

$$\Pi_t(q) = \max\left\{0, \max_{k\in[0,I]}\left[qm^k h^{N-k} - \rho k - w_t(q)(N-k)\right]\right\}$$

The following lemma establishes a central result for the subsequent analysis.

Lemma 1. *Even if novice labor is free ($w_t(q) = 0$), full automation ($k = I$) is strictly optimal for all experts with skill:*

$$q \geq \frac{\rho I}{h^N[(m/h)^I - 1]}$$

Proof. See Appendix B.2. □

Intuitively, high-skill experts benefit more from the productivity premium of machines ($m > h$), increasing their incentive to automate. A countervailing force is that more skilled experts also pay lower wages, increasing the incentive to use novice labor. The zero lower bound on wages, however, limits this labor cost advantage. As a result, for sufficiently skilled experts, the productivity premium of machines dominates, ensuring full automation. This intuition is central to understanding the impact and welfare effects of automation improvements in Section 5.

I then impose the following assumption:

---

future payments, whether in money or kind, to financial institutions or the experts themselves (Becker, 1964; Garicano and Rayo, 2017, 2025). Additionally, external borrowing may be restricted because neither the expert's mentorship nor the novice's learning effort is verifiable.

[12]Note that the model abstracts from the intensive margin of labor supply; this is purely an out-of-model interpretation.

Assumption 2 (Full Adoption of Automation Technologies)**.** *The lower bound of the initial skill distribution, x_0, satisfies:*

$$x_0 \geq \frac{\rho I}{h^N[(m/h)^I - 1]}$$

Since all experts in subsequent periods possess skills $q \geq x_0$ (as shown below), all active experts, regardless of their skill, optimally assign I tasks to machines and hire $N - I$ novices.

As formally shown in the Online Appendix I.2, this restriction entails no loss of generality for the asymptotic analysis. Because skill accumulation drives the minimum skill in period t above the full-automation skill threshold in finite time, relaxing this assumption only alters the early transition dynamics.

Given Assumption 2, the optimal income of expert q at time t reduces to:

$$\Pi_t(q) \equiv \max\{0,\ S(q) - w_t(q)(N - I)\} \tag{1}$$

where $S(q) \equiv qm^I h^{N-I} - \rho I$ is the surplus generated by an active expert with skill q before novice compensation. Expert q at time t is then active if and only if $S(q) \geq w_t(q)(N - I)$.

I restrict attention to equilibria in which the expert income schedule, $\Pi_t(q)$, is continuous and non-decreasing in q. The continuity restriction merely streamlines the exposition; one can show that any equilibrium featuring a non-decreasing income schedule is necessarily continuous. The monotonicity requirement, in turn, selects outcomes robust to the free disposal of skills—where an expert can costlessly mimic a less-skilled peer. This monotonicity immediately implies that the set of active experts at time t, denoted by A_t, is an upper interval:

$$A_t \equiv \{q \geq x_t : S(q) \geq w_t(q)(N - I)\} = [a_t, \infty)$$

Finally, I assume that:

Assumption 3 (Viability with Free Labor)**.** *The parameters are such that:*

$$S(x_0) = x_0 m^I h^{N-I} - \rho I \geq 0$$

Since all experts in subsequent periods possess skills $q \geq x_0$ (as shown below), this condition ensures that all experts would choose to be active if novice labor were free.

Novices' Problem.— Given the current wage schedule w_t and next period's expert income schedule Π_{t+1}, a novice at time t aims to maximize lifetime income. Because novices supply one unit of labor inelastically, they simply seek to match with the active expert delivering the highest total payoff. Specifically, working for an expert of skill q yields:

$$w_t(q) + \beta \Pi_{t+1}(q)$$

where $\Pi_{t+1}(q)$ is the optimal income of an expert with skill q in period $t + 1$, as defined by (1).

Within-Period Competitive Matching.— Having established experts' and novices' problems, I characterize the equilibrium outcome at time t, taking F_t and Π_{t+1} as given. This requires determining the wage schedule and matching arrangement that prevails in that period.

I assume a competitive and frictionless matching market, except for the zero lower bound on wages: there are no search frictions, informational asymmetries, or coordination failures. Expert-novice matching at time t can therefore be represented as an assignment from novices to the set of active experts A_t, with inactive experts receiving no novices. Because novices are ex ante identical, the assignment is summarized by a finite Borel measure $\mathcal{L}_t$ on A_t, where $\mathcal{L}_t(E)$ denotes the mass of novices assigned to experts in the measurable set $E \subseteq A_t$.

A feasible assignment must satisfy two conditions. First, each active expert optimally hires exactly $N - I$ novices. Thus,

$$\mathcal{L}_t(E) = (N - I) \int_E dF_t(q), \quad \text{for any measurable } E \subseteq A_t \tag{2}$$

Second, novices have unit mass and supply labor inelastically, so all novices must be assigned: $\mathcal{L}_t(A_t) = 1$. Since $A_t = [a_t, \infty)$, combining both feasibility conditions implies the following aggregate market-clearing condition:

$$1 = (N - I)\big[1 - F_t(a_t)\big]. \tag{3}$$

The next definition adapts the canonical static, frictionless assignment framework (e.g., Becker, 1973; Eeckhout and Kircher, 2011, Section 2.1) to an environment with a zero lower bound on wages:

Definition 1 (Within-Period Competitive Matching)**.** *Given F_t and Π_{t+1}, a within-period competitive matching consists of a set of active experts $A_t = [a_t, \infty)$, an assignment $\mathcal{L}_t$ on A_t, and a non-negative wage schedule $w_t(q)$ such that:*

1. ***Feasibility:*** *The assignment satisfies (2) and (3).*
2. ***Expert Participation:***

$$\Pi_t(q) = \begin{cases} 0 & \text{if } q \notin [x_t, a_t) \\ S(q) - (N - I)w_t(q) \geq 0 & \text{if } q \in [a_t, \infty) \end{cases}$$

3. ***Stability:*** *There do not exist an expert $q' \in [x_t, \infty)$, a wage $w' \geq 0$, and an active expert $q \in [a_t, \infty)$ such that:*

$$S(q') - (N - I)w' > \Pi_t(q') \quad \textit{and} \quad w' + \beta\Pi_{t+1}(q') > w_t(q) + \beta\Pi_{t+1}(q).$$

Online Appendix II.1 establishes that a unique within-period competitive matching exists if Π_{t+1} is continuous and non-decreasing, and F_t admits a continuous and strictly positive density on its support $[x_t, \infty) \subseteq [x_0, \infty)$.[13] In equilibrium, the aggregate market-clearing condition (3) uniquely determines the active set $A_t = [a_t, \infty)$, and the assignment $\mathcal{L}_t$ follows immediately from (2).

[13] As the law of motion for F_t derived below establishes, equilibrium dynamics ensure that F_t inherits these properties from the initial distribution, F_0.

Because $N - I > 1$ (Assumption 1), only a strict subset of experts is active in each period, i.e., $a_t > x_t$. This implies that the marginal active expert earns zero income, $\Pi_t(a_t) = 0$, yielding a wage of $w_t(a_t) = S(a_t)/(N - I)$. The strict monotonicity of $S(\cdot)$ and Assumption 3 guarantee this wage is strictly positive ($S(a_t) > S(x_t) \geq S(x_0) > 0$). The lifetime income of a novice ("novice utility") matched with expert a_t at time t is then:

$$U_t \equiv w_t(a_t) + \beta\Pi_{t+1}(a_t) = \frac{S(a_t)}{N - I} + \beta\Pi_{t+1}(a_t)$$

The stability of the competitive matching also requires the equilibrium wage schedule to satisfy:

$$w_t(q) = \begin{cases} S(a_t)/(N - I) & \text{if } q \in [x_t, a_t) \\ \max\{0, U_t - \beta\Pi_{t+1}(q)\} & \text{if } q \in [a_t, \infty) \end{cases} \tag{4}$$

For inactive experts ($q < a_t$), the wage ensures they cannot profitably operate.[14] For active experts ($q \geq a_t$), the schedule implies that those with $\beta\Pi_{t+1}(q) < U_t$ pay strictly positive wages and deliver novices the baseline utility U_t. Conversely, experts with $\beta\Pi_{t+1}(q) \geq U_t$ pay zero wages and deliver a lifetime income strictly exceeding U_t. Consequently, U_t represents the minimum lifetime income any novice can secure at time t.

Although the framework abstracts from explicit "queuing" or "rationing" by specifying only a realized assignment rather than an explicit matching protocol, these two cases admit a natural interpretation. Positions that deliver exactly U_t are *price-cleared*: competition drives wages down to offset the future value of superior training, leaving novices indifferent among these positions. In contrast, positions delivering strictly more than U_t are *oversubscribed*: the zero lower bound prevents further adjustment. Novices therefore strictly prefer these matches, generating latent "excess demand."

Intertemporal Linkages.— Having characterized the equilibrium at time t given F_t and Π_{t+1}, I now turn to the intertemporal linkages.

The first linkage is the backward-looking relationship linking future expert incomes to the wages paid and incomes earned by current experts. Because novices inherit their mentor's skill, experts extract the value of their mentorship through lower wages, subject to the zero lower bound. Substituting the equilibrium wage schedule (4) into the optimal income equation (1) formalizes this dependence:

$$\Pi_t(q) = \max\{0, S(q) - (N - I)\overbrace{\max\{0, U_t - \beta\Pi_{t+1}(q)\}}^{=w_t(q)}\}$$

This recursion also establishes that the monotonicity of the experts' income schedule is dynamically consistent. Specifically, if $\Pi_{t+1}(q)$ is non-decreasing in skill, $w_t(q)$ is non-increasing, meaning that novices are willing to accept lower wages to match with more skilled experts. Because the surplus $S(q)$ is strictly increasing, this implies that current expert income, $\Pi_t(q)$, is necessarily non-decreasing. Monotonicity of the income schedule thus propagates backward in time.

[14] Strictly speaking, equilibrium only bounds the wages of inactive experts ($w_t(q) \geq S(q)/(N - I)$), so the exact schedule for $q < a_t$ is indeterminate. Setting $w_t(q) = w_t(a_t) = S(a_t)/(N - I)$ is without loss of generality, as this shadow wage deters participation by any $q < a_t$.

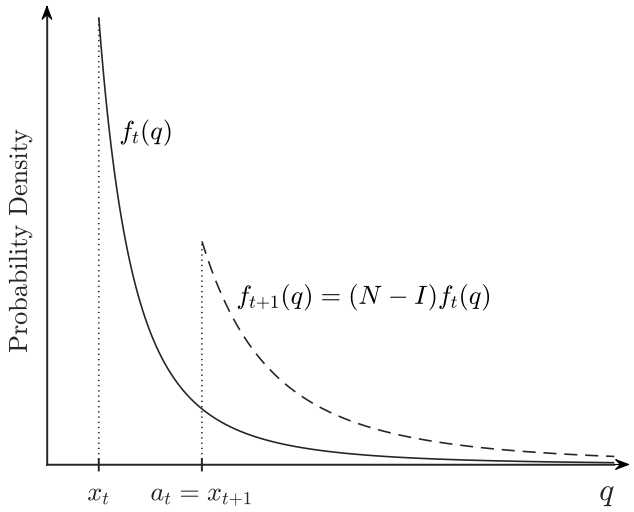


Figure 1: The Evolution of the Distribution of Experts' Skills over Time

The second linkage is the forward-looking law of motion for the skill distribution, F_t. Because each novice inherits her mentor's skill, and active experts $q \in A_t = [a_t, \infty)$ each hire $N - I$ novices, the law of motion for the skill distribution is:

$$F_{t+1}(q) = \begin{cases} 0 & \text{if } q \leq a_t \\ (N-I)\left[F_t(q) - F_t(a_t)\right] & \text{if } q > a_t \end{cases} \tag{5}$$

Equation (5) specifies a well-defined CDF, with the aggregate market-clearing condition (3) ensuring F_{t+1} integrates to one. Figure 1 plots the probability density functions $f_t(q) \equiv dF_t(q)/dq$ and $f_{t+1}(q) \equiv dF_{t+1}(q)/dq$. The density f_{t+1} results from truncating f_t at the marginal active expert a_t and scaling the remaining density by the experts' *span of control*, $N - I$. This truncation implies the support of F_{t+1} is $[x_{t+1}, \infty) = [a_t, \infty) \subset [x_t, \infty) \subset [x_0, \infty)$. Moreover, because the initial distribution F_0 admits a continuous and strictly positive density on $[x_0, \infty)$, equation (5) guarantees F_t maintains these properties on its support for all $t > 0$.

The law of motion dictates that the minimum skill of the $t+1$ cohort equals the skill of the marginal active expert at time t, $x_{t+1} = a_t$. Furthermore, the within-period competitive matching establishes that only a strict subset of experts is active, $a_{t+1} > x_{t+1}$. Together, these results imply a strictly increasing sequence of entry cutoffs, $a_{t+1} > a_t$. A novice inheriting the marginal skill a_t at time t therefore becomes inactive as an expert in the subsequent period, yielding $\Pi_{t+1}(a_t) = 0$. Consequently, the baseline novice utility at time t, U_t, simplifies to:

$$U_t \equiv w_t(a_t) + \beta \overbrace{\Pi_{t+1}(a_t)}^{=0} = \frac{S(a_t)}{N-I} \tag{6}$$

Dynamic Competitive Equilibrium.— Given the current skill distribution F_t and anticipated continuation income schedule Π_{t+1}, the within-period matching determines the active set $A_t = [a_t, \infty)$, baseline novice utility U_t, wage and income schedules w_t and Π_t, and assignment $\mathcal{L}_t$ at period t. The

active set then dictates the evolution of the state variable F_{t+1}, which in turn determines the subsequent matching market and the continuation income Π_{t+1}. A dynamic competitive equilibrium requires this sequence to be mutually consistent for all $t \geq 0$.

Definition 2 (Dynamic Competitive Equilibrium)**.** *A dynamic competitive equilibrium consists of the following sequences.*

- *Skill distributions $\{F_t\}_{t\geq 0}$, evolving according to the law of motion (5) with F_0 given.*
- *Marginal active experts $\{a_t\}_{t\geq 0}$, satisfying the aggregate market-clearing condition (3).*
- *Assignments $\{\mathcal{L}_t\}_{t\geq 0}$, constructed using (2).*
- *Baseline novice utilities $\{U_t\}_{t\geq 0}$ alongside wage and income schedules $\{w_t, \Pi_t\}_{t\geq 0}$, jointly determined by (1), (4), and (6).*

The equilibrium features a block-recursive structure. Given the current distribution F_t, the aggregate market-clearing condition (3) uniquely determines the active cutoff a_t. Because F_t and a_t fully characterize F_{t+1} via equation (5), the dynamical system of allocations $\{a_t, F_t\}_{t\geq 0}$ evolves autonomously forward in time. In contrast, equilibrium prices—utilities, wages, and incomes—are forward-looking. Because they depend on future states, they cannot be determined from the current state alone. How this system is solved becomes evident in Section 4.3.

For now, I simply state the existence and uniqueness of the equilibrium, directing the interested reader to Online Appendix II.2 for the formal argument. With these technical foundations secured, Section 4 proceeds directly to characterize the economy's long-run outcomes.

Proposition 1 (Existence and Uniqueness of a Dynamic Competitive Equilibrium)**.** *For any initial distribution F_0 satisfying the model's assumptions, there exists a unique dynamic competitive equilibrium within the class where expert income $\Pi_t(q)$ is continuous and non-decreasing in q. This equilibrium is block-recursive: the sequence of allocations $\{a_t, F_t\}_{t\geq 0}$ is determined independently of the sequence of prices $\{U_t, w_t, \Pi_t\}_{t\geq 0}$. Moreover, for all $t \geq 0$, the wage schedule $w_t(q)$ is non-increasing in q.*

Proof. See Online Appendix II.2 for the constructive proof. Online Appendix II.3 additionally characterizes the cross-sectional properties of the equilibrium at any given date t. □

Aggregate Output and Welfare.— Since only active experts produce, aggregate output at time t is:

$$Y_t \equiv \left[\int_{a_t}^{\infty} q dF_t(q)\right] m^I h^{N-I} \tag{7}$$

Welfare at time t is defined as the discounted sum of the aggregate income of novices and experts from period t onward:

$$W_t \equiv \sum_{j=t}^{\infty} \beta^{j-t} \left[\int_{a_j}^{\infty} S(q) dF_j(q)\right] = \sum_{j=t}^{\infty} \beta^{j-t} \Big(Y_j - \rho I[1 - F_j(a_j)]\Big) \tag{8}$$

This is also equivalent to the discounted sum of aggregate production minus machine costs.

4 Long-Run Equilibrium Outcomes

This section characterizes the economy's long-run behavior. I show that sustaining strictly positive output growth requires the initial skill distribution F_0 to have fat tails. When this condition holds, the asymptotic growth rate converges to a constant that depends on both the experts' span of control and the tail thickness of F_0. I then derive the asymptotic distribution of normalized expert skills, the limiting wage and income schedules, and the long-run baseline novice utility.

These results set the stage for Section 5, which examines the implications of different technological advances.

4.1 Growth Rate of Output

For fixed technology parameters, aggregate output in (7) depends solely on the skill distribution F_t and the entry cutoff a_t. As established in Section 3.2, both objects are determined independently of equilibrium prices: the sequence $\{a_t, F_t\}_{t\geq 0}$ solves the dynamical system defined by the market-clearing condition (3) and the law of motion (5). Given F_0, the unique solution is:

$$\begin{aligned} 1 &= (N-I)^{t+1}[1-F_0(a_t)], \text{ for } t \geq 0 \\ 1-F_t(q) &= (N-I)^t[1-F_0(q)], \text{ for } t \geq 1 \text{ and } q \geq a_t \end{aligned} \tag{9}$$

Since $N - I > 1$ and F_0 has unbounded support, the sequence $\{a_t\}_{t\geq 0}$ is strictly increasing and diverges as $t \to \infty$. Substituting this solution into (7) yields:

$$\begin{aligned} Y_t &= m^I h^{N-I}(N-I)^t[1-F_0(a_t)]\,\mathbb{E}_0[q \mid q \geq a_t] \\ &= \frac{m^I h^{N-I}}{N-I}\,\mathbb{E}_0[q \mid q \geq a_t] \end{aligned}$$

where the second equality follows from (9), and $\mathbb{E}_0[\cdot]$ denotes the expectation with respect to the initial distribution F_0.

Because $a_t \to \infty$ and $\mathbb{E}_0[q \mid q \geq a_t] \geq a_t$, the truncated mean diverges and aggregate output increases without bound ($Y_t \to \infty$). Sustaining a strictly positive long-run growth rate, however, depends on the right-tail thickness of the initial distribution. To formalize this, I classify the tail behavior of F_0 using regular variation:

Definition 3 (Tail Behavior)**.**

1. ***Regularly Varying (Fat Tails).*** *$1 - F_0$ is regularly varying with index $-1/\theta$ (denoted $1 - F_0 \in \mathrm{RV}_{-1/\theta}$) if, for all $\lambda > 0$,*
$$\lim_{q\to\infty} \frac{1-F_0(\lambda q)}{1-F_0(q)} = \lambda^{-1/\theta}.$$
2. ***Rapidly Varying (Thin Tails).*** *$1 - F_0$ is rapidly varying (denoted $1 - F_0 \in \mathrm{RV}_{-\infty}$) if, for all $\lambda > 1$,*
$$\lim_{q\to\infty} \frac{1-F_0(\lambda q)}{1-F_0(q)} = 0.$$

Remark 1. *Under regular variation, finite mean requires $0 < \theta < 1$. I maintain this restriction throughout.*

Remark 2. *Regular variation corresponds to power-law (Pareto-type) tails, as in Pareto and Fréchet distributions. Rapid variation corresponds to tails that decay faster than any power law, including exponential, Weibull, and log-normal distributions.*

As shown in Embrechts et al. (1997, Appendix A.3) and Perla and Tonetti (2014), these tail classes imply the following asymptotic limits for conditional means:

$$\lim_{a\to\infty} \frac{\mathbb{E}_0[q \mid q \geq a]}{a} = \begin{cases} \dfrac{1}{1-\theta} & \text{if } 1 - F_0 \in \mathrm{RV}_{-1/\theta} \\ 1 & \text{if } 1 - F_0 \in \mathrm{RV}_{-\infty}. \end{cases} \tag{10}$$

These limits yield the following result:

Proposition 2 (Asymptotic Growth Rate of Output)**.** *In equilibrium,*

$$\lim_{t\to\infty} \frac{a_{t+1}}{a_t} = \lim_{t\to\infty} \frac{Y_{t+1}}{Y_t} \equiv 1 + g = \begin{cases} (N-I)^\theta & \text{if } 1 - F_0 \in \mathrm{RV}_{-1/\theta} \\ 1 & \text{if } 1 - F_0 \in \mathrm{RV}_{-\infty} \end{cases}$$

Proof. See Appendix C.1. □

Although the formal proof is in the appendix, I provide a brief sketch of the mechanics behind the proposition here. First, note that (9) implies the following relationship between successive cutoffs:

$$\frac{1 - F_0(a_{t+1})}{1 - F_0(a_t)} = \frac{1}{N - I}.$$

Define the cutoff growth factor by $\lambda_t \equiv a_{t+1}/a_t > 1$. If $1 - F_0$ is regularly varying with index $-1/\theta$, then, since $a_t \to \infty$, regular variation yields:

$$\frac{1 - F_0(a_{t+1})}{1 - F_0(a_t)} = \frac{1 - F_0(\lambda_t a_t)}{1 - F_0(a_t)} \approx \lambda_t^{-1/\theta} \text{ for large } t.$$

Combining this approximation with the cutoff equation above implies $\lambda_t^{-1/\theta} \to 1/(N-I)$ as $t \to \infty$, and hence $a_{t+1}/a_t \to (N-I)^\theta = 1 + g$.

To obtain the asymptotic growth rate of output, decompose the ratio Y_{t+1}/Y_t using $Y_t \propto \mathbb{E}_0[q \mid q \geq a_t]$. Since the cutoff diverges ($a_t \to \infty$), the limiting behavior of the mean-to-cutoff ratios as $t \to \infty$ is governed by (10). Consequently:

$$\lim_{t\to\infty} \frac{Y_{t+1}}{Y_t} = \left(\lim_{t\to\infty} \frac{a_{t+1}}{a_t}\right)\left(\lim_{t\to\infty} \frac{\mathbb{E}_0[q \mid q \geq a_{t+1}]}{a_{t+1}}\right)\left(\lim_{t\to\infty} \frac{a_t}{\mathbb{E}_0[q \mid q \geq a_t]}\right) = 1 + g$$

The rapidly varying case can be viewed heuristically as the limit $\theta \to 0$ of the regularly varying case, for which the growth factor $1 + g = (N-I)^\theta$ converges to one. Consistent with this intuition, if $1 - F_0$ is rapidly varying, then $a_{t+1}/a_t \to 1$ and $Y_{t+1}/Y_t \to 1$ as $t \to \infty$. This completes the sketch.

The forces underlying this result are as follows. In this economy, output grows because the most skilled experts each transfer their skill to $N - I > 1$ novices. This replication expands the mass of high-skilled experts over time, raising the skill of the marginal active expert a_t.

Replication alone, however, is insufficient for strictly positive asymptotic growth. Sustained growth requires the cutoff a_t to rise at a constant proportional rate, which depends on the thickness of the upper tail of the skill distribution in period t. Because skills are inherited, the mass of experts above any given threshold at time t is simply a rescaled version of the upper tail of the initial distribution F_0 (see (9)). Consequently, as a_t increases, the economy draws on progressively higher skill levels from that initial distribution.

If F_0 features a fat upper tail, scaling a_t by a constant factor reduces the tail mass only proportionally. Replication thus operates on progressively higher skill levels at an asymptotically constant rate. If instead the upper tail is thin, scaling a_t reduces the tail mass more than proportionally. The cutoff continues to rise, but its proportional increments shrink over time, and asymptotic growth vanishes.

These two forces—replication of high-skill experts and scarcity in the upper tail—are summarized by the asymptotic growth factor $1 + g = (N - I)^\theta$. Experts' span of control, $N - I$, reflects the replication of skills in each period, while θ captures the thickness of the distribution of experts' skills. As $N - I \to 1$ or $\theta \to 0$, the growth factor converges to one, and asymptotic growth vanishes. Crucially, since replication is tied directly to the span of control, automation (I)—by substituting for novices—affects knowledge transmission. This creates a tension between maximizing current output through automation and sustaining long-run growth through the training of future experts, which I examine in Section 5.[15]

In sum, sustaining strictly positive long-run growth ($g > 0$) requires two conditions: expertise must replicate to more than one successor ($N - I > 1$), and the initial skill distribution must feature a sufficiently fat upper tail ($1 - F_0 \in \mathrm{RV}_{-1/\theta}$). Assumption 1 ensures the first condition; without it, the replication force vanishes and asymptotic growth collapses to zero (see Online Appendix I.1 for a formal proof).

To focus on the case of sustained long-run growth, I additionally assume hereafter that $1 - F_0 \in \mathrm{RV}_{-1/\theta}$. This requires $\beta(1 + g) < 1$ to ensure overall welfare remains finite. Thus, for the remainder of the paper, I assume:

Assumption 4 (Sustained Growth and Bounded Welfare)**.** $1 - F_0 \in \mathrm{RV}_{-1/\theta}$*, where* $\theta \in (0, 1)$*. Furthermore,* $\beta < \frac{1}{1+g}$*.*

4.2 Distribution of Experts' Skill

I now turn to the asymptotic distribution of experts' skills. Since the minimum skill level x_t diverges ($x_t = a_{t-1} \to \infty$), the distribution of absolute skills does not converge. I therefore focus on the distribution of normalized skills, defined as the ratio of an expert's skill to the minimum skill of the current cohort, $z \equiv q/x_t$.

[15]The implication that long-run growth is maximized at $I = 0$ rests on two assumptions: (i) all routine tasks are equally valuable for tacit knowledge transmission, and (ii) tasks are non-scalable. Relaxing either assumption implies that automation need not reduce long-run growth, and may even increase it (see Section 6).

Formally, let $\Phi(z)$ denote the asymptotic CDF of z:

$$1 - \Phi(z) \equiv \lim_{t\to\infty} \mathbb{P}_t\left(\frac{q}{x_t} > z\right) = \lim_{t\to\infty} \frac{1 - F_0(zx_t)}{1 - F_0(x_t)}$$

where the second equality follows from (9) (see Appendix C.2 for details). A direct application of regular variation then yields the following result:

Proposition 3 (Asymptotic Distribution of Experts' Skills)**.** *Suppose* $1 - F_0 \in \mathrm{RV}_{-1/\theta}$*. Then, the asymptotic distribution of normalized skills is a Pareto distribution with scale* 1 *and shape* $1/\theta$*:*

$$\Phi(z) = 1 - z^{-1/\theta}, \text{ for } z \geq 1.$$

Consequently, for large t*, the absolute distribution of experts' skills* $F_t(q)$ *is approximately Pareto with a time-varying scale* x_t *and a constant shape parameter* $1/\theta$*.*

Proof. See Appendix C.2 □

Intuitively, inder regular variation, the upper tail of F_0 is asymptotically Pareto. As established in Section 4.1, the equilibrium dynamics repeatedly truncate this initial distribution at progressively higher cutoffs ($a_t \to \infty$). As the truncation point diverges, the economy draws experts exclusively from the region where F_0 exhibits exact Pareto behavior. The normalized skill distribution therefore converges to a Pareto law with scale 1 and shape $1/\theta$.

Proposition 3 further implies that the asymptotic distribution of normalized skills depends solely on the initial tail parameter θ, independent of preference or technology. This follows from the fact that the Pareto distribution is invariant to lower truncation and constant reweighting—precisely the operations governing the evolution of expert skills. Consequently, while changes in technology or preferences alter the long-run growth rate (Proposition 2) and asymptotic price schedules, the long-run relative distribution of talent remains unchanged.

4.3 Long-Run Baseline Utility, and Income and Wage Schedules

Finally, I characterize the long-run normalized value of baseline novice utility, $u \equiv \lim_{t\to\infty} U_t/x_t$, alongside the asymptotic normalized expert income and wage schedules:

$$\pi(z) \equiv \lim_{t\to\infty} \frac{\Pi_t(zx_t)}{x_t}, \qquad \omega(z) \equiv \lim_{t\to\infty} \frac{w_t(zx_t)}{x_t}$$

Before deriving closed-form expressions, it is instructive to describe the intuition behind them. This is particularly true for $\pi(z)$ and $\omega(z)$, as these schedules are dynamically coupled: an expert's current income depends on the wage paid to novices today, which in turn depends on the novices' future income as experts tomorrow.

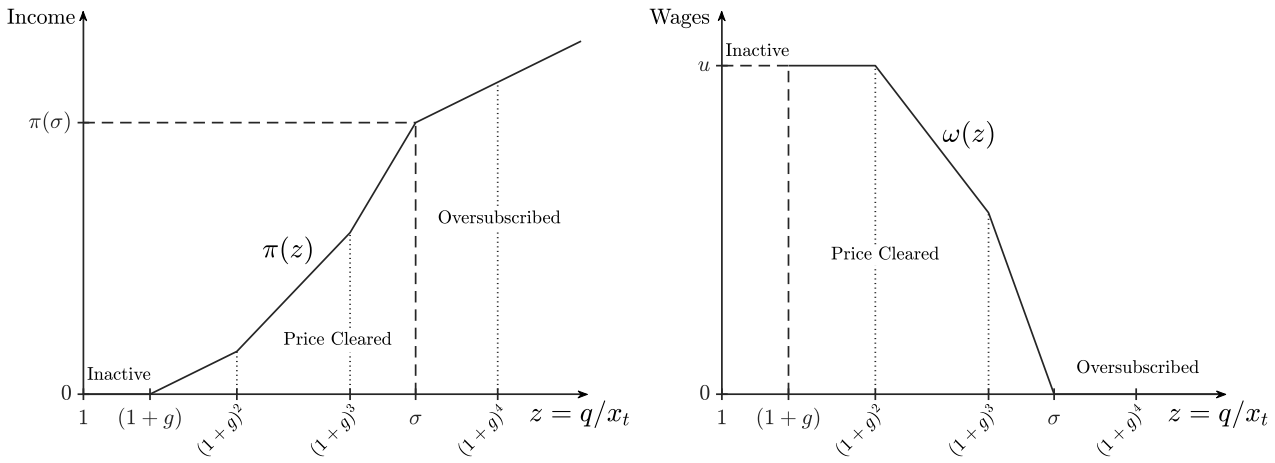
To start, panels (a) and (b) of Figure 2 plot $\pi(z)$ and $\omega(z)$, respectively. As shown, the expert population partitions into three intervals defined by the endogenous cutoffs $1 + g$ and σ:

1. **Inactive:** Experts in $[1, 1+g)$ who do not produce.
2. **Price-Cleared:** Experts in $[1+g, \sigma)$ who pay strictly positive wages.
3. **Oversubscribed:** Experts in $[\sigma, +\infty)$ for whom the zero lower bound binds.

These intervals and the resulting schedules are governed by the asymptotic limit of the marginal expert's normalized skill: $\lim_{t\to\infty}(a_t/x_t) = 1+g \equiv (N-I)^\theta$. This implies the absolute entry cutoff grows asymptotically by the factor $1+g$. Consequently, any fixed absolute skill level eventually becomes obsolete, so all future experts with that skill are inactive and earn zero income. The asymptotic schedules $\pi(z)$ and $\omega(z)$ are then constructed by tracing a skill's economic lifecycle backward from obsolescence, subject to the zero lower bound on wages.

Figure 2 illustrates this process. The horizontal axis in both panels is partitioned into intervals defined by consecutive powers of the growth factor, $[(1+g)^j, (1+g)^{j+1})$ for $j = 0, 1, \ldots$. All experts in $[1, 1+g)$ are inactive. Experts with $z \in [1+g, (1+g)^2)$ are active, but their novices inherit a relative skill $z/(1+g) \in [1, 1+g)$ and become inactive in the subsequent period. Expecting zero future income, these novices must be paid the baseline utility $\omega(z) = u$ upfront, generating the initial upward-sloping segment of $\pi(z)$ and the initial flat segment of $\omega(z)$.

In the next interval, $z \in [(1+g)^2, (1+g)^3)$, novices inherit a relative skill $z/(1+g) \in [1+g, (1+g)^2)$. These novices become active experts in the following period and pay the baseline wage u to their successors (who become inactive experts two periods later). This implies that current novices have a strictly positive continuation value that is strictly increasing in z, which current experts extract by reducing the upfront wage. Consequently, $\omega(z)$ begins to fall strictly below u and decreases in z,



(a) Asymptotic Income Schedule

(b) Asymptotic Wage Schedule

Figure 2: The Construction of the Asymptotic Income and Wage Schedules

Notes. Parameter Values: $N = 2.7$, $I = 1$, $h = 1$, $m = 2$, $\beta = 0.96^{10}$, and $\theta = 0.5$. Panel (a) shows the asymptotic normalized income schedule $\pi(z)$, and panel (b) shows the asymptotic normalized wage schedule $\omega(z)$. The vertical dotted lines denote consecutive powers of the growth factor $(1+g)^j$, marking the skill thresholds where the remaining active lifespans of expert dynasties increase by one period before obsolescence. These intervals create the piecewise kinks in both schedules throughout the price-cleared region $z \in [1+g, \sigma)$.

generating the first downward sloped region of $\omega(z)$.

As z crosses each successive $(1+g)^j$ threshold, the transmitted skill survives an additional period before obsolescence. This extended lifespan increases the continuation value of transmitted knowledge, which current experts extract by further reducing the upfront wage. Iterating this logic generates a sequence of kinks in both $\pi(z)$ and $\omega(z)$ throughout the price-cleared region, $z \in [1+g, \sigma)$. This process continues until $z = \sigma$, where the value of the transmitted knowledge exactly offsets the asymptotic novice utility u. At this threshold, the zero lower bound binds, and $\omega(z)$ remains zero thereafter.

The oversubscribed region ($z \geq \sigma$) rationalizes the severe rationing of entry-level positions observed at elite professional services firms. For instance, McKinsey and Boston Consulting Group accept only 0.6% and 1% of applicants, respectively (Vlamis et al., 2025). Similarly, fewer than 1% of applicants are admitted to the internship programs at JPMorgan and Goldman Sachs, which serve as the primary pipelines for full-time entry-level roles (Clarke, 2025).

While mechanisms outside the formal model—such as the screening of heterogeneous applicants—undoubtedly contribute to these low acceptance rates, the framework establishes that they are a natural consequence of the incomplete contracts governing the transmission of tacit knowledge. Because novices can compensate experts solely with their labor, the zero lower bound on wages caps the implicit price of elite mentorship: experts with $z \geq \sigma$ cannot fully extract the value of the tacit knowledge they transmit, so novices strictly prefer these matches, resulting in latent "excess demand."[16]

I now formalize this intuition and derive closed-form expressions for $\pi(z)$, $\omega(z)$, and u.

Preliminaries. Recall from Section 3.2 that an active expert with absolute skill q generates a surplus $S(q) \equiv qm^I h^{N-I} - \rho I$ before novice compensation. Because the minimum skill level diverges ($x_t = a_{t-1} \to \infty$), the rental cost becomes asymptotically negligible. Thus:

$$\lim_{t\to\infty} \frac{S(zx_t)}{x_t} = zm^I h^{N-I}$$

Long-Run Novice Utility. Novices matched with price-cleared experts attain a lifetime utility $U_t = S(a_t)/(N-I)$. The normalized baseline utility therefore converges to:

$$u \equiv \lim_{t\to\infty} \frac{U_t}{x_t} = \frac{m^I h^{N-I}}{N-I}(1+g)$$

Asymptotic Income and Wage Schedules. Because the normalized skill of the marginal expert converges to $\lim_{t\to\infty}(a_t/x_t) = 1+g$, experts with relative skill $z \in [1, 1+g)$ are inactive. These experts earn $\pi(z) = 0$ and face a shadow wage $\omega(z) = u$, which renders production unprofitable.

To characterize $\pi(z)$ for $z \geq 1+g$, substitute the wage policy $w_t(q) = \max\{0, U_t - \beta\Pi_{t+1}(q)\}$ into the active expert's income $\Pi_t(q) = S(q) - w_t(q)(N-I)$ to obtain the recursion:

$$\Pi_t(q) = \min\left\{S(q), S(q) - (N-I)U_t + \beta(N-I)\Pi_{t+1}(q)\right\}$$

[16] As discussed in Section 3.2, oversubscription is an informal interpretation, as the model characterizes only realized matches. Furthermore, as noted in Section 3.1, a wage of zero captures both literal unpaid labor and negligible effective compensation driven by demanding hours relative to nominal pay.

Normalizing by x_t and taking the limit as $t \to \infty$ yields (see Appendix C.3 for details):

$$\pi(z) = m^I h^{N-I} \min\left\{z, z - (1+g) + \frac{\beta(N-I)(1+g)}{m^I h^{N-I}} \pi\left(\frac{z}{1+g}\right)\right\}, \quad \text{for } z \geq 1+g$$

Zero income for the marginal active expert implies continuous pasting to the inactive region, establishing the boundary condition $\pi(z) = 0$ for all $z \in [1, 1+g]$.

Appendix C.3 solves this functional equation, yielding the closed-form asymptotic income schedule:

$$\pi(z) = m^I h^{N-I} \begin{cases} 0 & \text{if } z \in [1, 1+g) \\ \gamma(z) & \text{if } z \in [1+g, \sigma) \\ z & \text{if } z \in [\sigma, +\infty) \end{cases}$$

where:

$$\gamma(z) = \sum_{j=0}^{n(z)-1} [\beta(N-I)]^j \left(z - (1+g)^{j+1}\right) \quad \text{with} \quad n(z) \equiv \left\lfloor \frac{\ln z}{\ln(1+g)} \right\rfloor$$

and $\sigma \in (1+g, +\infty)$ is the unique solution to $\sigma = \gamma(\sigma)$.

The term $m^I h^{N-I} \gamma(z)$ represents the *value* of a dynasty of experts with skill z prior to obsolescence, defined as the present discounted sum of its excess output relative to the marginal active expert in each period. The index $n(z)$ denotes the number of periods until z's obsolescence, so $n(z) - 1$ is the maximum integer j for which the excess output condition $z \geq (1+g)^{j+1}$ holds.

Absent the zero lower bound on wages, an expert with relative skill z could always appropriate the full value of the dynasty by charging her novices for the skill provided, a transfer that recursively incorporates the value these novices will subsequently charge as experts to their own novices. However, because wages cannot be negative, an expert's current income cannot exceed the physical output generated within the period, $z m^I h^{N-I}$. Oversubscription thus emerges exactly at the skill threshold σ where the value of the dynasty meets this limit of current production, $m^I h^{N-I} \gamma(\sigma) = m^I h^{N-I} \sigma$, yielding the condition $\sigma = \gamma(\sigma)$.

Finally, to characterize $\omega(z)$ for $z \geq 1+g$, normalize the active expert's income $\Pi_t(q) = S(q) - w_t(q)(N-I)$ by x_t. Taking the limit as $t \to \infty$ and rearranging yields:

$$\omega(z) = \frac{z m^I h^{N-I} - \pi(z)}{N-I} = u - \beta(1+g)\pi\left(\frac{z}{1+g}\right), \quad \text{for } z \geq 1+g$$

where the last equality follows from the functional equation of $\pi(z)$. Evaluating $\omega(z)$ using the piecewise solution for π, and imposing the shadow wage u for inactive experts ($z \in [1, 1+g)$), yields the complete asymptotic wage schedule:

$$\omega(z) = \frac{m^I h^{N-I}}{N-I} \begin{cases} 1+g & \text{if } z \in [1, 1+g) \\ z - \gamma(z) & \text{if } z \in [1+g, \sigma) \\ 0 & \text{if } z \in [\sigma, +\infty) \end{cases}$$

5 The Impact of Different Technological Advances

This section examines the effects of technological advances on aggregate output and welfare. I focus on two different forms of technical change: improvements in entry-level automation and the creation of new tasks. Section 6 extends the analysis to labor- and capital-augmenting technologies.

For ease of exposition, I assume the absolute skill distribution at the time of the shock τ is exactly Pareto with scale x_τ and shape $1/\theta$. This assumption admits two possible interpretations:

1. **Exact Initial Distribution**: If the initial distribution F_0 is exactly Pareto with scale x_0 and shape $1/\theta$, equation (5) ensures that the skill distribution at date τ is also exactly Pareto with scale $x_\tau = x_0(N-I)^{\theta\tau}$ and shape $1/\theta$.

2. **Asymptotic Approximation**: If F_0 is not exactly Pareto, the setup approximates an economy that experiences the shock at an arbitrarily distant date, by which time the skill distribution has converged to its stable Pareto form (Proposition 3).

5.1 Improvements in Entry-Level Automation

To establish a baseline, consider the evolution of the economy from period τ onward in the absence of a shock. The distribution of expert skill then remains exactly Pareto, with its scale parameter evolving geometrically as $x_t = x_\tau(N-I)^{\theta(t-\tau)}$ for all $t \geq \tau$. Aggregate output along this trajectory is therefore:

$$Y_t = a_t \left[\frac{m^I h^{N-I}}{(N-I)(1-\theta)}\right] = x_\tau(N-I)^{\theta(t-\tau)} \left[\frac{m^I h^{N-I}}{(1-\theta)(N-I)^{1-\theta}}\right], \text{ for } t \geq \tau$$

where the second equality follows from $a_t/x_t = a_t/a_{t-1} = (N-I)^\theta$.

Suppose an unanticipated technological shock at date τ automates an additional measure Δ of tasks, where $N - I - \Delta > 1$. This restriction ensures that the long-run consequences of automation derived below do not rely on mass unemployment or stagnation, but emerge in an economy that sustains market-clearing and positive long-run growth (see Online Appendix I.1 for details).[17] Additionally, I assume that machine productivity in these newly automated tasks equals that in previously automated tasks, m. This assumption is not crucial for the ensuing analysis.

The shock induces active experts to immediately reallocate tasks, assigning $I + \Delta$ to machines and $N - I - \Delta$ to novices from date τ onward. Letting tildes denote the shocked trajectory, the absolute skill distribution at date τ remains predetermined, $\tilde{F}_\tau(q) = F_\tau(q)$. Given this initial condition and the updated span of control, the subsequent equilibrium allocations $\{\tilde{a}_t = \tilde{x}_{t+1}, \tilde{F}_t\}_{t \geq \tau}$ solve the same dynamical system of Section 4.1 (equations (3) and (5)). This yields:

$$\tilde{x}_t = x_\tau(N-I-\Delta)^{\theta(t-\tau)} \text{ and } 1 - \tilde{F}_t(q) = \left(\frac{q}{\tilde{x}_t}\right)^{-1/\theta}, \text{ for } t \geq \tau$$

[17] Moreover, in Section 6.3, I perform back-of-the-envelope calculations using U.S. data, which indicate that driving the span of control below unity would require a shock reducing the expert span by 24% to 49% depending on the specification—an exceptionally severe disruption.

Consequently, the minimum expert skill continues to grow geometrically, albeit by a permanently lower factor $1+\tilde{g} = (N-I-\Delta)^{\theta} < 1+g$. Moreover, the absolute skill distribution at time t remains exactly Pareto with shape $1/\theta$, but with new scale $\tilde{x}_t$. From (7), aggregate output along the new equilibrium trajectory is:

$$\tilde{Y}_t = x_\tau (N-I-\Delta)^{\theta(t-\tau)} \left[\frac{m^{I+\Delta} h^{N-I-\Delta}}{(1-\theta)(N-I-\Delta)^{1-\theta}} \right], \text{ for } t \geq \tau$$

The ratio of $\tilde{Y}_t$ to Y_t then summarizes both the short- and long-run implications of the shock:

$$\frac{\tilde{Y}_t}{Y_t} = \underbrace{\overbrace{\left(\frac{m}{h}\right)^{\Delta}}^{\text{Intensive margin gains}} \quad \overbrace{\left(1-\frac{\Delta}{N-I}\right)^{-(1-\theta)}}^{\text{Extensive margin gains}}}_{\text{Immediate adoption gains } (>1)} \underbrace{\left(1-\frac{\Delta}{N-I}\right)^{\theta(t-\tau)}}_{\text{Long-run losses } (\leq 1)}, \quad \text{for } t \geq \tau. \tag{11}$$

This equation immediately leads to the following result:

Proposition 4 (Output Effects of an Automation Shock)**.** *Suppose the economy experiences an unanticipated automation shock at date τ. Then, relative to the no-shock baseline:*

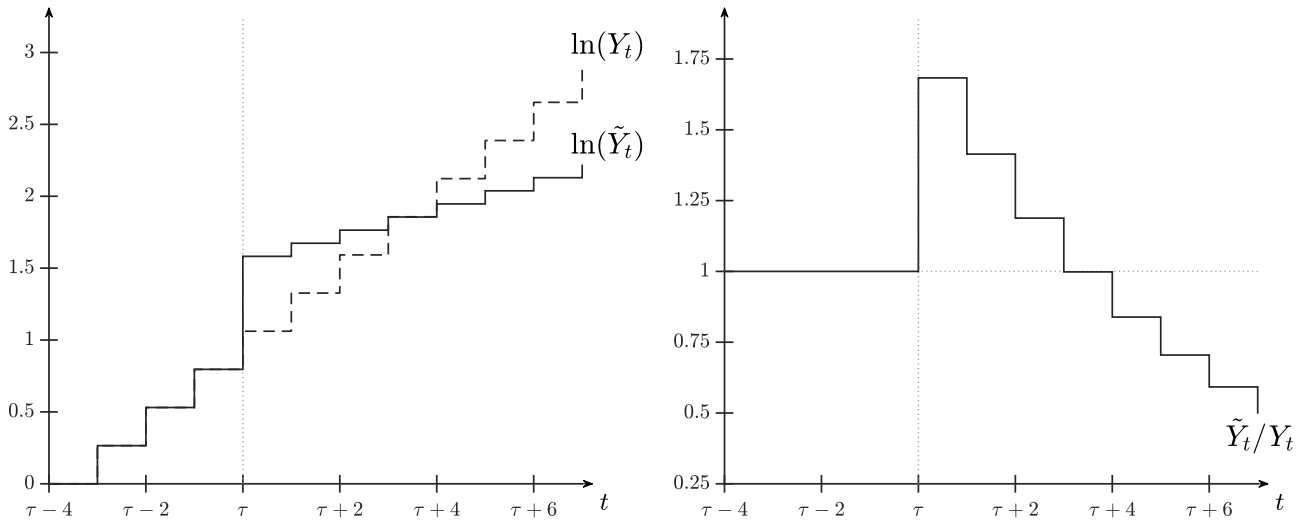
1. ***Immediate Adoption Gains**: Output is strictly higher on impact ($\tilde{Y}_\tau > Y_\tau$).*
2. ***Permanent Growth Slowdown**: The growth rate of output is strictly lower for all $t \geq \tau$ ($\tilde{g} < g$).*
3. ***Long-Run Reversal**: There exists a date $t^* > \tau$ such that output is strictly lower for all $t > t^*$ ($\tilde{Y}_t < Y_t$).*

Proof. The result follows directly from (11). □

Proposition 4 is illustrated in Figure 3. Panel (a) plots the log-level trajectories of aggregate output (normalized so that $Y_{\tau-4} = 1$). Panel (b) plots their ratio, $\tilde{Y}_t/Y_t$. On impact, the shock generates an upward shift in the level of output. However, as the flatter slope in Panel (a) shows, this initial expansion comes at the expense of a permanently lower growth rate. Consequently, the output ratio in Panel (b) initially jumps upwards but then persistently decays over time, eventually falling strictly below one.

Equation (11) isolates these opposing forces. The first two terms—both strictly greater than one and constant for all $t \geq \tau$—capture the immediate adoption gains along two distinct margins. The first term, $(m/h)^{\Delta}$, reflects the intensive margin: because machines outperform novices in newly automated tasks ($m > h$), the output of all existing expert-novice teams increases on impact. The second term, $(1-\Delta/(N-I))^{-(1-\theta)}$, reflects the extensive margin: as highly skilled experts demand fewer novices, displaced labor reallocates downward to clear the market. This allows a new tier of previously inactive, lower-skill experts to profitably produce, further expanding aggregate output.

In contrast, the third term, $(1-\Delta/(N-I))^{\theta(t-\tau)}$, reflects the long-run losses. At $t = \tau$, this term equals one because the aggregate stock of tacit knowledge is predetermined. For $t > \tau$, however, it falls strictly below one and monotonically decreases. This decay is the dynamic counterpart to the equilibrium reallocation of labor: as highly skilled experts demand fewer novices, displaced novices



(a) Log Output Trajectories

(b) Relative Output

Figure 3: The Short- and Long-Run Effects of an Automation Improvement on Aggregate Output

Notes. Parameter values: $N = 2.7$, $I = 1$, $h = 1$, $m = 2$, $\beta = 0.96^{10}$, $\theta = 0.5$, and $\Delta = 0.5$. Panel (a) plots the trajectories of aggregate log output without (Y_t) and with ($\tilde{Y}_t$) the automation shock. Both series are normalized so that the output four periods prior to the shock equals 1 ($Y_{\tau-4} = 1$). Panel (b) plots the ratio of these trajectories, $\tilde{Y}_t/Y_t$. The vertical dotted line denotes the date of the automation shock, $t = \tau$.

are forced to learn from less-skilled experts. These novices acquire less tacit knowledge and become worse mentors themselves. The best practices of top experts thus diffuse more slowly, imposing a compounding drag on future output.

A common intuition is that automation harms skill accumulation by depriving novices of hands-on practice. While this mechanism is plausible, Proposition 4 establishes that it is not strictly necessary: a compositional degradation of mentors alone suffices to impair growth. Anecdotal reports are consistent with this compositional shift. Amid the recent cooling of the entry-level white-collar labor market—which, as discussed below, some attribute to nascent AI adoption—graduates are increasingly sorting into smaller, less productive firms that previously struggled to attract such talent (Ellis, 2026; Lahart, 2026).

Proposition 4 highlights an intertemporal trade-off: automation increases immediate aggregate output but imposes a permanent growth slowdown. To evaluate the welfare implications of this trade-off, I compute the shock's net effect using the definition from Section 3.2. Using (8), overall welfare at date τ with and without the shock is:

$$W_\tau = \frac{Y_\tau}{1-\beta(1+g)} - \frac{\rho I}{(N-I)(1-\beta)}$$
$$\tilde{W}_\tau = \frac{\tilde{Y}_\tau}{1-\beta(1+\tilde{g})} - \frac{\rho(I+\Delta)}{(N-I-\Delta)(1-\beta)}$$

I then have the following result:

Proposition 5 (Welfare Effects of an Automation Shock)**.** *There exists a unique threshold* $\beta^* \in (0, \frac{1}{1+g})$

such that the automation shock is strictly welfare-decreasing ($\tilde{W}_\tau < W_\tau$) if and only if the discount factor satisfies $\beta > \beta^*$.

Proof. See Appendix D.1 □

Proposition 5 establishes that an improvement in entry-level automation can reduce aggregate welfare.[18] This inefficiency stems directly from the zero lower bound on wages. Constrained by this friction, oversubscribed experts fail to internalize the full value of their mentorship. The result is socially excessive automation that forces displaced novices to learn from less-skilled mentors.

This market failure does not arise simply because machines outperform novices in routine tasks. In a frictionless economy, novices would borrow against future earnings to outbid machines and preserve socially efficient training—especially since this same improvement in automation makes their future expertise even more valuable. Instead, the zero lower bound forces novices to finance elite mentorship solely through their labor—the very currency that automation devalues.

The mechanism underlying Propositions 4 and 5 operates through a displacement margin: because production is non-scalable, automation reduces knowledge transmission by shrinking the number of novices attached to each expert-led project. Section 6.1 shows that this is not the only margin through which automation can operate: if it also frees expert time and expands experts' scale of operation, an offsetting scale effect arises that can attenuate—and, in principle, reverse—the propositions' conclusions.

Even so, at least in these early stages of adoption, the preliminary evidence on generative AI points more clearly to displacement than to offsetting scale effects. It documents a relative decline in entry-level hiring in AI-exposed, knowledge-intensive sectors—such as Information, Professional Services, and Finance—where the transmission of tacit skills is paramount. Whether this decline reflects realized AI displacement, anticipatory hiring reductions ahead of AI adoption, or confounding sectoral or macroeconomic trends remains debated, leaving the permanence of this shift uncertain. Yet, the immediate observable consequence is a relative narrowing of the novice pipeline in AI-exposed sectors.[19]

For instance, Hosseini and Lichtinger (2025) find that GenAI-adopting firms reduce junior employment by 9% relative to non-adopters, while senior employment remains unaffected. Likewise, Brynjolfsson et al. (2025a) document a 16% relative decline in early-career employment across highly AI-exposed occupations, with no corresponding impact on senior positions. Both estimates are robust to controlling for several confounding macro-sectoral variables and trends. These results are further corroborated by sharp drops in entry-level white-collar vacancies (Berger et al., 2024), declin-

[18] Conditional on this inefficiency, the role of the discount factor is mechanical: a patient society ($\beta > \beta^*$) places greater weight on the compounding growth losses of future cohorts relative to the immediate output gains of current ones.

[19] While the model cannot account for transitory novice displacement—such as that driven by the anticipation of AI—the framework readily captures the macroeconomic implications of such a shock. Mechanically, a temporary reduction in the span of control with no associated improvement in productivity generates a permanent negative level effect on output, the magnitude of which depends on the shock's depth and duration. This stands in contrast to the permanent negative growth effect associated with structural automation.

ing employment shares for young workers in AI-exposed sectors (Atkinson and Yamco, 2026), and similar displacement effects in the United Kingdom (Klein Teeselink, 2025).[20]

Thus, preliminary empirical evidence suggests early iterations of generative AI may constitute a form of "seniority-biased technological change" (Hosseini and Lichtinger, 2025): it substitutes for entry-level execution while complementing expert judgment.

5.2 Task-Creating Technologies

While preliminary evidence points to a relative decline in entry-level hiring within AI-exposed occupations, the historical record demonstrates that technological advances also create new labor-intensive tasks (Autor et al., 2024).[21] Early examples involving generative AI include data annotation, prompt engineering, and forward-deployed software integration. Motivated by this precedent, I now explore the macroeconomic effects of a task-creating technology.

Formally, consider an unanticipated technological shock at date τ that introduces a new production technology alongside the baseline one. Under this new technology, the set of automatable tasks is $[0, I - \Delta]$ for a given $\Delta \in (0, I)$, where machines and novices contribute factors m and h per task, respectively. The remaining tasks in $(I - \Delta, N]$ are strictly non-automatable: novices contribute the baseline factor h per task on standard tasks in $(I - \Delta, N - \Delta]$, and a strictly higher factor $H > m > h$ on newly introduced tasks in $(N - \Delta, N]$.

This formulation, inspired by Acemoglu and Restrepo (2018) and Autor et al. (2024), captures the introduction of a task-creating technology. The new technology leads to the creation of an additional measure Δ of labor-intensive tasks that require human execution. As in those papers, I assume that the creation of these new tasks simultaneously renders obsolete an equal measure of previously automatable tasks to keep the total measure of tasks in a production process fixed at N.

Because the baseline technology remains available, the expert's optimization problem now entails three choices: (i) whether to be active; (ii) which technology to adopt; and (iii) the optimal allocation of tasks between machines and novices. I then impose the following condition:

Assumption 5 (Full Adoption of New Technology)**.** *The productivity of novices on new tasks satisfies:*

$$\left(\frac{H}{m}\right)^{\Delta} \geq 1 + \frac{\Delta}{N - I}$$

[20] Strictly speaking, the model cannot account for the occupation-specific AI effects documented in this empirical literature, as it features a single sector. As such, the framework should be interpreted either as capturing economy-wide averages or as a partial-equilibrium approximation of an exposed industry. Furthermore, while the evidence documents relative declines in entry-level employment, the model holds aggregate novice employment constant. This is entirely consistent with an economy-wide interpretation, since aggregate labor supply is relatively inelastic. Alternatively, under a single-industry interpretation, such declines likely reflect inter-sectoral mobility and labor market frictions—forces omitted here for tractability.

[21] For instance, Autor et al. (2024) estimate that 60% of U.S. employment in 2018 consisted of job titles introduced since 1940.

As shown in Appendix D.2, this condition guarantees that all active experts adopt the new technology from τ onwards; if it does not hold, the least-skilled active experts will either reject the new technology immediately at τ or eventually abandon it as absolute skills grow. Assumptions 2 and 5 together therefore imply that an active expert with skill q produces:

$$qm^{I-\Delta}h^{N-I}H^{\Delta}$$

Because all active experts adopt the new technology, the equilibrium dynamics mirror the baseline model, except that experts hire $N - I + \Delta$ novices and output is scaled by the productivity premium of new tasks, $(H/m)^{\Delta}$. Consequently, aggregate output and welfare along the new trajectory are:

$$\tilde{Y}_t = x_\tau (N - I + \Delta)^{\theta(t-\tau)} \left[\frac{m^{I-\Delta} h^{N-I} H^{\Delta}}{(1-\theta)(N - I + \Delta)^{1-\theta}} \right], \text{ for } t \geq \tau$$

$$\tilde{W}_\tau = \frac{\tilde{Y}_\tau}{1 - \beta(1+\tilde{g})} - \frac{\rho(I-\Delta)}{(N - I + \Delta)(1-\beta)}$$

where $1 + \tilde{g} \equiv (N - I + \Delta)^{\theta}$ is the new growth factor of output. The ratio of aggregate output with and without the shock is therefore:

$$\frac{\tilde{Y}_t}{Y_t} = \underbrace{\overbrace{\left(\frac{H}{m}\right)^{\Delta}}^{\text{Intensive margin gains}} \overbrace{\left(1 + \frac{\Delta}{N-I}\right)^{-(1-\theta)}}^{\text{Extensive margin losses}}}_{\text{Immediate adoption gains } (>1)} \underbrace{\left(1 + \frac{\Delta}{N-I}\right)^{\theta(t-\tau)}}_{\text{Long-run gains } (\geq 1)}$$

I then have the following result:

Proposition 6 (Output and Welfare Effects of a Task-Creating Technology)**.** *Suppose the economy experiences an unanticipated shock at date τ that introduces a task-creating technology. Then, relative to the no-shock baseline:*

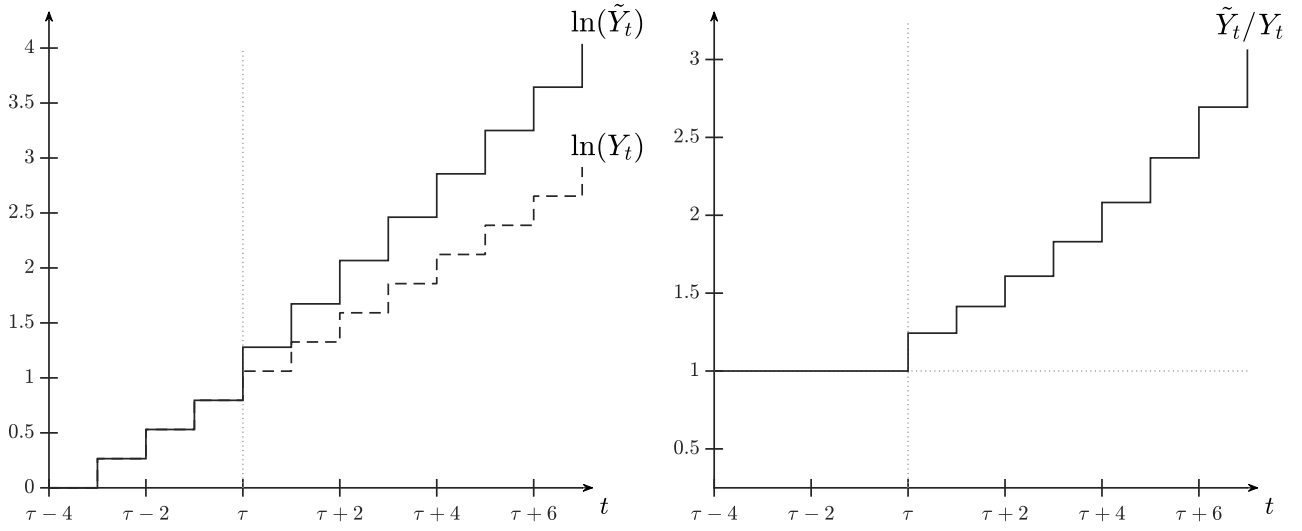
1. ***Immediate Adoption Gains**: Output is strictly higher on impact ($\tilde{Y}_\tau > Y_\tau$).*
2. ***Growth Acceleration**: The growth rate of output is strictly higher for all $t \geq \tau$ ($\tilde{g} > g$).*
3. ***Welfare Gains**: The shock is strictly welfare-increasing ($\tilde{W}_\tau > W_\tau$).*

Proof. Growth acceleration and welfare gains are immediate from comparing $\tilde{g}$ and $\tilde{W}_\tau$ with g and W_τ from Section 5.1. Immediate adoption gains, in turn, follows because:

$$\left(\frac{H}{m}\right)^{\Delta} \left(1 + \frac{\Delta}{N-I}\right)^{-(1-\theta)} \geq \left(1 + \frac{\Delta}{N-I}\right)^{\theta} > 1$$

where the first inequality holds by Assumption 5. □

Figure 4 illustrates the output effects of a task-creating technology. Unlike improvements in automation, this technology generates both immediate and long-run gains. By inducing the highest-skilled experts to hire additional novices, it widens their span of control and permanently raises the long-run growth rate. Although this expanded span of control raises the entry skill cutoff—thereby



(a) Log Output Trajectories

(b) Relative Output

Figure 4: The Short- and Long-Run Effects of Task-Creating Technology

Notes. Parameter values: $N = 2.7$, $I = 1$, $h = 1$, $m = 2$, $\beta = 0.96^{10}$, $\theta = 0.5$, $\Delta = 0.5$, and $H = 4$. Panel (a) plots the trajectories of aggregate log output without (Y_t) and with ($\tilde{Y}_t$) the introduction of the task-creating technology. Both series are normalized so that the output four periods prior to the shock equals 1 ($Y_{\tau-4} = 1$). Panel (b) plots the ratio of these trajectories, $\tilde{Y}_t/Y_t$. The vertical dotted line denotes the date of the task-creating shock, $t = \tau$.

generating an extensive margin loss on impact—the optimality of adoption (Assumption 5) implies that the intensive margin gains strictly dominate. Thus, aggregate output unambiguously increases on impact.

In practice, technological advances frequently exhibit both automating and task-creating dimensions. The electronic spreadsheet, for instance, automated the manual arithmetic previously performed by clerical staff. Simultaneously, by lowering the cost of calculation, it generated new, labor-intensive tasks—such as standardizing raw data and auditing complex formulas—that expanded the routine execution duties of junior analysts.[22]

These competing forces can be formalized by combining the shocks studied above. When a technology simultaneously automates a subset of existing tasks and introduces a measure of new ones, aggregate output always rises on impact. The effect on the growth rate, however, depends on the net change in the span of control of experts active before the shock. If the span widens, the intergenerational transmission of expertise strengthens, accelerating long-run growth. If it contracts, transmission weakens, producing a permanent growth slowdown.

[22] As documented by Autor et al. (2024), however, recent technological advances have proven more task-automating than task-creating. Using textual analysis to distinguish patents that generate new occupational specialties (augmentation) from those that replace existing worker tasks (automation), they find that the labor-reinstating force of augmentation has attenuated since 1980, while the employment-eroding force of automation has intensified.

6 Discussion and Extensions

This section examines the implications of key assumptions and extends the baseline model in several directions. The Online Appendix provides further extensions and robustness checks.

6.1 Scale Effects and Other Technological Shocks

Scale Effects.— The baseline model assumes that production is non-scalable. Due to limited time and attention, each expert oversees only the fixed measure N of tasks required for a single project. To relax this assumption, let the expert choose a continuous scale of operation, μ, representing the measure of concurrent projects she oversees. The total output of an expert with skill q is therefore:

$$\mu \cdot q m^I h^{N-I}$$

The scale μ is constrained by the expert's unit of time and the total time $\mathcal{M}(I, m, h)$ required to manage a single project, yielding $\mu = 1/\mathcal{M}(I, m, h)$. This project-level time aggregates the supervision costs of its constituent tasks via a constant-returns-to-scale CES function, where $\eta > 0$ governs the complexity of supervision. The time required to oversee an individual task is inversely proportional to the productivity of the executing agent: $1/m$ for a machine and $1/h$ for a novice. Weighting these per-task time costs by their task shares and scaling by project size N, total project time is:

$$\mathcal{M}(I, m, h) = N \left[\frac{I}{N} \left(\frac{1}{m} \right)^{\eta} + \frac{N-I}{N} \left(\frac{1}{h} \right)^{\eta} \right]^{\frac{1}{\eta}}$$

Because an expert manages μ projects and hires $N - I$ novices per project, her total span of control is $\ell = \mu(N - I) = (N - I)/\mathcal{M}(I, m, h)$. Since ℓ is constant over time, equilibrium dynamics mirror the baseline model, except that the long-run growth factor of aggregate output becomes $1 + g = \ell^{\theta}$.

This formulation departs from standard task-based frameworks, which typically feature the continuous scaling of inputs within individual tasks (see Appendix B.1). Such a departure serves three purposes. First, it more faithfully captures the expert's role as an orchestrator who integrates tasks, rather than a standard factor of production. Second, it explicitly incorporates the expert's time constraint, formalizing the premise that time acts as the fundamental bottleneck to scaling embodied expertise. Finally, it represents a minimal deviation from the baseline, allowing for the analysis of scale effects without requiring a full recharacterization of the equilibrium.

Consider then the automation shock studied in Section 5.1. Although output continues to increase on impact, the effect on the span of control—and therefore on long-run growth—is now ambiguous. Differentiating $\ln \ell$ with respect to I isolates the competing forces:

$$\frac{\partial \ln \ell}{\partial I} = \underbrace{-\frac{1}{N-I}}_{\text{displacement effect}} \quad \underbrace{-\frac{\partial \ln \mathcal{M}(I, m, h)}{\partial I}}_{\text{scale effect}}$$

The first term is the displacement effect analyzed in Section 5.1: allocating tasks to machines reduces the measure of novices trained per project. The second term captures a new scale effect. Because

machines are more productive and require less expert supervision ($m > h$), automation lowers total project management time ($\partial \mathcal{M}/\partial I < 0$), allowing experts to oversee more concurrent projects. The net impact on long-run growth depends on which of these opposing forces dominates.

When the displacement effect dominates, introducing scale effects does not alter the qualitative results of Section 5.1. This occurs in two different cases. First, displacement prevails when automation is "so-so" (Acemoglu and Restrepo, 2019): machines are adopted but are only marginally more productive than novices ($m \approx h$). Here, replacing a novice with a machine yields minimal time savings. As $m \to h$, the scale effect vanishes ($\partial \mathcal{M}/\partial I \to 0$), and displacement entirely drives the impact on the span of control.

Second, the displacement effect dominates when task supervision is highly interdependent (high η). As $\eta \to \infty$, the CES time aggregator converges to a Leontief technology. In this limit, tasks are supervised in parallel, so total project time is dictated entirely by the slowest inputs—the human novices. Consequently, even if machines are substantially more productive than novices ($m \gg h$), the expert cannot reallocate the time saved by the automation improvement. The scale effect is therefore negligible, and the automation shock unambiguously reduces the expert's span of control.

By contrast, when the scale effect dominates, improvements in entry-level automation accelerate long-run growth. In this case, the expert's expanded scale of operation outweighs the displacement of novices per project. Top experts hire a larger aggregate measure of novices, accelerating the diffusion of best practices. The macroeconomic effects of the shock thus mirror those of a task-creating technology, even without introducing new labor-intensive tasks. As the evidence reviewed in Section 5.1 suggests, however, the empirically salient margin so far appears to be displacement rather than expert scaling.

Labor- and Capital-Augmenting Technologies.— Beyond automating and creating tasks, AI tools can increase the productivity of novices and machines at existing tasks. Consistent with labor augmentation, Brynjolfsson et al. (2025b) document that AI assistants lead to a 15% increase in the resolution rate of customer-support agents by providing real-time response suggestions. For capital augmentation, de Souza (2025) reports that AI-powered maintenance reduces unplanned machine downtime by 50% through analysis of real-time sensor data.

Formally, labor- and capital-augmenting technologies correspond to exogenous increases in h and m, respectively. In the baseline model without scale effects, these improvements generate a one-time level effect on aggregate output but leave the experts' span of control—and thus long-run growth—unchanged. Scale effects alter this result. By reducing task execution time, both technologies lower total supervision time per project. Experts therefore oversee more concurrent projects, accelerating knowledge transmission and long-run growth.[23]

[23]The growth effects of these shocks are asymmetric when task supervision is highly interdependent ($\eta \to \infty$). Under parallel supervision, human execution speed bounds total project time because machines are already faster than novices ($m > h$). Capital-augmenting improvements therefore yield no marginal time savings. Only labor-augmenting technologies relax this constraint and accelerate long-run growth.

Span of Control as a Sufficient Statistic.— As the preceding discussion shows, new technologies set multiple competing forces in motion: they can simultaneously automate tasks, create new labor-intensive tasks, and augment both labor and capital in existing ones. Furthermore, even pure automation can accelerate the diffusion of tacit knowledge if the resulting time savings enable experts to scale their operations. Consequently, the net effect of a new technology on knowledge transmission is theoretically ambiguous.

The framework cuts through this complexity by establishing that the span of control of the most skilled experts is a sufficient statistic for the rate of knowledge transmission—and thus, for asymptotic growth. Consequently, the effect of a technological shock on knowledge transmission depends entirely on whether the mass of novices mentored by top experts expands or contracts.

Admittedly, this sufficient statistic result relies on the model's stark simplifying assumptions—most notably, perfect knowledge transmission between experts and novices, and a one-to-one mapping between routine tasks and span of control. While relaxing these assumptions introduces additional margins of adjustment (as discussed throughout this section), the core economic logic holds. The span of control of the most skilled experts remains an informative and practical proxy for a technology's net effect on the intergenerational transmission of tacit knowledge.

6.2 A Quantitative Illustration

Intertemporal Tradeoff and the Break-Even Change in the Span of Control.— The preliminary evidence in Section 5.1 suggests generative AI is displacing novices in AI-exposed sectors. According to Proposition 4, this displacement yields immediate output gains but compromises future output by disrupting the transmission of tacit knowledge.

A full structural estimation of this effect is beyond the scope of this paper. It is nevertheless instructive to ask whether this intertemporal tradeoff operates at a quantitatively relevant scale. To this end, I carry out a simple back-of-the-envelope calculation. Taking recent macroeconomic projections of the one-off output gains from AI adoption as given, I compute the reduction in the experts' net span of control required to completely dissipate these gains after exactly one generation.

Importantly, this exercise does not take a stance on whether generative AI will ultimately accelerate or weaken knowledge transmission. As discussed in Section 6.1, the technology can set multiple competing forces in motion. Rather, the calculation is conditional: assuming AI creates displacement, how severe must the contraction in effective mentorship be to completely offset the projected adoption gains in just one generation, and is this required reduction empirically plausible?

As in Section 5, suppose the adoption of a new technology at date τ generates an immediate aggregate output gain $\tilde{Y}_\tau / Y_\tau > 1$. Let ℓ denote the pre-shock aggregate span of control (corresponding to $N - I$ in the baseline model or the scale-adjusted span in Section 6.1), and let $\tilde{\ell}$ denote the post-shock net span of control. In either specification, the economy breaks even exactly one generation after the

shock when:

$$Y_\tau \ell^\theta = \tilde{Y}_\tau (\tilde{\ell})^\theta \iff 1 - \frac{\tilde{\ell}}{\ell} = 1 - \left(\frac{\tilde{Y}_\tau}{Y_\tau}\right)^{-\frac{1}{\theta}} \tag{12}$$

Calculating this break-even reduction requires only the tail parameter θ and the immediate output gain $\tilde{Y}_\tau / Y_\tau$. Following Lucas (2009), Lucas and Moll (2014), and De la Croix et al. (2018), I set $\theta = 0.5$, a value consistent with the observed dispersion of earnings in U.S. data. For the immediate output gain, I consider two recent projections:

1. **Acemoglu (2024)** projects a one-off TFP increase of 0.71% realized over ten years, assuming AI affects 5% of all tasks. Adjusting for capital deepening (dividing by a standard labor share of 0.6) yields a one-off increase in aggregate output of 1.18%, implying $\tilde{Y}_\tau / Y_\tau = 1.0118$.

2. **Aghion and Bunel (2024)** project a larger one-off TFP increase of 7% over the same horizon, assuming AI affects 30% of tasks. The same capital deepening adjustment yields an 11.67% output increase, implying $\tilde{Y}_\tau / Y_\tau = 1.1167$.

Equation (12) then implies that the net span of control must decrease by $1 - (1.0118)^{-2} \approx 2.3\%$ under the projections of Acemoglu (2024) to completely dissipate the immediate gains within one generation. Under Aghion and Bunel's (2024) projections, the corresponding reduction is $1 - (1.1167)^{-2} \approx 19.8\%$.

Evaluating the plausibility of these reductions requires comparing them against the underlying task exposures of 5% and 30%. This mapping, however, is not exact. First, the exposure measures used by Acemoglu (2024) and Aghion and Bunel (2024) do not distinguish between task automation and augmentation. Second, while the model assumes a one-to-one mapping between routine tasks and the span of control, experts and novices in practice interact through complex bundles of responsibilities (Autor and Thompson, 2025; Garicano et al., 2026).

The quantitative question is therefore whether the net effect of automation and augmentation within these bundles shrinks effective mentorship capacity by the required 2.3% or 19.8%. Because these breakeven thresholds remain below the gross task exposure estimates in both projections, the intertemporal tradeoff identified in the model is a quantitatively relevant possibility.

An External Validity Check.—The derived break-even thresholds depend on projected AI adoption gains and the tail parameter θ of the skill distribution. While adoption gains are hypothetical, θ can be empirically disciplined. Standard calibrations set $\theta = 0.5$ to match the cross-sectional dispersion of earnings. To verify whether this imported parameter yields reasonable predictions in the present setting, I substitute empirical estimates of the span of control into the model's growth equation and compare the implied growth rates with historical U.S. data.

In the model, the generational growth factor of output is ℓ^θ. Letting T denote the number of years an individual spends in each stage—first as a novice, then as an expert—the annualized growth rate is $\ell^{\theta/T} - 1$. To evaluate this rate, I draw on Caicedo et al. (2019), setting $T = 22$ to match

their assumption of a 44-year career. For ℓ, I use their estimates of employment shares across U.S. organizational layers, which adapt the occupational classification of Caliendo et al. (2015) for French hierarchies. They document a ratio of workers to first-layer managers—an empirical counterpart to ℓ—of 1.56 in 1990 and 1.31 in 2010.

Given $\theta = 0.5$, these values imply annualized growth rates of 1.02% (for $\ell = 1.56$) and 0.62% (for $\ell = 1.31$). For an empirical benchmark, FRED data indicate that U.S. real GDP per capita grew from \$40,361 to \$70,279 chained 2017 dollars between the first quarter of 1990 and the fourth quarter of 2025, yielding a compound annual growth rate of 1.56%. Against this benchmark, the model accounts for 40% to 65% of aggregate growth. Evaluated against the more canonical 150-year historical average of 2% U.S. per-capita growth—the benchmark used by Lucas (2009), Lucas and Moll (2014), and Caicedo et al. (2019)—the model accounts for 31% to 51% of aggregate growth.

Under either metric, the mechanism captures a significant share of aggregate growth but leaves a fraction unexplained. Part of this missing growth likely reflects the fact that technological change is ongoing rather than one-off: machines keep improving, new technologies create new tasks, and new knowledge is continually generated. The baseline exercise abstracts from these additional sources of growth in order to isolate the diffusion margin.

In the Online Appendix, I formalize one such omitted margin by developing an extension in which growth stems from both the diffusion of tacit knowledge and the generation of new tacit ideas. This extension can account for the "missing growth."

The Threat of Stagnation.— In the baseline model, strictly positive asymptotic growth requires experts' span of control to exceed unity ($N - I > 1$). Should a technological shock drive this span below one, the intergenerational transmission of knowledge breaks down, and long-run growth disappears.

Additional back-of-the-envelope calculations yield a rough estimate of the shock strength required to induce stagnation. Given the empirical span of control estimated by Caicedo et al. (2019), driving the post-shock span to the stagnation boundary in the model without innovation ($\tilde{\ell} \leq 1$) requires a proportional reduction of $1 - 1/\ell$. This implies a 36% reduction from the 1990 estimate ($\ell = 1.56$) and a 24% reduction from the 2010 estimate ($\ell = 1.31$).

As shown in the Online Appendix, allowing growth to stem from both the diffusion of tacit knowledge and the generation of new tacit ideas makes the economy even more resilient to stagnation: reaching the stagnation boundary requires a 49.4% contraction relative to the 1990 estimate and a 39.7% contraction relative to the 2010 estimate. These calculations suggest that completely halting long-run growth requires an exceptionally severe disruption.

6.3 Further Considerations

Relaxing the Strict Division of Labor.— The baseline framework imposes a strict division of labor: experts exclusively apply tacit knowledge as orchestrators, while novices execute routine tasks. This assumption forces the least-skilled experts in any cohort into inactivity, yielding zero second-period income. Suppose instead that experts face an endogenous occupational choice between acting as or-

chestrators and supplying one unit of routine labor.[24] Although this flexibility alters microeconomic allocations, the core macroeconomic results remain unchanged.

Experts' endogenous occupational choice induces perfect sorting by age. Facing a zero continuation value, older low-skilled experts (who would otherwise be inactive) maximize static payoffs by supplying routine labor to the lowest-skilled orchestrators, who pay the highest upfront wages to compensate for inferior mentorship. Forward-looking novices, in contrast, value skill acquisition and outbid these older workers for elite orchestrators. The intergenerational transmission of the upper tail is thus preserved.

With novices still matching with top experts, aggregate responses to technological shocks parallel the baseline, operating through a new extensive margin. An automation shock shrinks elite orchestrators' span of control, displacing novices downward. To clear the market, the marginal experts who previously supplied routine labor transition to orchestrators. This entry of lower-skilled mentors dilutes knowledge transmission and permanently slows growth. Conversely, a task-creating technology expands the span of control and pulls novices upward. Unable to hire novices, the lowest-skilled orchestrators switch to routine labor. This improves the quality of the mentor pool and accelerates long-run growth.

More on Contract Incompleteness.— As discussed in Section 3.1, the zero lower bound on wages serves as a reduced-form proxy for the contract incompleteness inherent in the transfer of tacit knowledge. This formulation avoids taking a stance on the specific contracting friction and naturally generates the severe rationing of entry-level positions observed at elite professional services firms (Section 4.3).

At the same time, treating this bound as exogenous abstracts from the possibility that technological change alters the severity of the underlying contractual frictions themselves. A richer framework would endogenize the relationship between technology and contractual frictions. Whether this interaction attenuates or amplifies the forces emphasized here remains an important direction for future research.[25]

More on Knowledge Transmission.— The baseline model assumes that all routine tasks provide uniform insight into the expert's tacit knowledge. I abstract from task heterogeneity because its equilibrium implications are straightforward. Automating menial tasks that offer no meaningful interaction with the mentor increases production efficiency without hindering knowledge transmission—and may even enhance it via the scale effect discussed above.

The framework further assumes that novices perfectly acquire their mentor's skill. Given its uncodifiable nature, the transmission of tacit knowledge is probably noisy: apprentices may fall short of or surpass their mentor's expertise by drawing on their own insights. Moreover, changes in commu-

[24] This binary choice between occupations captures experts' limited time and attention.

[25] For a model in which technology interacts with the contractual environment governing training, see Garicano and Rayo (2025). In their setting, novices' inability to commit to employment post-training forces experts to transfer knowledge gradually. AI then affects apprenticeship viability through two competing channels: while automating entry-level tasks compresses the training window and limits the surplus experts can extract, a concurrent increase in the future value of expertise can offset this reduction by relaxing the novice's no-quit constraint.

nication technologies and other workplace trends—such as remote work—may endogenously alter the extent of this transmission. Extending the model along these lines may be a fruitful direction for future research.

7 Final Remarks

I conclude by summarizing the paper's key insights, outlining some qualifications, and discussing the policy implications that follow from the results.

The Main Takeaways.— The analysis rests on two premises. First, early-career work serves not only as an input into current production but also as the main channel through which tacit knowledge is transmitted across generations. Second, because tacit knowledge is embodied and non-verifiable, the training relationships through which it is acquired are contractually incomplete: novices cannot pledge future human capital to finance access to better mentors and must instead compensate experts through concurrent labor.

Together, these premises imply that technological change affects the intergenerational transmission of tacit knowledge by reshaping who works with whom and, therefore, who learns from whom. Because private decisions need not yield socially efficient training, technological improvements that displace novices from the most skilled experts weaken the diffusion of best practices and can reduce aggregate welfare.

Qualifications.—The analysis has focused on standard technological improvements—such as automation, task creation, and factor augmentation—that affect the creation and execution of tasks. AI, however, may differ from earlier technological changes in its ability to encode and apply tacit knowledge at scale (Ide and Talamàs, 2025). This raises the possibility that AI could democratize expertise either by *democratizing performance*—narrowing the productivity gap between experts—or by *democratizing learning*—facilitating novices' acquisition of tacit knowledge through exposure to a broad range of examples and on-demand interactive guidance.

A more radical possibility is that AI becomes sufficiently transformative to render some existing forms of tacit knowledge obsolete, fundamentally reshaping the skills required by future generations. In that case, productivity growth would depend on the development of new forms of tacit expertise that today's experts may not possess, including the ability to effectively prompt, direct, and evaluate AI systems. Even in that scenario, however, existing tacit knowledge is unlikely to become irrelevant. Human judgment, grounded in experience, would remain important for assessing when AI output is reliable and when it is not.

These possibilities may have important implications for the intergenerational transmission of tacit knowledge and the long-run trajectory of aggregate productivity. Although fully characterizing them lies beyond the scope of the present paper, the framework developed here provides a natural foundation for studying these forces.

Policy Implications.— The contractual incompleteness inherent in the transmission of tacit knowledge implies that technological improvements need not be socially efficient, creating scope for policy intervention. At the same time, the analysis shows that new technologies can set multiple, and often competing, forces in motion. Even entry-level automation can strengthen knowledge transmission if the resulting time savings allow top experts to scale their operations. The relevant policy question is therefore not whether technological progress should be resisted, but how its gains can be realized without undermining the intergenerational transmission of tacit knowledge.

The model provides two broad insights for approaching this question. First, the span of control of top experts is a useful diagnostic of a technology's likely effect on knowledge transmission. Second, once a technology is judged to weaken knowledge transmission, the relevant policy margins are identified by the source of the inefficiency. Policy should therefore aim to relax the frictions that prevent novices from efficiently accessing high-quality mentorship. This may justify subsidies for training-intensive entry-level arrangements, taxes on entry-level automation, or relaxing minimum-wage regulations for young workers.

APPENDIX

A Glossary of Symbols

Symbol	Definition	Symbol	Definition
Exogenous Primitives: Basic Parameters & Initial Conditions			
t	Discrete time index.	$f_0(q)$	Initial PDF of expert skills.
β	Discount factor.	x_0	Minimum of the support of F_0.
q	Absolute skill level.	θ	Tail thickness parameter.
$F_0(q)$	Initial CDF of expert skills.	τ	Date of technological shock.
Exogenous Primitives: Production & Technology			
N	Measure of tasks per project.	ρ	Machine rental rate.
I	Measure of automatable tasks.	Δ	Measure of newly automated or created tasks.
m	Machine productivity per task.	H	Novice productivity on new tasks.
h	Novice productivity per task.		
Endogenous Variables (Finite Time)			
$F_t(q)$	CDF of expert skills at t.	A_t	Set of active experts at t.
$f_t(q)$	PDF of expert skills at t.	a_t	Skill of the marginal active expert at t.
x_t	Minimum of the support of F_t.	$\mathcal{L}_t(E)$	Mass of novices asigned to expert set E at t.
k	Measure of tasks assigned to machines.	U_t	Baseline novice utility at t.
$S(q)$	Surplus of expert q before labor costs.	Y_t	Aggregate output at t.
$w_t(q)$	Wage schedule at t.	W_t	Aggregate welfare at t.
$\Pi_t(q)$	Expert income schedule at t.		
Endogenous Variables (Normalized Asymptotics)			
g	Asymptotic growth rate of output.	$\omega(z)$	Asymptotic normalized wage schedule.
z	Normalized skill level.	σ	Skill threshold where zero lower bound binds.
$\Phi(z)$	Asymptotic CDF of normalized skills.	$\gamma(z)$	Asymptotic value factor for dynasty z.
u	Asymptotic normalized baseline utility.	$n(z)$	Periods until obsolescence of dynasty z.
$\pi(z)$	Asymptotic normalized income schedule.		
Extensions (Section 6)			
μ	Measure of concurrent projects.	ℓ	Total span of control (scale effects included).
$\mathcal{M}(I,m,h)$	Project management time.	T	Years an individual spends in each career stage.
η	Complexity of task supervision.		

B Derivations and Omitted Proofs for Section 3

B.1 More on the Production Function

In this appendix, I show that the production function of the baseline model is a special case of a more general formulation. In this general formulation, an expert with skill q produces final output $y(q)$ by combining her skill with a continuum of routine tasks $s \in [0, N]$, aggregated via a Cobb–Douglas technology:

$$\ln y(q) = \ln q + \int_0^N \ln \mathcal{y}(s)\, ds$$

where $\mathcal{y}(s)$ denotes the output of task s. Tasks can be performed by either machines or novice labor according to the following technology:

$$\mathcal{y}(s) = \begin{cases} \psi_h(s)\chi(s) + \psi_m(s)[1 - \chi(s)] & \text{if } s \in [0, I] \\ \psi_h(s)\chi(s) & \text{if } s \in (I, N] \end{cases}$$

where $\psi_h(s)$ and $\psi_m(s)$ denote the productivity of novices and machines in task s, respectively, and $\chi(s) \in \{0, 1\}$ indicates whether task s is performed by a novice ($\chi(s) = 1$) or a machine ($\chi(s) = 0$).

The baseline specification corresponds to the special case where $\psi_h(s) = h$ for all $s \in [0, N]$ and $\psi_m(s) = m$ for all $s \in [0, I]$. Under these assumptions, total output depends only on the measure $k \in [0, I]$ of tasks assigned to machines. Substituting these constant productivities into the aggregator and exponentiating yields:

$$y(q) = q m^k h^{N-k}, \text{ where } k \in [0, I]$$

The fundamental difference between this specification and standard task-based models (e.g. Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2018, 2019) is that in the present setting, tasks are non-scalable. Formally, this restricts $\chi(s)$ to be a binary indicator. In contrast, standard models allow for continuous input choices (the intensive margin), replacing the discrete allocation $\chi(s)$ with variable quantities of labor and capital. As discussed in Sections 3.1 and 6.1, the assumption of non-scalability reflects experts' limited time and attention.

B.2 Proof of Lemma 1

This appendix shows that even with free novice labor, full automation ($k = I$) is optimal for all experts with:

$$q \geq \frac{\rho I}{h^N[(m/h)^I - 1]}$$

An active expert chooses $k \in [0, I]$ to maximize:

$$\max_{k \in [0,I]} \left\{ q m^k h^{N-k} - \rho k - w_t(q)(N - k) \right\}$$

Since $m > h$, the objective function is strictly convex in k, implying the optimum is a corner solution $k^* \in \{0, I\}$. Full automation ($k = I$) dominates no automation ($k = 0$) if and only if:

$$q(m^I h^{N-I} - h^N) \geq I(\rho - w_t(q))$$

Since $w_t(q) \geq 0$, the condition for full automation is most restrictive when labor is free. Setting $w_t(q) = 0$ and solving for q in the above inequality immediately gives the desired result. □

C Derivations and Omitted Proofs for Section 4

C.1 Proof of Proposition 2

Case 1: $1 - F_0 \in \text{RV}_{-1/\theta}$**.** I first show that $\lim_{t\to\infty}(a_{t+1}/a_t) = (N-I)^\theta$. Define $\xi_t \equiv (N-I)^{t+1}$ and $J(q) \equiv 1/(1-F_0(q))$. According to (9), $J(a_t) = \xi_t$ and $J(a_{t+1}) = (N-I)\xi_t$. Because F_0 admits a continuous and strictly positive PDF, J is strictly increasing, ensuring its inverse J^{-1} is well-defined. Therefore, $a_t = J^{-1}(\xi_t)$ and $a_{t+1} = J^{-1}((N-I)\xi_t)$, yielding:

$$\frac{a_{t+1}}{a_t} = \frac{J^{-1}((N-I)\xi_t)}{J^{-1}(\xi_t)}$$

Since $1 - F_0 \in \text{RV}_{-1/\theta}$, I have $J \in \text{RV}_{1/\theta}$. Because $1/\theta > 0$, an application of Bingham et al. (1989, Theorem 1.5.12) yields $J^{-1} \in \text{RV}_\theta$. Finally, since $\xi_t \to \infty$ as $t \to \infty$, it immediately follows that:

$$\lim_{t\to\infty}\left(\frac{a_{t+1}}{a_t}\right) = \lim_{\xi_t\to\infty}\left(\frac{J^{-1}((N-I)\xi_t)}{J^{-1}(\xi_t)}\right) = (N-I)^\theta$$

I now show that $\lim_{t\to\infty}(Y_{t+1}/Y_t) = (N-I)^\theta$. Let $r(a) \equiv \mathbb{E}_0[q \mid q \geq a]/a$. As shown in the main text, Y_{t+1}/Y_t can be decomposed as follows:

$$\frac{Y_{t+1}}{Y_t} = \left(\frac{a_{t+1}r(a_{t+1})}{a_t r(a_t)}\right).$$

By (10), $\lim_{a\to\infty} r(a) = (1-\theta)^{-1}$. Because the sequence of equilibrium cutoffs $\{a_t\}$ is strictly increasing and diverges to infinity as $t \to \infty$, the sequential characterization of limits directly guarantees that the sequences $\{r(a_t)\}$ and $\{r(a_{t+1})\}$ converge to the exact same limit $(1-\theta)^{-1}$ as $t \to \infty$. Applying this to the decomposed output ratio, and using the fact that $\lim_{t\to\infty}(a_{t+1}/a_t) = (N-I)^\theta$, yields the result.

Case 2: $1 - F_0 \in \text{RV}_{-\infty}$**.** I first show that $\lim_{t\to\infty}(a_{t+1}/a_t) = 1$. Let $\lambda_t \equiv a_{t+1}/a_t$. Since the sequence of equilibrium cutoffs $\{a_t\}$ is strictly increasing, $\lambda_t > 1$ for all t. Suppose, by way of contradiction, that $\limsup_{t\to\infty} \lambda_t > 1$. Then, there exists a subsequence $\{t_j\}_{j\geq 0}$ and some $\epsilon > 0$ such that $\lambda_{t_j} \geq 1+\epsilon$ for all j. This implies that $a_{t_j+1} \geq (1+\epsilon)a_{t_j}$. Because the survival function $1-F_0$ is strictly decreasing, it immediately follows that $1 - F_0(a_{t_j+1}) \leq 1 - F_0((1+\epsilon)a_{t_j})$. Dividing both sides by $1 - F_0(a_{t_j})$, and noting from (9) that the one-period forward ratio $(1 - F_0(a_{t_j+1}))/(1 - F_0(a_{t_j}))$ is identically equal to $1/(N-I)$, yields:

$$\frac{1}{N-I} \leq \frac{1 - F_0((1+\epsilon)a_{t_j})}{1 - F_0(a_{t_j})}.$$

By the definition of rapid variation, the right-hand side converges to zero as $j \to \infty$ because $a_{t_j} \to \infty$. However, the left-hand side is identically $1/(N-I) > 0$, yielding a contradiction. Thus, $\limsup_{t\to\infty} \lambda_t \leq 1$. Since $\lambda_t > 1$, it must be that $\lim_{t\to\infty} \lambda_t = 1$.

The proof that $\lim_{t\to\infty}(Y_{t+1}/Y_t) = 1$ is analogous to the regularly varying case, except that now $\lim_{a\to\infty} r(a) = 1$ (see (10)) and $\lim_{t\to\infty}(a_{t+1}/a_t) = 1$. □

C.2 Proof of Proposition 3

By construction, $\mathbb{P}_t\left(q/x_t > z\right) = 1 - F_t(zx_t)$, and due to (9), $1 - F_t(zx_t) = (N-I)^t[1 - F_0(zx_t)]$. Hence, $\mathbb{P}_t\left(q/x_t > z\right) = (N-I)^t[1 - F_0(zx_t)]$. Note that (9) also implies that $1 = (N-I)^t[1 - F_0(a_{t-1})]$, and since $a_{t-1} = x_t$, it follows that $(N-I)^t = [1 - F_0(x_t)]^{-1}$. Combining everything yields the relationship stated in Section 4.2:

$$1 - \Phi(z) \equiv \lim_{t\to\infty} \mathbb{P}_t\left(\frac{q}{x_t} > z\right) = \lim_{t\to\infty} \frac{1 - F_0(zx_t)}{1 - F_0(x_t)}$$

From the proof of Proposition 2, the sequence of entry cutoffs diverges, meaning the minimum skill of the active cohort $x_t = a_{t-1} \to \infty$ as $t \to \infty$. Because the initial survival function is regularly varying ($1 - F_0 \in \mathrm{RV}_{-1/\theta}$), the definition of regular variation directly implies that for any constant $z \geq 1$:

$$\lim_{t\to\infty}\left(\frac{1 - F_0(zx_t)}{1 - F_0(x_t)}\right) = \lim_{x_t\to\infty}\left(\frac{1 - F_0(zx_t)}{1 - F_0(x_t)}\right) = z^{-1/\theta}.$$

Substituting this limit into the survival function trivially implies $\Phi(z) = 1 - z^{-1/\theta}$. □

C.3 Characterization of the Asymptotic Income Schedule

This appendix characterizes in detail the asymptotic income and wage schedules, $\pi(z)$ and $\omega(z)$.

1. Detailed Derivation of the Functional Equation for $\pi(z)$. The absolute income of active experts satisfies the recursive formulation:

$$\Pi_t(q) = \min\left\{S(q), S(q) - (N-I)U_t + \beta(N-I)\Pi_{t+1}(q)\right\}, \quad \text{for } q \geq a_t$$

with the boundary condition $\Pi_t(q) = 0$ for all $q < a_t$. Evaluate this expression at $q = zx_t$, divide by x_t, and take the limit as $t \to \infty$. From Section 4.3, the normalized limits are $\lim_{t\to\infty} S(zx_t)/x_t = zm^I h^{N-I}$ and $\lim_{t\to\infty}(N-I)U_t/x_t = (1+g)m^I h^{N-I}$. Furthermore, the normalized entry cutoff converges to $\lim_{t\to\infty}(a_t/x_t) = 1 + g$, and the continuation to:

$$\lim_{t\to\infty}\frac{\Pi_{t+1}(zx_t)}{x_t} = \lim_{t\to\infty}\left(\frac{x_{t+1}}{x_t}\right)\frac{\Pi_{t+1}\left(z\frac{x_t}{x_{t+1}}x_{t+1}\right)}{x_{t+1}} = (1+g)\pi\left(\frac{z}{1+g}\right)$$

Substituting these limits and factoring out $m^I h^{N-I}$ yields the asymptotic functional equation:

$$\pi(z) = m^I h^{N-I}\min\left\{z, z - (1+g) + \frac{\beta(N-I)}{m^I h^{N-I}}(1+g)\pi\left(\frac{z}{1+g}\right)\right\}, \quad \text{for } z \geq 1+g.$$

Because the finite-time boundary condition dictates zero income for all inactive experts, the normalized schedule pastes continuously to the inactive region, establishing the asymptotic boundary condition $\pi(z) = 0$ for all $z \in [1, 1+g]$. □

2. Uniqueness of the Income Schedule. Global uniqueness is established by forward induction on the intervals $I_j \equiv [(1+g)^j, (1+g)^{j+1})$ for integers $j \geq 0$.

The boundary condition imposes $\pi(z) = 0$ for all $z \in I_0$. Assume $\pi(z)$ is uniquely determined on I_{j-1}. For any $z \in I_j$, the continuation skill $\frac{z}{1+g} \in I_{j-1}$. Because the functional equation determines $\pi(z)$ exclusively as a mapping of $\pi\left(\frac{z}{1+g}\right)$, uniqueness on I_{j-1} guarantees uniqueness on I_j. By induction, the schedule is uniquely determined for all $z \geq 1$. □

3. Characterization of the Solution. Since global uniqueness has been established, it suffices to verify that the following piecewise schedule satisfies the functional equation globally:

$$\pi(z) = m^I h^{N-I} \begin{cases} 0 & \text{if } z \in [1, 1+g) \\ \gamma(z) & \text{if } z \in [1+g, \sigma) \\ z & \text{if } z \in [\sigma, +\infty) \end{cases}$$

where $\gamma(z)$ is the finite sum defined in Section 4.3:

$$\gamma(z) = \sum_{j=0}^{n(z)-1} [\beta(N-I)]^j \left(z - (1+g)^{j+1}\right), \quad \text{with} \quad n(z) \equiv \left\lfloor \frac{\ln z}{\ln(1+g)} \right\rfloor$$

and $\sigma \in (1+g, +\infty)$ is the unique threshold satisfying $\sigma = \gamma(\sigma)$.

To do this, the verification proceeds in two steps. First, I establish the algebraic properties of $\gamma(z)$ and prove the existence of the unique threshold σ. Second, I show that the proposed schedule indeed solves the functional equation when $z \in [1+g, \sigma)$ and $z \geq \sigma$.

For the first step, note that $\gamma(z)$ is the unique solution to the following recursion:

$$\gamma(z) = z - (1+g) + \beta(N-I)(1+g)\gamma(z/(1+g)), \quad \text{for } z \geq 1+g$$

with the boundary condition $\gamma(z) = 0$ for $z \in [1, 1+g)$. Differentiating the recursion directly yields the slope $\gamma'(z) = 1 + \beta(N-I)\gamma'(z/(1+g))$.

Define $D(z) \equiv \gamma(z) - z$. Clearly, $D(1+g) = -(1+g) < 0$. Moreover, when $z \in [1+g, (1+g)^2)$, the continuation skill $z/(1+g)$ falls in the inactive region where $\gamma'(z/(1+g)) = 0$. This yields $\gamma'(z) = 1$, meaning $D'(z) = 0$, so $D(z)$ is constant and equal to $D(z) = -(1+g) < 0$. For any $z > (1+g)^2$, however, the continuation skill satisfies $z/(1+g) > 1+g$, implying $\gamma'(z/(1+g)) \geq 1$. Thus, $\gamma'(z) \geq 1 + \beta(N-I) > 1$, implying $D'(z) > 0$ for all $z > (1+g)^2$.

Because $D(z)$ is continuous, strictly negative on the initial interval, and strictly increasing to $+\infty$ thereafter, it admits a unique root $\sigma > (1+g)^2$. This guarantees $D(\sigma) = 0$, or equivalently, $\gamma(\sigma) = \sigma$. Consequently, $\gamma(z) < z$ for $z \in [1+g, \sigma)$, and $\gamma(z) \geq z$ for $z \geq \sigma$.

For the second step, I evaluate the right-hand side of the functional equation using the candidate $\pi(z)$ in the two active regions:

(a) For $z \in [1+g, \sigma)$: The continuation skill satisfies $z/(1+g) < \sigma$. The candidate therefore evaluates to $\pi(z/(1+g)) = m^I h^{N-I} \gamma(z/(1+g))$. Substituting this into the functional equation yields:

$$\pi(z) = m^I h^{N-I} \min\left\{z, \underbrace{z - (1+g) + \beta(N-I)(1+g)\gamma(z/(1+g))}_{=\gamma(z)}\right\}$$

Thus, $\pi(z) = m^I h^{N-I} \min\{z, \gamma(z)\}$. Because $\gamma(z) < z$ in this region, the $\min$ operator strictly selects the second argument, matching the candidate $\pi(z) = m^I h^{N-I}\gamma(z)$.

(b) For $z \geq \sigma$: To match the candidate $\pi(z) = m^I h^{N-I} z$, the $\min$ operator must strictly select the first argument, z. This requires:

$$z - (1+g) + \beta(N-I)(1+g)\frac{\pi(z/(1+g))}{m^I h^{N-I}} \geq z$$

or equivalently, $\beta(N-I)\frac{\pi(z/(1+g))}{m^I h^{N-I}} \geq 1$. Note then that $\sigma = \gamma(\sigma)$ if and only if $\beta(N-I)\gamma(\sigma/(1+g)) = 1$. Since $\sigma/(1+g) < \sigma$, the candidate schedule dictates $\frac{\pi(\sigma/(1+g))}{m^I h^{N-I}} = \gamma(\sigma/(1+g))$. Thus, the required condition binds with equality exactly at $z = \sigma$. Because the candidate schedule $\pi(z)$ is strictly increasing, it follows that $\beta(N-I)\frac{\pi(z/(1+g))}{m^I h^{N-I}} \geq 1$ for all $z \geq \sigma$. Thus, $\pi(z) = m^I h^{N-I} z$. □

D Derivations and Omitted Proofs for Section 5

D.1 Proof of Proposition 5

Let $\delta(\beta)$ denote the change in total welfare as a function of $\beta \in (0, \frac{1}{1+g})$:

$$\delta(\beta) \equiv \tilde{W}_\tau - W_\tau = \frac{\tilde{Y}_\tau}{1-\beta(1+\tilde{g})} - \frac{Y_\tau}{1-\beta(1+g)} - \frac{\rho\Delta N}{(N-I)(N-I-\Delta)(1-\beta)}$$

First, note that $\lim_{\beta\to 1/(1+g)} \delta(\beta) = -\infty$. Next, I establish that $\delta(0) > 0$, which immediately implies that $\delta(\beta)$ has at least one root in $\beta^* \in (0, \frac{1}{1+g})$.

Observe that the date-τ skill distribution is predetermined and the mass of active experts increases on impact ($\tilde{a}_\tau < a_\tau$). Thus, $\delta(0)$ can be written as:

$$\delta(0) = \int_{a_\tau}^{\infty} [\tilde{S}(q) - S(q)] dF_\tau(q) + \int_{\tilde{a}_\tau}^{a_\tau} \tilde{S}(q) dF_\tau(q)$$

where $S(q) \equiv q m^I h^{N-I} - \rho I$ and $\tilde{S}(q) \equiv q m^{I+\Delta} h^{N-I-\Delta} - \rho(I+\Delta)$. Per Assumption 2, it follows that $\tilde{S}(q) > S(q)$ for all $q \geq x_0$. Furthermore, Assumption 3 ensures $S(x_0) > 0$. Together, these conditions imply that $\tilde{S}(q) > S(q) > 0$ for all active experts ($q \geq \tilde{a}_\tau \geq x_0$). Since both integrands are strictly positive, it immediately follows that $\delta(0) > 0$.

To establish the uniqueness of β^*, note that for all $\beta \in \left(0, \frac{1}{1+g}\right)$, the roots of $\delta(\beta)$ coincide with those of $P(\beta) \equiv \delta(\beta)(1-\beta)\big[1-\beta(1+g)\big]\big[1-\beta(1+\tilde{g})\big]$. Expanding $P(\beta)$ yields:

$$P(\beta) = \tilde{Y}_\tau(1-\beta)\big[1-\beta(1+g)\big] - Y_\tau(1-\beta)\big[1-\beta(1+\tilde{g})\big] - C\big[1-\beta(1+g)\big]\big[1-\beta(1+\tilde{g})\big]$$

where $C \equiv \frac{\rho\Delta N}{(N-I)(N-I-\Delta)} > 0$. Because each of its three terms is the product of two linear functions of β, $P(\beta)$ is a polynomial of degree at most two. Evaluating $P(\beta)$ at the interval boundaries yields $P(0) = \delta(0) > 0$ and

$$P\left(\frac{1}{1+g}\right) = -Y_\tau\left(1-\frac{1}{1+g}\right)\left(1-\frac{1+\tilde{g}}{1+g}\right) < 0.$$

Because $P(0) > 0$ and $P\left(\frac{1}{1+g}\right) < 0$, the fact that $P(\beta)$ is a polynomial of degree at most two immediately implies that it crosses zero exactly once in $\left(0, \frac{1}{1+g}\right)$. Hence, β^* is unique. □

D.2 Full Adoption of a Task-Creating Technology

This appendix proves that under Assumption 5, all active experts strictly prefer to adopt the task-creating technology for all $t \geq \tau$. The proof proceeds in two steps. First, I show that the incentive to adopt the new technology is strictly increasing in expert skill, q. Second, I establish that under Assumption 5, adoption is optimal for the marginal active expert, a_t. Monotonicity then ensures that all active experts $q \geq a_t$ adopt the technology for every $t \geq \tau$.

Step 1. Suppose there is an equilibrium where all active experts adopt the new technology, and let $w_t(q)$ be the equilibrium wage schedule. In this equilibrium, dynamics mirror the baseline model, except that active experts hire $N - I + \Delta$ novices and output is scaled by the productivity premium of new tasks, $(H/m)^{\Delta}$.

Consider an active expert with skill $q \geq a_t$ at time t. Given $w_t(q)$, the optimal profits under the baseline and new task-creating technologies are, respectively:

$$\Pi_{old}(q) = qm^I h^{N-I} - \rho I - w_t(q)(N - I)$$
$$\Pi_{new}(q) = qm^{I-\Delta} h^{N-I} H^{\Delta} - \rho(I - \Delta) - w_t(q)(N - I + \Delta)$$

The net incentive to adopt the new technology is the profit difference:

$$\Pi_{new}(q) - \Pi_{old}(q) = qm^{I-\Delta} h^{N-I}(H^{\Delta} - m^{\Delta}) + \rho\Delta - w_t(q)\Delta$$

Because $H > m$, the first term is strictly increasing in q. By Proposition 1, the wage schedule $w_t(q)$ is non-increasing in q, implying the cost term $-w_t(q)\Delta$ is non-decreasing in q. Therefore, the net incentive to adopt is strictly increasing in q.

Step 2. I now establish that under Assumption 5, the marginal active expert a_t strictly prefers the new technology. Combined with the monotonicity established in Step 1, this guarantees that all active experts optimally adopt the task-creating technology for all $t \geq \tau$, confirming the conjectured equilibrium.

In this equilibrium, the marginal expert earns zero profit under the new technology, implying $\Pi_{new}(a_t) = 0$. This boundary condition pins down the wage paid by the marginal expert:

$$w_t(a_t) = \frac{a_t m^{I-\Delta} h^{N-I} H^{\Delta} - \rho(I - \Delta)}{N - I + \Delta}$$

Given $\Pi_{new}(a_t) = 0$, the marginal expert strictly prefers the new technology if and only if deviating to the baseline technology yields strictly negative profits, $\Pi_{old}(a_t) < 0$. Substituting $w_t(a_t)$ into the expression for $\Pi_{old}(a_t)$, multiplying by $N - I + \Delta$, and rearranging terms yields:

$$a_t m^I h^{N-I}(N - I)\left[1 + \frac{\Delta}{N - I} - \left(\frac{H}{m}\right)^{\Delta}\right] < \rho\Delta N$$

Under Assumption 5, the productivity premium of new tasks satisfies:

$$\left(\frac{H}{m}\right)^{\Delta} \geq 1 + \frac{\Delta}{N - I}$$

This condition guarantees that the bracketed term on the left-hand side is non-positive. Since the right-hand side, $\rho\Delta N$, is strictly positive, the strict inequality $\Pi_{old}(a_t) < 0$ always holds.

Conversely, when Assumption 5 does not hold, the left-hand side is strictly positive and proportional to a_t. Because sustained growth dictates that $a_t \to \infty$ as $t \to \infty$, the left-hand side grows without bound. Consequently, there necessarily exists a time $t \geq \tau$ at which $\Pi_{old}(a_t) > 0$. At this point, the least-skilled active experts strictly prefer the baseline technology, unraveling the full-adoption equilibrium. □

References

ACEMOGLU, D. (2024): "The Simple Macroeconomics of AI," *Economic Policy*, 40, 13–58.

ACEMOGLU, D. AND D. AUTOR (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, Elsevier, vol. 4, 1043–1171.

ACEMOGLU, D., D. KONG, AND A. OZDAGLAR (2026): "AI, Human Cognition and Knowledge Collapse," NBER Working Paper 34910.

ACEMOGLU, D., F. KONG, AND P. RESTREPO (2024): "Tasks At Work: Comparative Advantage, Technology and Labor Demand," NBER Working Paper No. 32872.

ACEMOGLU, D. AND J. LOEBBING (2024): "Automation and Polarization," Working Paper.

ACEMOGLU, D. AND P. RESTREPO (2018): "The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 108, 1488–1542.

——— (2019): "8. Artificial Intelligence, Automation, and Work," in *The Economics of Artificial Intelligence: An Agenda*, ed. by A. Agrawal, J. Gans, and A. Goldfarb, University of Chicago Press, 197–236.

——— (2022): "Tasks, Automation, and the Rise in U.S. Wage Inequality," *Econometrica*, 90, 1973–2016.

AGHION, P. AND S. BUNEL (2024): "AI and Growth: Where Do We Stand," LSE Working Paper.

AGHION, P., B. F. JONES, AND C. I. JONES (2017): "Artificial Intelligence and Economic Growth," NBER Working Paper No. 23928.

ALTMAN, S. (2025): "The Gentle Singularity," June, 2025. Available at https://blog.samaltman.com/the-gentle-singularity (accessed June 15, 2025).

AMODEI, D. (2024): "Machines of Loving Grace: How AI Could Transform the World for the Better," October, 2024. Available at https://darioamodei.com/machines-of-loving-grace (accessed June 15, 2025).

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

ATKINSON, T. AND S. YAMCO (2026): "Young Workers' Employment Drops in Occupations with High AI Exposure," January 6, 2026. Available at https://www.dallasfed.org/research/economics/2026/0106 (accessed March 2, 2026).

AUTOR, D. (2024): "Applying AI to Rebuild Middle Class Jobs," Tech. rep., National Bureau of Economic Research.

AUTOR, D., C. CHIN, A. SALOMONS, AND B. SEEGMILLER (2024): "New Frontiers: The Origins and Content of New Work, 1940–2018," *The Quarterly Journal of Economics*, 139, 1399–1465.

AUTOR, D. AND N. THOMPSON (2025): "Expertise," *Journal of the European Economic Association*, 23, 1203–1271.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration*," *The Quarterly Journal of Economics*, 118, 1279–1333.

BEANE, M. (2019): "Shadow Learning: Building Robotic Surgical Skill When Approved Means Fail," *Administrative Science Quarterly*, 64, 87–123.

——— (2024a): "How AI Could Keep Young Workers From Getting the Skills They Need," *The Wall Street Journal*, July 26, 2024. Available at https://www.wsj.com/lifestyle/careers/ai-training-young-employees-ed08cedc (accessed July 20, 2025).

——— (2024b): *The Skill Code: How to Save Human Ability in an Age of Intelligent Machines*, New York: Harper Business.

BECKER, G. S. (1964): *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, University of Chicago press.

——— (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81, 813–846.

BENHABIB, J., J. PERLA, AND C. TONETTI (2021): "Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier," *Econometrica*, 89, 2261–2301.

BERAJA, M. AND E. TALAMÀS (2026): "The Value of Organizational Learning Technologies," Working Paper.

BERGER, P. G., W. CAI, L. QIU, AND C. XINYI SHEN (2024): "Employer and Employee Responses to Generative AI: Early Evidence," Available at SSRN 4874061.

BINGHAM, N. H., C. M. GOLDIE, AND J. L. TEUGELS (1989): *Regular Variation*, vol. 27 of *Encyclopedia of Mathematics and its Applications*, Cambridge: Cambridge University Press, first paperback edition (with additions) ed.

BLOOMBERGTV (2025): "AI Boom, Entry-Level Bust: Why College Grads Are Struggling to Land Jobs," August, 9, 2025. Available at https://www.bloomberg.com/news/videos/2025-08-09/ai-boom-entry-level-bust-grads-struggle-to-land-jobs-video (accessed August 11, 2025).

BROWN, J. S. AND P. DUGUID (1991): "Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation," *Organization Science*, 2, 40–57.

BRYNJOLFSSON, E., B. CHANDAR, AND R. CHEN (2025a): "Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence," Working Paper.

BRYNJOLFSSON, E., D. LI, AND L. RAYMOND (2025b): "Generative AI at Work," *The Quarterly Journal of Economics*, 140, 889–942.

BRYNJOLFSSON, E. AND T. MITCHELL (2017): "What Can Machine Learning Do? Workforce Implications," *Science*, 358, 1530–1534.

BUERA, F. J. AND R. E. LUCAS (2018): "Idea Flows and Economic Growth," *Annual Review of Economics*, 10, 315–345.

BUERA, F. J. AND E. OBERFIELD (2020): "The Global Diffusion of Ideas," *Econometrica*, 88, 83–114.

CAICEDO, S., J. LUCAS, ROBERT E., AND E. ROSSI-HANSBERG (2019): "Learning, Career Paths, and the Distribution of Wages," *American Economic Journal: Macroeconomics*, 11, 49–88.

CALIENDO, L., F. MONTE, AND E. ROSSI-HANSBERG (2015): "The Anatomy of French Production Hierarchies," *Journal of Political Economy*, 123, 809–852.

CLARKE, P. (2025): "Record Number of Graduates Apply to be Wall Street Bankers," *Financial News*, September 1, 2025. Available at https://www.fnlondon.com/articles/record-number-of-graduates-apply-to-be-wall-street-bankers-0ac689a8 (accessed February 22, 2026).

DE LA CROIX, D., M. DOEPKE, AND J. MOKYR (2018): "Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Preindustrial Economy," *The Quarterly Journal of Economics*, 133, 1–70.

DE SOUZA, G. (2025): "Artificial Intelligence in the Office and the Factory: Evidence from Administrative Software Registry Data," FRB of Chicago Working Paper No. 2025-11.

DEMIRER, M., J. J. HORTON, N. IMMORLICA, B. LUCIER, AND P. SHAHIDI (2026): "Chaining Tasks, Redefining Work: A Theory of AI Automation," NBER Working Paper 34859.

EATON, J. AND S. KORTUM (1999): "International Technology Diffusion: Theory and Measurement," *International Economic Review*, 40, 537–570.

——— (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

EECKHOUT, J. AND P. KIRCHER (2011): "Identifying Sorting—In tTheory," *The Review of Economic Studies*, 78, 872–906.

ELLIS, L. (2026): "Even MBAs From Top Business Schools Are Struggling to Get Hired," *The Wall Street Journal*, January 19, 2026. Available at https://www.wsj.com/lifestyle/careers/even-mbas-from-top-business-schools-are-struggling-to-get-hired-11f4a167 (accessed February 26, 2026).

EMBRECHTS, P., C. KLÖPPELBERG, AND T. MIKOSCH (1997): *Modelling Extremal Events*, vol. 33 of *Applications of Mathematics (New York)*, Berlin: Springer-Verlag, for Insurance and Finance.

FORAY, D. (2004): *Economics of Knowledge*, MIT press.

FREUND, L. B. AND L. F. MANN (2025): "Job Transformation, Specialization, and the Labor Market Effects of AI," CESifo Working Paper No. 12072.

FUDENBERG, D., G. GEORGIADIS, AND L. RAYO (2021): "Working to Learn," *Journal of Economic Theory*, 197, 105347.

FUDENBERG, D. AND L. RAYO (2019): "Training and Effort Dynamics in Apprenticeship," *American Economic Review*, 109, 3780–3812.

GANS, J. S. (2025): "Growth in AI Knowledge," NBER Working Paper No. 33907.

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

——— (2025): "The AI Becker Problem," July, 23, 2025. Available at https://www.siliconcontinent.com/p/the-ai-becker-problem (accessed July 23, 2025).

GARICANO, L., J. LI, AND Y. WU (2026): "Weak Bundle, Strong Bundle: How AI Redraws Job Boundaries," Working Paper.

GARICANO, L. AND L. RAYO (2017): "Relational Knowledge Transfers," *American Economic Review*, 107, 2695–2730.

——— (2025): "Training in the Age of AI: A Theory of Apprenticeship Viability," Working Paper.

GARICANO, L. AND E. ROSSI-HANSBERG (2004): "Inequality and the Organization of Knowledge," *American Economic Review Papers and Proceedings*, 94, 197–202.

——— (2006): "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, 121, 1383–1435.

——— (2015): "Knowledge-Based Hierarchies: Using Organizations to Understand the Economy," *Annual Review of Economics*, 7, 1–30.

GLICKMAN, B. AND A. SAEEDY (2025): "110-Hour Workweeks Drove Young Bankers at a Boutique Firm to the Brink," *The Wall Street Journal*, April 30, 2025. Available at https://www.wsj.com/finance/banking/banking-culture-robert-w-baird-aeccc947 (accessed February 25, 2026).

HOSSEINI, S. AND G. LICHTINGER (2025): "Generative AI as Seniority-Biased Technological Change: Evidence from U.S. Résumé and Job Posting Data," Available at SSRN 5425555.

HUMPHRIES, J. (2006): "English Apprenticeship: A Neglected Factor in the First Industrial Revolution," in *The Economic Future in Historical Perspective*, British Academy.

IDE, E. AND E. TALAMÀS (2024): "The Turing Valley: How AI Capabilities Shape Labor Income," ArXiv preprint arXiv:2408.16443.

——— (2025): "Artificial Intelligence in the Knowledge Economy," *Journal of Political Economy*, 133, 3762–3800.

——— (2026): "The Impact of AI on Global Knowledge Work," *Journal of Monetary Economics*, 157, 103876.

JAROSCH, G., E. OBERFIELD, AND E. ROSSI-HANSBERG (2021): "Learning From Coworkers," *Econometrica*, 89, 647–676.

JOHNSTON, A. AND C. MAKRIDIS (2025): "The Labor Market Effects of Generative AI: A Difference-in-Differences Analysis of AI Exposure," Available at SSRN 5375017.

JONES, C. I. AND C. TONETTI (2026): "Past Automation and Future AI: How Weak Links Tame the Growth Explosion," Working paper.

KLEIN TEESELINK, B. (2025): "Generative AI and labor market outcomes: Evidence from the United Kingdom," Available at SSRN 5516798.

KORINEK, A. AND D. SUH (2024): "Scenarios for the Transition to AGI," NBER Working Paper No. 32255.

KORTUM, S. S. (1997): "Research, Patenting, and Technological change," *Econometrica*, 1389–1419.

KOSMYNA, N., E. HAUPTMANN, Y. T. YUAN, J. SITU, X.-H. LIAO, A. V. BERESNITZKY, I. BRAUNSTEIN, AND P. MAES (2025): "Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task," ArXiv preprint arXiv:2506.08872.

LAHART, J. (2026): "Economists Are Studying the Slowing Job Market—and Feeling It Themselves," *The Wall Street Journal*, January 18, 2026. Available at https://www.wsj.com/economy/jobs/economists-job-market-hiring-2213807b (accessed February 26, 2026).

LAVE, J. AND E. WENGER (1991): *Situated Learning: Legitimate Peripheral Participation*, Cambridge University Press.

LEVY, S. (2025): "Demis Hassabis Embraces the Future of Work in the Age of AI," *Wired*, June 4, 2025. Available at https://www.wired.com/story/google-deepminds-ceo-demis-hassabis-thinks-ai-will-make-humans-less-selfish/ (accessed June 15, 2025).

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

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

MOKYR, J. (2002): *The Gifts of Athena: Historical Origins of the Knowledge Economy*, Princeton, NJ: Princeton University Press.

——— (2019): *The Economics of Apprenticeship*, Cambridge University Press, 20–43.

MOLL, B., L. RACHEL, AND P. RESTREPO (2022): "Uneven Growth: Automation's Impact on Income and Wealth Inequality," *Econometrica*, 90, 2645–2683.

OYER, P. (2006): "Initial Labor Market Conditions and Long-Term Outcomes for Economists," *Journal of Economic Perspectives*, 20, 143–160.

PERLA, J. AND C. TONETTI (2014): "Equilibrium Imitation and Growth," *Journal of Political Economy*, 122, 52–76.

PERLA, J., C. TONETTI, AND M. E. WAUGH (2021): "Equilibrium Technology Diffusion, Trade, and Growth," *American Economic Review*, 111, 73–128.

POLANYI, M. (1962): *Personal Knowledge: Towards a Post-Critical Philosophy*, London: Routledge, first published 1958. 2005 Taylor & Francis e-Library edition.

——— (1966): *The Tacit Dimension*, Doubleday & Co.

ROOSE, K. (2025): "For Some Recent Graduates, the A.I. Job Apocalypse May Already Be Here," *The New York Times*, May 30, 2025. Available at https://www.nytimes.com/2025/05/30/technology/ai-jobs-college-graduates.html (accessed August 11, 2025).

SAEEDY, A. (2024a): "Bank of America Urges Bankers to Sound Alarm on Overwork After WSJ Investigation," *The Wall Street Journal*, August 13, 2024. Available at https://www.wsj.com/finance/bank-of-america-tells-bankers-to-report-bosses-who-pressure-them-to-underreport-hours-758cb44b (accessed February 25, 2026).

——— (2024b): "How Bank of America Ignores Its Own Rules Meant to Prevent Dangerous Workloads," *The Wall Street Journal*, August 11, 2024. Available at https://www.wsj.com/finance/banking/bank-of-america-worker-death-policies-89eff5f6 (accessed February 25, 2026).

STIGLER, S. M. (1977): "Fractional Order Statistics, With Applications," *Journal of the American Statistical Association*, 72, 544–550.

THE JOURNAL PODCAST (2025): "AI Is Coming for Entry-Level Jobs," *The Journal - WSJ Podcast*, Monday, July 7, 2025. Available at https://www.wsj.com/podcasts/the-journal/ai-is-coming-for-entry-level-jobs/e9f9eb31-14ad-498d-94f3-11ce91e9c464 (accessed July 19, 2025).

VLAMIS, K., L. VARANASI, AND T. PARADIS (2025): "MBB Explained: How Hard it is to Get Hired and What it's Like to Work for the Prestigious Strategy Consulting Firms, McKinsey, Bain, and BCG," *Business Insider*, January 29, 2025. Available at https://www.businessinsider.com/mbb-explained-mckinsey-bain-bcg-compare-big-three-strategy-consultants-2024-11 (accessed February 22, 2026).

ZEIRA, J. (1998): "Workers, Machines, and Economic Growth," *The Quarterly Journal of Economics*, 113, 1091–1117.